فصلنامه راهبرد مدیریت مالی
دانشگاه الزهرا
سال دوازدهم، شماره چهل و چهارم، بهار ۱۴۰۳
صفحات ۱۹۰-۱۶۱

مقاله پژوهشی

مقایسه عملکرد سیستم های معاملات الگوریتمی مبتنی بر یادگیری ماشین در بازار رمز ارزها[۱]

عماد کوشا[۲]، محسن صیقلی[۳]، ابراهیم عباسی[۴]

تاریخ دریافت: ۱۴۰۱/۰۶/۲۷ تاریخ پذیرش: ۱۴۰۲/۰۱/۲۳

چکیده

هدف از این پژوهش استفاده از مدل یادگیری جمعی برای ترکیب پیش‌بینی های مدل‌های جنگل تصادفی، حافظه طولانی کوتاه‌مدت و شبکه عصبی بازگشتی جهت ارائه یک سیستم معاملاتی الگوریتمی مبتنی بر آن می‌باشد. در این پژوهش یک مدل پیش‌بینی مبتنی بر مدل یادگیری ماشین جمعی ارائه شده است و عملکرد آن با هر یک از زیر الگوریتم ها و داده‌های واقعی مقایسه می‌شود. در این پژوهش در مرحله اول با استفاده از سه مدل یادگیری ماشین، سقف و کف قیمت بیت‌کوین پیش‌بینی شده است. در مرحله دوم، خروجی‌های مدل‌ها به عنوان متغیرهای ویژگی به مدل‌های XGboost و LightGBM جهت پیش‌بینی سقف و کف‌ها ارائه شده است. سپس در مرحله سوم خروجی‌های مرحله دوم، با الگوی دسته بندی رای گیری جمعی برای پیش‌بینی سقف و کف بعدی، ترکیب می‌شوند. داده‌های سقف و کف قیمت بیت‌کوین در تایم فریم ۱ ساعته از تاریخ ۲۰۱۸/۱/۱ الی آخر ۲۰۲۲/۶/۳۰ به عنوان متغیر هدف و ۳۱ اندیکاتور تحلیل تکنیکال به عنوان متغیر ویژگی برای سه مدل در مرحله اول استفاده شده‌اند. در نهایت مقادیر پیش‌بینی و سیستم های معاملاتی الگوریتمی با داده‌های واقعی برای ۳ مدل و مدل یادگیری جمعی معرفی شده مورد ارزیابی و مقایسه قرار گرفتند. نتایج به دست آمده نشان‌دهنده ارتقا عملکرد دقت و صحت مدل یادگیری جمعی پیشنهاد شده در پیش‌بینی سقف و کف بیت‌کوین و همچنین، عملکرد بهتر آن نسبت به زیر الگوریتم‌ها می‌باشد.

واژگان کلیدی: معاملات الگوریتمی، پیش‌بینی سقف و کف قیمت، یادگیری ماشین جمعی، XGBoost، LightGBM

طبقه‌بندی موضوعی: *G17,B17,C53, F17,F19*

---

۱. کد DOI مقاله: 10.22051/JFM.2024.41815.2742

۲. دانشجوی دکتری، گروه مدیریت مالی، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران. E mail: emadkoosha92@gmail.com

۳. استادیار، گروه مدیریت، دانشکده مدیریت، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران. Email: seighaly@gmail.com

۴. استاد، گروه مدیریت، دانشکده علوم اجتماعی و اقتصادی، دانشگاه الزهرا، تهران، ایران. Email: abbasiebrahim2000@Alzahra.ac.ir

مقدمه

پیش‌بینی قیمت در بازارهای مالی یکی از مهم‌ترین عوامل موفقیت برای معامله گران و سرمایه‌گذاران می‌باشد. افزایش دقت در پیش‌بینی قیمت می‌تواند در موارد زیر به فعالین بازارهای مالی کمک کند (جیانک و همکاران، ۲۰۱۹):

۱- افزایش کارایی استراتژی‌های معامله‌گری

۲- سرمایه‌گذاران می‌توانند بر اساس نتایج پیش‌بینی به دنبال پوشش ریسک خود باشند.

۳- سفته‌بازان و آربیتراژگران می‌توانند از نتایج پیش‌بینی خود بازدهی خود را افزایش و ریسک را کاهش دهند.

۴- می‌توانند با پیش‌بینی شاخص‌های مهم در مورد بازار های مالی به‌طورکلی، اظهار نظر نمایند.

بنابراین، پیش‌بینی قیمت برای فعالین بازارهای مالی، اهمیت و ضرورت ویژه‌ای دارد. در بین بازارهای مالی، بازار رمز ارزها[1] با توجه به رشد خیره‌کننده آن‌ها در سال‌های ۲۰۱۵ تا ۲۰۱۷ و ۲۰۱۹ تا ۲۰۲۱ توانسته نظر سرمایه‌گذاران و فعالین را به خود جلب کند. بیت‌کوین به عنوان بزرگترین رمز ارز، یک پول دیجیتالی است که توسط ناکاموتو در سال ۲۰۰۸ ارائه شد. این رمز ارز از فناوری بلاکچین[2] و رمزنگاری[3] استفاده کرده است تا افراد بتوانند همتا به همتا به صورت امن تراکنش مالی انجام دهند (ناکاماتو، ۲۰۰۸). از دلایل مهم رشد بیت‌کوین در سال‌های ۲۰۱۹ تا ۲۰۲۱، می‌توان به اثر هاوینگ[4] (نصف شدن جایزه شبکه بیت‌کوین) در سال ۲۰۲۰، پاندامی کرونا و سیاست‌های انبساطی فدرال رزرو و همچنین حمایت شرکت‌ها و سازمآن‌های بزرگ در دنیا اشاره کرد. پیش‌بینی قیمت بیت‌کوین به عنوان بزرگترین رمز ارز در این بازار می‌تواند به شکل نماینده عمل کند چرا که همبستگی سایر رمز ارزها (آلتکوین) با بیت‌کوین بالاست. در نتیجه پیش‌بینی قیمت بیت‌کوین می‌تواند برای فعالین در بازارهای مالی بین‌المللی با هدف سپر تورمی، را می‌توان اهمیت و ضرورت این پژوهش دانست.

در این پژوهش، به دنبال پیش‌بینی قیمت بیت‌کوین به عنوان نماینده بازار رمز ارزها با استفاده از مدل‌های یادگیری ماشین و سپس، ارائه سیستم معاملاتی مبتنی بر آن هستیم. گوپتا و نیین[5] (۲۰۲۱) در پژوهشی نشان دادند که مدل‌های سری زمانی به طور خاص مدل میانگین متحرک وزنی[6] و ARIMA در پیش‌بینی قیمت بیت‌کوین نسبت به مدل‌های یادگیری ماشین یا یادگیری ماشین عمیق عملکرد بهتری در پیش قیمت یک روز دارند. همچنین، توانستند در بهترین حالت ۸۶ درصد دقت را ارئه کنند. البته پژوهش‌هایی نشان دادند مدل‌های یادگیری ماشین، با توجه به غلبه بر محدودیت‌های خطی مدل‌های کلاسیک توانسته‌اند دقت پیش‌بینی را افزایش دهند (فقیهی‌نژاد و مینایی، ۱۳۹۷). این پژوهش، به جای

1. Cryptocurrencies
2. Blockchain
3. Cryptography
4. Halving
5. Gupta and Nain
6. Weighted Moving Average

رویکرد سری زمانی و پیش‌بینی یک داده بعد به دنبال پیش کف و سقف بعدی برای بیت‌کوین در تایم فریم ۱ ساعته است. با رویکرد گسسته و دسته‌بندی سقف یا کف می‌توان دقت مدل‌های یادگیری ماشین را در پیش‌بینی ارتقا داد.

بنابراین، هدف این پژوهش پیش‌بینی سقف‌ها و کف‌های بیت‌کوین با استفاده از مدل‌های یادگیری ماشین نظارت شده می‌باشد. مدل پیشنهادی در این پژوهش سه مرحله اصلی دارد. در مرحله اول، سقف‌وکف‌های قیمت بیت‌کوین با استفاده از مدل‌های جنگل تصادفی[۱]، حافظه طولانی کوتاه‌مدت[۲] و شبکه عصبی بازگشتی[۳] پیش‌بینی می‌شوند. در مرحله دوم، نتیجه حاصله به عنوان ورودی به مدل‌های XGBoost[۴] و LightGBM[۵] ارائه می‌شود و نتایج این دو مدل با استفاده از یادگیری جمعی[۶] رای گیری[۷] ترکیب شده و نتیجه پیش‌بینی نهایی قابل مقایسه با داده واقعی می‌باشد. در گام آخر، بر اساس هر یک از مدل‌های RNN، LSTM، RF و یادگیری جمعی پیشنهادی یک سیستم معاملاتی برای بازه زمانی داده‌های تست اجرا و ارزیابی می‌شوند. عملکرد این سیستم های معاملاتی به طور مجزا بررسی و با یکدیگر مقایسه می‌شوند. این مراحل به صورت دقیق در بخش های بعدی پژوهش توضیح داده شده است.

مبانی نظری و مروری بر پیشینه پژوهش:

پیش‌بینی قیمت دارایی‌های مالی، یکی از موضوعات مهم برای فعالین، سرمایه گذاران، معامله گران و آربیتراژگران در بازارهای مالی است. به‌طورکلی، روش‌های پیش‌بینی قیمت در بازارهای مالی را می‌توان به دو دسته اصلی کلاسیک و هوشمند تقسیم‌بندی کرد. در روش‌های کلاسیک فرض بر این است مقادیر آینده با متغیرهای توضیح‌دهنده که می‌تواند عوامل موثر بر قیمت یا خود قیمت با وقفه زمانی باشد پیش‌بینی می‌شوند. در این مدل‌های نیاز است رابطه میان متغیرهای توضیح‌دهنده و وابسته مشخص و با استفاده از پیش آزمون‌ها داده‌ها را با فروض مدل‌ها تطبیق داده تا نتایج مطلوب دست پیدا کند. در مقابل روش‌های هوشمند هستند که از یادگیری ماشین استفاده می‌کنند. در این رویکرد، رابطه از قبل مشخص نیست و این الگوریتم ها به دنبال کشف یک رابطه ریاضی هستند.

یادگیری ماشین حوزه مطالعاتی است که به رایانه‌ها توانایی یادگیری بدون برنامه‌ریزی آشکار روی می‌دهد (ساموئل، ۱۹۵۹). در حال حاضر، یادگیری ماشین به سه دسته اصلی؛ یادگیری نظارت شده[۸]، نظارت نشده[۹] و تقویتی[۱۰] تقسیم‌بندی می‌شود. یادگیری ماشین نظارت شده، به دنبال ارتباط بین مقادیر

1. Random Forest (RF)
2. Long-Short Term Memory (LSTM)
3. Recurrent Neural Network (RNN)
4. Extreme Gradient Boosting (XGBoost)
5. Light Gradient Boosting Machine (LightGBM)
6. Ensemble learning
7. Voting ensemble classifier
8. Supervised Learning
9. Unsupervised learning
10. Reinforcement learning

متغیرهای ویژگی با متغیر هدف است. در نتیجه، از این مدل‌ها بیشتر در مسائل پیش‌بینی استفاده می‌شود. دو نوع اصلی یادگیری ماشین نظارت‌شده؛ طبقه‌بندی و رگرسیون مطرح می‌باشد. با توجه با اینکه پژوهش حاضر به دنبال پیش‌بینی بازار با رویکرد گسسته است، از الگوریتم‌های یادگیری ماشین طبقه‌بندی استفاده شده است. می‌توان گفت با توجه به اینکه در این پژوهش هدف ۰ (در مسیر کف بازار) و ۱ (در مسیر سقف) بازار است، یک بردار هدف دو دویی وجود دارد که با استفاده از مدل‌های RF، RNN و LSTM پیش‌بینی می‌شود. در مورد نحوه پیش‌بینی، در قسمت روش پژوهش به طور کامل توضیح داده شده است.

مرور پیشینه پژوهش نشان می‌دهد، استفاده از الگوریتم های یادگیری ماشین توانسته به صورت معنی داری به دسته‌بندی و پیش‌بینی قیمت در بازارهای مالی کمک کند (تان[1] و همکاران ۲۰۱۹، متور[2] و همکاران ۲۰۱۹). در حالی که مدل‌های یادگیری ماشین در پیش‌بینی بازارهای مالی توانسته توجه بسیاری از فعالین را به خود جذب کند، در مطالعات اخیر، مدل‌های متنوع یادگیری جمعی توانسته اند دقت پیش‌بینی افزایش دهند (نتی[3] و همکاران، ۲۰۲۰). فقیهی نژاد و مینایی (۱۳۹۷) پژوهش‌هایی که در حوزه پیش‌بینی قیمت با استفاده از یادگیری ماشین (مدل‌های هوشمند) انجام شده‌اند را به سه گروه تقسیم کرده اند؛ گروه اول، پژوهش‌هایی هستند که تنها از یک مدل یا الگوریتم استفاده برای پیش‌بینی استفاده می‌کنند. گروه دوم، از چند مدل به صورت ترکیبی استفاده می‌کنند. گروه سوم، استفاده از مدل‌های یادگیری جمعی برای تجمیع خروجی‌ها مدل‌ها استفاده می‌نمایند. در ادامه پژوهش‌هایی که از مدل‌های یادگیری جمعی برای پیش‌بینی قیمت استفاده کردند، مرور می‌شوند.

یانگ[4] و همکاران (۲۰۲۱)، از مدل‌های XGBoost و LightGBM برای پیش‌بینی قیمت سهام استفاده کردند. نتیجه این پژوهش، عملکرد بهتر مدل ترکیبی از هر یک از مدل‌های به صورت مجزا در پیش‌بینی قیمت سهام را نشان داد.

لی و پن[5] (۲۰۲۲)، در پژوهشی یک مدل نوین یادگیری عمیق را برای پیش‌بینی حرکت قیمت سهام پیشنهاد دادند. این مدل از مدل یادگیری جمعی برای ترکیب دو شبکه عصبی بازگشتی استفاده کرده است. در این پژوهش از داده شاخص S&P 500 استفاده شده است. نتیجه این پژوهش نشان‌دهنده کاهش ۵۷,۷٪ میانگین مجذور خطا و افزایش شاخص‌ها صحت به میزان ۴۰٪، پوشش به اندازه ۵۰٪ و F1 به اندازه ۴۴,۷۸٪ شده است.

سان[6] و همکاران (۲۰۲۰)، در پژوهشی با استفاده از GBDT و LightGBM روند بازار رمز ارزها را پیش‌بینی کردند. برای این پیش‌بینی از ۴۲ نوع رمز ارز با شاخص‌های اقتصادی کلیدی استفاده شده است. نتیجه پژوهش نشان‌دهنده عملکرد پایدار و بهتر LightGBM نسبت به سایر مدل‌ها می‌باشد.

1. Tan
2. Mathur
3. Nti
4. Yang
5. Li and Pan
6. Sun

بورگس و نوس[1] (۲۰۲۰)، یک سیستم مبتنی بر یادگیری ماشینی را برای توسعه یک استراتژی سرمایه‌گذاری با قابلیت معامله در بازارهای رمز ارزها پیشنهاد کردند. آن‌ها از مدل‌های رگرسیون لجستیک، RF، دسته‌بندی کننده بردار پشتیبان و تقویت درخت گرادیان برای پیش‌بینی با کمک شاخص‌های تحلیل تکنیکال به عنوان متغیرهای ویژگی استفاده کردند. آنها نشان دادند که صرف نظر از روش نمونه‌گیری مجدد مورد استفاده، همه الگوریتم های یادگیری بهتر از استراتژی خرید و نگه داشتن (B&H) در اکثریت قریب به اتفاق ۱۰۰ بازار مورد آزمایش قرار گرفتند.

لامیری و بیکیروس[2] (۲۰۱۹)، از سه مدل متفاوت LSTM، GRNN و نزدیک‌ترین همسایه برای بررسی رفتاز غیر خطی بیت‌کوین استفاده کردند. مطابق با نتیجه یون و همکاران (۲۰۲۱) آن‌ها دریافتند که LSTM خطای کمتری نسبت به سایر روش‌ها دارد و می‌تواند الگو ها را بهتر تشخیص دهد.

ژانگ و همکاران[3] (۲۰۱۸)، مدلی را با استفاده از جنگل‌های تصادفی، یادگیری عدم تعادل، انتخاب ویژگی و برش برای پیش‌بینی حرکت قیمت سهام و فاصله نرخ رشد (یا کاهش) آن پیشنهاد کردند. این مدل با استفاده از بیش از ۷۰ شاخص تحلیل تکنیکال به عنوان ورودی، اهداف پیش‌بینی را در ۴ کلاس می‌توان طبقه‌بندی کرد: بالا، پایین، مسطح و ناشناخته. این مدل با استفاده از بیش از ۴۰۰ سهم در بازار شنژن ارزیابی شد. نتایج نشان دادند که از نظر دقت و بازده در هر معامله مدل پیشنهادی این پژوهش از شبکه عصبی مصنوعی، ماشین بردار پشتیبان[4] و کی-نزدیک‌ترین همسایه‌ها[5] بهتر عمل می‌کند.

آگروال[6] و همکاران (۲۰۲۰)، با هدف کشف ماهیت قیمت بیت‌کوین و پیش‌بینی آن از مدل‌های یادگیری ماشین استفاده کردند. در این پژوهش از مدل یادگیری جمعی (CEEMD) استفاده شده است. قیمت روزانه بیت‌کوین از سال ۲۰۱۲ الی ۲۰۱۸ مورد استفاده قرار گرفته است تا به سه قسمت؛ کوتاه، میان و بلندمدت تقسیم شوند. این پژوهش از مدل ماشین بردار پشتیبان استفاده کرده است که می‌تواند برای ۵ گام کوتاه‌مدت بیت‌کوین را پیش‌بینی نماید.

دنیز[7] و همکاران (۲۰۱۹)، در پژوهشی در گام اول مهندسی ویژگی و الگوریتم‌های متفاوت انتخاب متغیر ویژگی برای پیش‌بینی قیمت بیت‌کوین پیش‌بینی اجرا نمودند. سپس از مدل‌های شبکه عصبی مصنوعی، ماشین بردار پشتیبان و مدل‌های یادگیری جمعی مبتنی بر شبکه عصبی بازگشتی و خوشه‌بندی میانگین K تایی برای پیش‌بینی قیمت بیت‌کوین استفاده کردند. نتیجه حاصل از این پژوهش افزایش ۱۰درصد دقت پیش‌بینی نتایج پژوهش‌های قبلی است.

1. Borges and Neves
2. Lahmiri and Bekiros
3. Zhang et al
4. Support vector machines - SVMs
5. k-nearest neighbors' algorithm
6. Aggarwal
7. Dennys

نتی[1] و همکاران (۲۰۲۰)، یک پژوهش تحلیلی مقایسه‌ای گسترده از مدل‌های یادگیری جمعی اعم از ؛ Bagging، Boosting، Blending و Stacking انجام دادند. مدل‌های درخت تصمیم گیری، ماشین بردار پشتیبان، شبکه عصبی با ۲۵ حالت مختلف با یادگیری جمعی ترکیب شدند. این پژوهش از سال ۲۰۱۲ الی ۲۰۱۸ شاخص‌های بورس غنا، ژوهانسبورک، بمبئی و نیویورک با استفاده از مدل‌ها پیش‌بینی نموده و عملکرد آن‌ها را مورد مقایسه قرار داده است. نتیجه این پژوهش نشان می‌دهد مدل‌های Blending و Stacking دقت بالاتر و خطای کمتری نسبت به مدل‌های bagging و boosting دارد.

جاکارت و همکاران[2](۲۰۲۱)، پیش‌بینی پذیری بازار بیت‌کوین را در افق‌های پیش‌بینی ۱ تا ۶۰ دقیقه تجزیه و تحلیل کردند. آن‌ها دریافتند که RNN و طبقه بندی Gboost برای پیش‌بینی بازار عملکرد قابل قبولی دارند. آن‌ها از متغیرهای ویژگی متنوعی شامل تحلیل تکنیکال، داده‌های بلاکچین، احساسات و غیره استفاده کردند. به علاوه، قابلیت پیش‌بینی در افق‌های طولانی‌تر افزایش می‌یابد.

یون[3] و همکاران (۲۰۲۱)، یک مدل ترکیبی از XGBoost و الگوریتم ژنتیک با یک فرآیند مهندسی ویژگی گسترده بر روی بیش از ۶۰ شاخص تحلیل تکنیکال برای پیش‌بینی‌های بازار سهام پیشنهاد کردند. مدل ترکیبی آنها می‌تواند از نظر عملکرد و تفسیرپذیری بهتر از مدل‌های LSTM عمل کند.

نتی آدکویا و ویوری[4] (۲۰۲۰) یک طبقه‌بندی گروه همگن[5] بر اساس الگوریتم ژنتیک برای انتخاب ویژگی‌ها و بهینه‌سازی پارامترهای SVM برای پیش‌بینی حرکت قیمت ۱۰ روزه در بورس غنا پیشنهاد کرد. آنها از روش گروه رای اکثریت ساده برای ترکیب نتایج حاصل از ۱۵ مدل بردار ماشین پشتیبان مختلف با استفاده از ۱۴ شاخص تحلیل تکنیکال به عنوان ورودی استفاده کردند. نتایج تجربی آنها نشان داد مدل مجموعه آنها دقت پیش‌بینی بالاتری از حرکت قیمت سهام در مقایسه با درخت تصمیم، جنگل تصادفی و شبکه عصبی دارد.

آمپوما و همکاران [6] (۲۰۲۰) اثربخشی مدل‌های مختلف مجموعه مبتنی بر درخت از جمله جنگل تصادفی، XGBoost، Bagging، AdaBoost، Extra Trees Classifier و Voting Classifier را در پیش‌بینی جهت حرکت قیمت سهام مقایسه کرد. برای این مطالعه از هشت داده مختلف سهام از سه بورس اوراق بهادار NYSE، NASDAQ وNSE استفاده شد. آنها از تجزیه و تحلیل اجزای اصلی برای انتخاب ویژگی ۴۵ ورودی شامل ۴۰ شاخص تحلیل تکنیکال استفاده کردند. نتایج تجربی نشان داد طبقه‌بندی‌کننده Extra Trees نسبت به سایر مدل‌ها، در تمام رتبه‌بندی‌ها عملکرد بهتری داشت.

تا[7]و همکاران (۲۰۲۰) پرتفویی با استفاده از شبکه عصبی LSTM و سه روش بهینه‌سازی پرتفوی، به عنوان مثال؛ روش هم وزن شبیه‌سازی مونت کارلو و مدل MV ایجاد کردند. همچنین، آنها از رگرسیون خطی و SVM به عنوان مقایسه در فرآیند انتخاب سهام استفاده کردند. نتایج تجربی نشان داد که شبکه

1. Nti
2. Jaquart
3. Yun
4. Nti, Adekoya, and Weyori
5. Homogeneous Ensemble Classifier
6. Ampomah.
7. Ta

عصبی LSTM دارای دقت پیش‌بینی بالاتری نسبت به رگرسیون خطی و SVM است و پرتفوهای ساخته شده آن بهتر از سایرین عمل می‌کند. این مدل‌ها روش‌های مختلفی را برای انتخاب سهام اعمال می‌کنند، سپس، مدل‌های بهینه‌سازی سبد سهام را با سهام منتخب برای سرمایه‌گذاری تجاری ایجاد می‌کنند. این روش‌ها جهت امیدوار کننده‌ای برای ساخت مدل‌های پرتفو را در عمل نشان می‌دهند. با این حال، مدل‌های کلاسیک بهینه‌سازی اوراق بهادار اغلب برای سرمایه‌گذاری عملی کوتاه‌مدت نامناسب هستند. بنابراین، کشف رویکرد کارآمدتر برای ترکیب نتایج پیش‌بینی بازده با مدل‌های بهینه‌سازی پرتفو مهم است.

کیم[1](۲۰۲۱)، یک سیستم معاملاتی تطبیقی جدید ایجاد کردند که در آن از یادگیری ماشینی و آزمایش برگشتی[2] برای بازار اوراق قرضه استفاده می‌کند. که در آن یک مدل پیش‌بینی که اسپرد بین اوراق خزانه ۱۰ و ۳ ساله را پیش‌بینی می‌کند پیشنهاد شده است. متعاقباً برای تأیید عملکرد مدل پیش‌بینی، از آزمون برگشتی استفاده می‌شود. که AdaBoost از سایر مدل‌های پیش‌بینی بهتر عمل می‌کند. علاوه بر این، زمانی که آزمون برگشتی بر اساس نتایج مدل‌های پیش‌بینی‌کننده اعمال شد، تا ۵۴/۲ درصد در بازده سرمایه‌گذاری طی ۶ ماه کسب شد.

فقیهی نژاد و مینایی[3] (۲۰۲۱)، مدلی برای پیش‌بینی بازار سهام با استفاده از مدل‌های هوشمند و یادگیری ماشین ارائه کردند. در این پژوهش برای افزایش دقت از مدلی بر مبنای الگوریتم یادگیری جمعی با مدل‌های پایه شبکه عصبی استفاده شده است. نتایج این پژوهش نشان داد اولاً، پیش‌بینی رفتار بازار سهام با وجود ماهیت نوسانی و ناپایدار آن امکان‌پذیر است. ثانیاً، مدل پیشنهادی این پژوهش در مقایسه با سایر روش‌ها با دقت بیشتری می‌تواند بر نوسان‌های بازار غلبه کند.

عباسی و همکاران(۲۰۲۰)، عملکرد استراتژی‌های مبتنی بر اندیکاتورهای تحلیل تکنیکال را با استراتژی خرید و نگهداری مقایسه کردند و به این نتیجه رسیدند که استراتژی‌های مبتنی بر تحلیل تکنیکال می‌تواند بازدهی بهتری نسبت به خرید و نگهداری داشته باشد. بنابراین، استفاده از اندیکاتورها می‌تواند در پیش‌بینی کمک کننده باشد.

سهرابی و همکاران (۱۴۰۱)، رویکرد های جنگل تصادفی، ماشین بردار پشتیبان، شبکه عصبی مصنوعی و شبکه عصبی بازگشتی برای پیش‌بینی شاخص کل بورس اوراق بهادار تهران در دوره ۱۳۹۲ تا ۱۳۹۹ مورد استفاده قرار دادند. نتایج پیش‌بینی دوره‌های ۱، ۳ و ۶ روزه برای دوره خارج از نمونه نشان می‌دهد روش یادگیری ماشین مبتنی بر شبکه عصبی بازگشتی حافظه طولانی کوتاه‌مدت (LSTM) در مقایسه با سایر مدل‌های مورد بررسی نتیجه بهتری داشته است.

بشیری و پاریاب (۱۳۹۹)، عملکرد مدل‌های جنگل تصادفی، بردار پشتیبان، گرادیان تقویتی و شبکه عصبی چند لایه پرسپترون را برای پیش‌بینی قیمت بیت‌کوین با استفاده از داده ۹ رمز ارز دیگر مورد مقایسه قرار دادند. نتیجه این پژوهش نشان‌دهنده دقت بیشتر مدل گرادیان تقویتی است.

1. Misuk Kim
2. Back-testing
3. Faghihi and Minaei

قربانی و همکاران (۱۴۰۱)، برای پیش‌بینی تغییرات قیمتی بیتکوین با استفاده از الگوریتم‌های یادگیری ماشین همراه با استفاده از داده‌های تحلیل احساسات مردم نسبت به افراد مشهور مانند ایلان ماسک در شبکه‌های اجتماعی توییتر، ردیت و تلگرام مدل ارائه کردند. در این پژوهش الگوریتم xgboost با رکورد دقت بدست آمده ۸۸ درصدی در پیش‌بینی روند تغییرات قیمت بیتکوین عملکرد عالی را نشان داد.

برای جمع‌بندی، بر اساس پژوهش مروری کروانجو و آکای، بسیاری از پژوهش‌ها به دنبال مقایسه مدل‌های آماری و یادگیری ماشین بوده اند که به‌طورکلی، مدل‌های یادگیری ماشین در مقایسه با مدل‌های آماری توانسته اند عملکرد بهتری را نشان بدهند(کروانجو و آکای[1]، ۲۰۲۰). حال با توجه به جذابیت پیش‌بینی قیمت رمز ارزها (به ویژه، بیت‌کوین) برای فعالین، بسیاری دیگر از پژوهش‌ها به مقایسه عملکرد مدل‌های یادگیری ماشین و یا ترکیب آن‌ها در دقت پیش‌بینی پرداخته‌اند. در مطالعاتی که از یادگیری جمعی استفاده شده است، این رویکرد توانسته است عملکرد و دقت پیش‌بینی قیمت را افزایش دهد. لازم به ذکر است پژوهش‌های فوق بازارهای مالی متفاوتی مانند؛ بازار سهام[2]، اوراق قرضه[3]، شاخص‌های بورس، رمز ارزها و کالا مورد بررسی قرار گرفته است که به تقویت نتیجه‌گیری برتری یادگیری جمعی منجر می‌شود.

جدول ۱. دسته بندی پژوهش‌های مرتبط با پیش‌بینی بازارهای مالی با استفاده از یادگیری ماشین

ردیف	نام دسته	پژوهش‌ها	چکیده	نقاط قابل بهبود
۱	پیش‌بینی بر اساس ترکیب مدلهای یادگیری ماشین (استفاده از یادگیری جمعی)	امپوما[4] و همکاران (۲۰۲۰) فقیهی نژاد و بیدگلی (۲۰۲۱) لیوریس[5] و همکاران (۲۰۲۰) یه[6] و همکاران (۲۰۲۲) ادکویا و ویوری[7] (۲۰۱۹) یانگ[8] و همکاران (۲۰۲۱) ادکویا و ویوری[9] (۲۰۲۰) دنیس[10] و همکارن (۲۰۱۹) آگاروال[11] و همکاران (۲۰۲۰) لی و پن[12] (۲۰۲۲)	در این پژوهش‌ها، مدلهای یادگیری ماشین با استفاده از یادگیری جمعی با هم ترکیب شده‌اند. نشان داده شده است که یادگیری جمعی عملکرد بهتری نسبت به هر یک از مدلهای پیش‌بینی به صورت جداگانه دارد.	پیش‌بینی کف و سقف نمودار می‌تواند نسبت به بازدهی یا خود قیمت عملکرد بهتری داشته باشد.

1. Kervancı and F. Akay
2. Stock
3. Bonds
4. Ampomah
5. Livieris
6. Ye
7. Adekoya, Weyori
8 .Yang
9. Adekoya, Weyori
10. Dennys
11. Aggarwal
12. Li and Pan

ردیف	نام دسته	پژوهش‌ها	چکیده	نقاط قابل بهبود
۲	مقایسه عملکرد مدلهای یادگیری ماشین در پیش‌بینی قیمت	سان[1] و همکاران (۲۰۲۰) ژیانگ[2] و همکاران (۲۰۲۰) یون[3] و همکاران (۲۰۲۱) ریمن[4] و همکاران (۲۰۲۰) مانچاندا و آگاروال[5] (۲۰۲۱) لامیری و بیکیروس[6] (۲۰۱۹) جاکارت و همکاران[7] ۲۰۲۱ بورگس و نوس[8] (۲۰۲۰) ناگولا و الکساکیس[9] (۲۰۲۲) جی[10] و همکاران (۲۰۱۹) تا[11] و همکاران (۲۰۲۰)	در این پژوهش‌ها، عملکرد مدلهای مختلف یادگیری ماشین با هم مقایسه شده است.	پیش‌بینی کف و سقف نمودار می‌تواند نسبت به بازدهی یا خود قیمت عملکرد بهتری داشته باشد.

شکاف مطالعاتی پیشنهادی این پژوهش، رویکرد مدلسازی داده‌های قیمت می‌باشد. در بسیاری از پژوهش‌ها به دنبال پیش‌بینی عدد قیمت یا به دلایل مانایی و غیره، عدد بازدهی هستند. در صورتی که این رویکرد می‌تواند دقت پیش‌بینی را کاهش دهد. با اینکه قیمت یک داده سری زمانی[12] است، می‌توان به شکل گسسته به آن نگاه کرد. به معنی آن که، به جای اینکه پیش‌بینی قیمت یا بازدهی آن می‌بایست به دنبال پیش‌بینی وضعیت سقف یا کف بود. اینکه بدانیم در مسیر ساختن سقف قیمتی یا کف قیمتی هستیم، در صورتی که باعث افزایش دقت و کارایی مدل‌های پیش‌بینی شود، می‌توان بر اساس آن معاملات الگوریتمی مطمئن‌تری طراحی و پیاده‌سازی کرد. بنابراین، این پژوهش به دنبال پیش‌بینی وضعیت سقف یا کف قیمتی بر اساس مدل‌های یادگیری ماشین می‌باشد. در ادامه، روش پژوهش و چگونگی پیش‌بینی وضعیت سقف و کف مورد استفاده در این پژوهش ارائه می‌شود.

سئوالات پژوهش

با توجه به هدف این پژوهش که به دنبال پیش‌بینی سقف یا کف قیمت بیت‌کوین با استفاده از مدل‌های یادگیری ماشین است، سئوالات این پژوهش را می‌توان به شکل زیر ارائه نمود:

1. Sun
2. Jiang
3. Yun
4. Rahman
5. Manchanda and Aggarwal
6. Lahmiri and Bekiros
7. Jaquart
8. Borges and Neves
9. Nagula and Alexakis
10. Ji
11. Ta
12. Time Series

۱- عملکرد سیستم معاملاتی هوشمند مبتنی بر مدل‌های پیش‌بینی جنگل تصادفی، یادگیری حافظه بلند و کوتاه‌مدت، شبکه عصبی بازگشتی و یادگیری جمعی در مقایسه با داده واقعی و استراتژی خرید و نگهداری[1] چگونه است؟
۲- میزان دقت و صحت یادگیری ماشین جمعی در پیش‌بینی سقف‌وکف‌های بیت‌کوین چقدر است؟
۳- آیا میزان دقت و صحت پیش‌بینی سقف‌وکف‌های بیت‌کوین با استفاده از یادگیری ماشین جمعی نسبت به زیر الگوریتم ها بیشتر است؟

روش شناسی پژوهش

با توجه به اینکه بازار مالی هدف در این پژوهش بازار رمز ارزها (بیت‌کوین) می‌باشد، استفاده از پایگاه داده معتبر اهمیت دارد. در این پژوهش داده‌های قیمت کندل ها (OHLCV) بیت‌کوین در تایم فریم ۱ ساعته به عنوان بزرگترین رمز ارز به عنوان نماینده این بازار انتخاب شده است. برای استخراج این داده از ماژول پایتونی Historic-Crypto که داده‌ها را از API صرافی CoinBase Pro استخراج می‌کند، استفاده شده است.

بازار رمز ارزها با توجه به نصف شدن پاداش استخراج، رفتار سیکلی به صورت ۴ ساله دارد. بنابراین، از کل داده‌های بیت‌کوین که از سال ۲۰۱۰ در درسترس است، بازه زمانی ۲۰۱۸ الی ۲۰۲۲ انتخاب شده است. در سال ۲۰۱۸ بیت‌کوین بازاری رکودی و سپس نزولی، را تجربه کرده است. در سال‌های ۲۰۱۹ الی ۲۰۲۱ نیز یک روند صعودی با توجه پاندامی کرونا و نصف شدن جایزه در سال ۲۰۲۰ را تجربه کرده است. در سال ۲۰۲۲ نیز یک روند رکودی و نزولی که می‌توان گفت برگشت به سال ۲۰۱۸ است که در نتیجه نشان می‌دهد نمونه حاضر کلیه فازهای سیکلی را توضیح می‌دهد. ۷۰ درصد ابتدایی داده‌ها به عنوان داده‌های آموزشی به مدل داده می‌شود. ۲۰ درصد داده‌ها به عنوان داده‌های اعتبار سنجی در نظر گرفته شده و ۱۰ درصد انتهایی داده‌ها به عنوان داده‌های تست به مدل داده می‌شوند.

با استفاده از زبان برنامه‌نویسی پایتون و ماژول ها معتبر در آن این پژوهش انجام شده است. در بستر Google Colab با توجه به اشتراک‌گذاری GPU و همچنین، از کتابخانه‌های numpy، Pandas، ta، tensorflow، sklearn وScipy به طور خاص برای پیاده‌سازی استفاده شده است. برای ارزیابی عملکرد استراتژی‌ها از چارچوب پایتونی freqtrade استفاده شده است.

متغیرهای این پژوهش به دو دسته اصلی تقسیم می‌شوند. متغیر هدف، متغیری است که به دنبال پیش‌بینی آن هستیم. متغیر هدف در این پژوهش متغیر سقف یا کف (۱ یا ۰) است که با استفاده از اندیکاتور Awesome Oscillator (AO) از قیمت بسته شدن کندل‌های استفاده شده است. سایر متغیرهای اندیکاتورهای تحلیل تکنیکال به عنوان متغیر ویژگی، برای پیش‌بینی سقف یا کف‌های قیمتی می‌باشند. لازم به ذکر است برای انتخاب این متغیرها از کتابخانه‌های موجود همبستگی بالغ بر ۱۵۰ اندیکاتور و اسیلاتور با متغیر هدف مورد بررسی قرار گرفته و اندیکاتورهایی (عددی) که همبستگی پیرسون بالای ۷۰ و با p_value ۰/۰۵ داشتند، به عنوان متغیر ویژگی انتخاب شده‌اند. در ادامه لیست این

1. Buy and Hold

اندیکاتورها معرفی می‌شود. پس از آن داده‌های یادگیری به دنبال کشف ارتباط بین لیست اندیکاتورها و متغیر هدف یعنی سقف و کف‌ها با استفاده از مدل‌های RF، RNN و LSTM می‌باشد. اعداد مربوط به اندیکاتور و اسیلاتور با تقسیم شدن به عدد قیمت بسته شدن نرمال شده‌اند تا در یک مقیاس باشند. همان‌طوری که گفته شده ۷۰ درصد داده به عنوان داده یادگیری و ۳۰ درصد ما بقی جهت اجرای مراحل بعدی در نظر گرفته شده است. لیست اندیکاتور و اسیلاتورهای مورد استفاده در این پژوهش، به شرح ذیل می‌باشد:

جدول ۲. لیست متغیرهای ویژگی مورد استفاده در مرحله اول

نام در مدل	نام متغیر	نام در مدل	نام متغیر
volume_mfi	حجم (شاخص جریان پول)	'volume_sma_em'	میانگین حجم
'volatility_bbp'	نوسان BBP	'volatility_kcp'	نوسان KCP
'volatility_dcp	نوسان DCP	'trend_macd'	روند MACD
'trend_macd_diff'	اختلاف روند MACD	'trend_adx_pos'	روند ADX
'trend_vortex_ind_pos'	روندVORTEX	'trend_vortex_ind_diff'	اختلاف روند VOREXT
'trend_cci'	روند CCI	'trend_aroon_up'	روند AROON
'trend_stc'	روند STC	'momentum_rsi'	جهش RSI
'momentum_tsi	جهش TSI	'momentum_uo'	جهش UO
'momentum_stoch'	جهش STOCHAS'IC	'momentum_stoch_signal'	جهش STOCH SIGNAL
'momentum_wr'	جهش WR	'momentum_ao'	جهش AO
'momentum_roc'	جهش ROC	'ao'	AO
'RSI'	شاخص قدرت نسبی	'aboveEMA10'	بالا میانگین متحرک ۱۰ تایی
'aboveEMA15'	بالا میانگین متحرک ۱۵ تایی	'aboveEMA20'	بالا میانگین متحرک ۲۰ تایی
'aboveEMA30'	بالا میانگین متحرک ۳۰ تایی	'aboveEMA40'	بالا میانگین متحرک ۴۰ تایی
'aboveEMA50'	بالا میانگین متحرک ۵۰ تایی	'aboveEMA60'	بالا میانگین متحرک ۶۰ تایی

در لیست بالا اندیکاتورهای تحلیل تکنیکال مربوط به حجم معاملات، نوسان‌پذیری قیمت، مبتنی بر روند، جهش‌های قیمتی و همچنین، اندیکاتورهای باینری معرفی شده است. همانطور که گفته شد، این اندیکاتورها رابطه معنی‌داری با متغیر هدف (سقف یا کف) دارند. لازم به ذکر است که اندیکاتورهای باینری از قسمت مهندسی ویژگی این پژوهش به این لیست اضافه شده‌اند. زیرا این متغیرها در بهبود دقت مدل‌ها توانسته‌اند نقش به سزایی داشته باشند. برای توضیح بیشتر اندیکاتورهای باینری، اندیکاتور بالای EMA(10) را در نظر بگیرید که خود یک متغیر باینری ۰ و ۱ می‌باشد. در صورتی که قیمت بالای میانگین متحرک نمایی ۱۰ تایی باشد، عدد ۱ و اگر کمتر از آن باشد، عدد صفر را به خود اختصاص می‌دهد.

مراحل اجرای مدل به صورت کلی به شرح ذیل می‌باشد:

مرحله ۱ – داده‌های سقف و کف بیت‌کوین به عنوان متغیر هدف با استفاده از مدل‌های جنگل تصادفی، حافظه طولانی کوتاه‌مدت و شبکه عصبی بازگشتی با متغیرهای ویژگی جدول (۱)، پیش‌بینی می‌شوند. خروجی حاصل از این مرحله پیش‌بینی های سقف و کف هر یک از مدل‌ها به همراه امتیاز آن‌ها می‌باشد.

مرحله ۲- خروجی مرحله اول (پیش‌بینی سقف و کف‌ها به همراه امتیاز آن‌ها) به عنوان متغیر ویژگی به همراه سقف‌ها و کف‌ها به عنوان متغیر هدف به مدل‌های XGBoost و LightGBM برای یادگیری ارائه می‌شوند.

مرحله ۳- خروجی‌های حاصل از مرحله دوم به استفاده از الگوریتم یادگیری جمعی رأی‌گیری تا نتیجه پیش‌بینی نهایی ثبت شده و با داده واقعی مقایسه می‌شود.

مرحله ۴- گام آخر استفاده از نتایج مدل‌ها پیش‌بینی RNN، LSTM، RF و یادگیری جمعی پیشنهادی جهت سیگنال خرید و فروش در بازار رمز ارزها و مقایسه عملکرد استراتژی معاملاتی مبتنی بر آن‌ها با سیستم معاملاتی مبتنی بر داده‌های واقعی سقف و کف بیت‌کوین می‌باشد.

همان‌طور که در بخش‌های قبلی اشاره شد، متغیر هدف در این پژوهش مقدار دو دویی صفر و یک را اتخاذ می‌کند. منظور از صف، در مسیر کف و منظور از ۱، در مسیر سقف می‌باشد. با فرض یک‌طرفه بودن بازار و عدم دسترسی به تعهد فروش یا فروش استقراضی، می‌توان یک استراتژی معاملاتی را مورد ارزیابی قرار داد. در صورتی که پیش‌بینی مدل‌ها از مقدار صفر به مقدار یک تبدیل شد، به این معنی است که پیش‌بینی مدل این است در مسیر یک سقف قرار گرفته‌ایم. بنابراین، در حالت می‌توان خرید انجام داد. سپس، وقتی بر عکس مقدار ۱ مجدد به مقدار ۰ تبدیل شد، منظور این است که در مسیر کف قرار گرفته‌ایم. بنابراین، می‌توان دارایی را مورد نظر را فروخت. سیستم معاملاتی در هر معامله ۱۰ درصد از سرمایه اولیه را درگیر می‌کند. سرمایه اولیه هزار دلار در نظر گرفته شده است. در هر کندل این سیستم فقط می‌تواند ۱ معامله باز داشته باشد و آن هم فقط خرید می‌باشد. سپس، عملکرد سیستم معاملاتی برای هر یک از مدل‌ها، با مقدار واقعی و استراتژی خرید و نگهداری مقایسه خواهد شد. در جدول (۳)، شاخص‌های عملکردی سیستم معاملاتی که مورد مقایسه قرار می‌گیرد، قابل مشاهده می‌باشد:

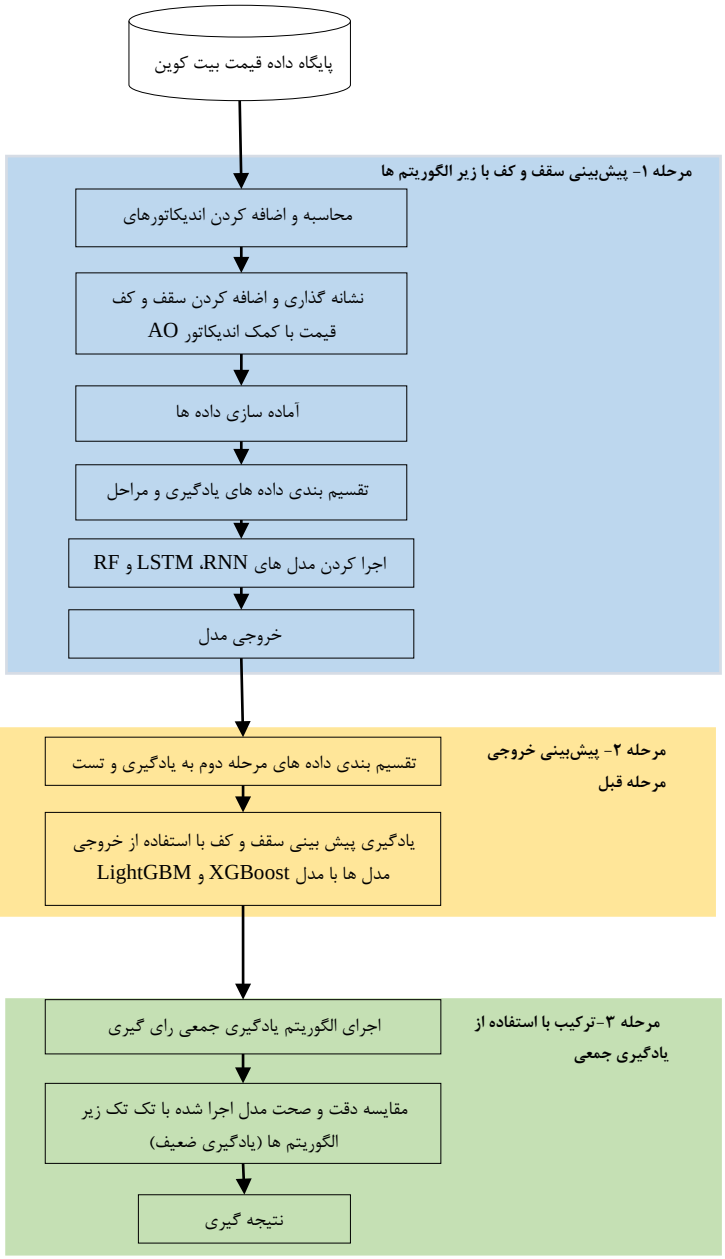
جدول ۳. شاخص‌های عملکردی سیستم معاملاتی

ردیف	نام شاخص عملکرد	توضیح
۱	تعداد معامله	تعداد کل معامله انجام گرفته در بازه تست. هر خرید و فروش یک معامله در نظر گرفته می‌شود.
۲	تاریخ شروع	تاریخ شروع بک تست
۳	تاریخ پایان	تاریخ پایان بک تست
۴	کل سود یا ضرر	جمع سود و ضرر کل معاملات
۵	بازدهی	حاصل تقسیم سود یا ضرر به سرمایه اولیه
۶	میانگین بازدهی هر معامله	مجموع بازدهی معاملات تقیسم بر تعداد معاملات
۷	درصد موفقیت	تعداد معاملات سود ده به کل معاملات
۸	شاخص سود دهی[1]	حاصل تقیسم مجموع سود به مجموع ضرر
۹	بیشترین افت سرمایه[2]	کمترین سرمایه ی در بازه تست
۱۰	بیشترین سرمایه در لحظه	بیشتر سرمایه در بازه تست

منبع: یافته‌های پژوهش

بازه زمانی داده‌های تست از ۲۰۲۲-۰۳-۱۱ تا ۲۰۲۲-۰۵-۳۰ می‌باشد. برای اجرا تست سیستم معاملاتی از فریم ورک پایتونی Freqtrade که مبتنی بر Docker می‌باشد، استفاده شده است.

1. Profit Factor
2. Max Draw Down (Max DD)

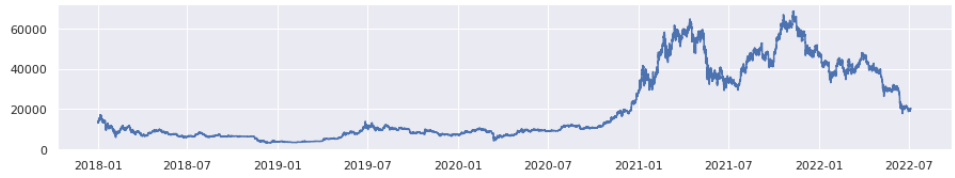


شکل ۱. مراحل اجرای پژوهش

تجزیه و تحلیل داده‌ها و آزمون فرضیه‌ها

مروری بر داده‌ها

همانطور که در قسمت قبل گفته شد داده‌ها قیمت بیت‌کوین در بازه زمانی ۲۰۱۸ الی ۲۰۲۲ در تایم فریم ۱ ساعته بعد از فرآیند تمیز شدن برای کل مراحل مورد استفاده قرار گرفتند. نمودار (۱)، قیمت بیت‌کوین به صورت خطی از سال ۲۰۱۸ تا آخر ۶ ماهه اول سال به شکل زیر می‌باشد:



نمودار ۱. نمودار خطی بیت‌کوین به دلار

داده‌های قیمت بیت‌کوین در بازه زمانی اشاره شده دارای آمار توصیفی طبق جدول (۴)، می‌باشند:

جدول ۴. آمار توصیفی قیمت بسته شدن کندل‌های ۱ ساعته بیت‌کوین از ۲۰۱۸ تا آخر نیم سال اول ۲۰۲۲

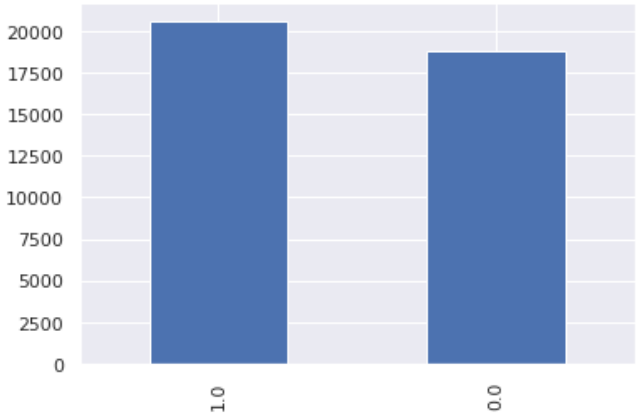
ردیف	نام شاخص	مقدار شاخص
۱	تعداد داده قیمت	۳۹،۴۸۴
۲	بزرگترین داده	۶۸،۶۳۹
۳	کوچکتری داده	۳،۱۳۹
۴	میانه	۱۰،۱۴۲
۵	میانگین	۲۰،۳۸۴
۶	مد	۶،۳۹۹
۷	انحراف معیار	۱۷،۸۷۶
۸	چولگی	۰/۹۷
۹	کشیدگی	-۰/۵۶

منبع: یافته‌های پژوهش

در نمودار (۲) برای داده آخر نمودار بیت‌کوین به همراه وضعیت سقف و کف را مشخص شده است که به عنوان داده هدف به مدل‌های یادگیری ماشین داده شده است:

نمودار ۲. قیمت بیت‌کوین به همراه وضعیت سقف و کف برای ۱۰۰ داده اخیر

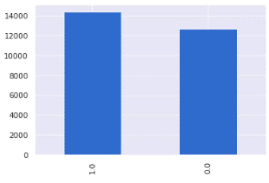
تعداد داده‌ها ۳۹،۴۸۴ می‌باشد. از کل این داده‌ها ۲۰،۶۶۳ (۵۲٪) داده وضعیت سقف و ۱۸،۸۲۱ (۴۷٪) داده وضعیت کف را نشان می‌دهد. این موضوع نشان از توزان نسبی میان تعداد داده سقف و کف را نشان می‌دهد.



نمودار ۳. تعداد وضعیت‌های سقف (۱) و کف (۰) در داده‌های قیمت بیت‌کوین برای کلیه مراحل

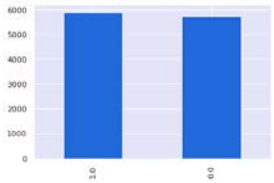
منبع: یافته‌های پژوهش

متوازن بودن داده‌های سقف و کف برای تمامی مراحل نیز می‌بایست به صورت مستقل بررسی شود. در نمودارهای ۴ تا ۶، تعداد سقف‌ها (۱) و کف‌ها (۰) به تفکیک مراحل پژوهش مشخص شده‌اند. این تعداد نشان‌دهنده توازن در متغیر هدف در تمامی مراحل می‌باشد.



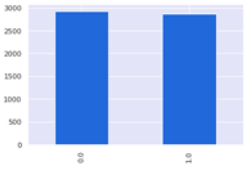
تعداد کل داده: ۲۷،۰۲۸
درصد داده سقف: ٪۵۳
درصد داده کف: ٪۴۶

نمودار ۴. تعداد سقف‌وکف‌های در مرحله (۱)



تعداد کل داده: ۱۱،۵۸۳
درصد داده سقف: ٪۵۰/۸
درصد داده کف: ٪۴۹/۱۹

نمودار ۴. تعداد سقف‌وکف‌های در مرحله (۲)



تعداد کل داده: ۵،۷۹۱
درصد داده سقف: ٪۵۰/۵۲
درصد داده کف: ٪۴۹/۴۷

نمودار ۵. تعداد سقف‌وکف‌های در مرحله (۳)

منبع: یافته‌های پژوهش

ساختار و پارامترهای ورودی مدل‌ها

مدل‌های مرحله اول دارای ساختار و پارامتر های وروی

جدول ۵. اطلاعات مربوط به ساختار مدل‌های LSTM و RNN در مرحله اول

ردیف	نام مدل	اطلاعات لایه	تعداد واحد	نسبت صفر کردن ورودی	تابع فعالیت
۱	RNN	لایه اول (لایه ورودی)	۶۴	-	hyperbolic tangent (tanh)
		لایه دوم	۶۴	-	hyperbolic tangent (tanh)
		لایه خروجی	۲ (باینری)	-	sigmoid
۲	LSTM	لایه اول	۶۴	-	hyperbolic tangent (tanh)
		Dropout	-	۰/۲	-
		لایه دوم	۶۴	-	hyperbolic tangent (tanh)
		Dropout	-	۰/۲	-
		لایه سوم (خروجی)	۲ (باینری)	-	softmax

منبع: یافته‌های پژوهش

لازم به ذکر است که پارامترهای ورودی با استفاده از جستجوی پنجره ای به دست آمده اند.

جدول ۶. شاخص‌های مربوط به توقف و اجرای یادگیری مدل‌های LSTM و RNN در مرحله اول

نام مدل	شاخص توقف یادگیری			شاخص‌های اجرای یادگیری		
	Monitor	patience	verbose	optimizer	loss	metrics
RNN	‘loss’	۳	۱	adam	binary_ crossentropy	['accuracy']
LSTM	‘loss’	۳	۱	adam	sparse_ categorical_ crossentropy	['accuracy']

منبع: یافته‌های پژوهش

جدول ۷. ساختار مدل جنگل تصادفی در مرحله اول

ردیف	نام مدل	Criterion	تعداد درخت	حداکثر عمق	تعداد کار همزمان	Minimum number of samples required to be at a leaf node
۱	Random Forest	gini	۲۰	۲۰	using all processors	۲

منبع: یافته‌های پژوهش

مدل XGBOOST و LightGBM

جدول ۸. ساختار مدل‌های XGBOOST و LightGBM

ردیف	نام مدل	تعداد درخت تقویت شده	نرخ یادگیری	حداکثر عمق درخت ها	نسبت نمونه فرعی	حداقل مجموع وزن نمونه مورد نیاز در یک بچه (برگ)	نسبت نمونه فرعی به آزمایشی	گاما
۱	XGBOOST	۲۱	۰/۳	۵	۰/۲۵	۴	۰/۲۱	۰/۲۲
۲	LightGBM	۱۵	۰/۱۹	۱	۰/۴	۳	۰/۵۸	-

منبع: یافته‌های پژوهش

لازم به ذکر است ما بقی پارامترها مقادیر پیش فرض در نظر گرفته شده است. مقادیر فوق نیز بر اساس تنظیم هایپر پارامترهای [1] به دست آمده‌اند. برای مثال نرخ یادگیری اعدادی بین ۰/۰۱ تا ۰/۵ را با گام‌های ۰/۰۱ طی کرده است تا به عدد بهینه ۰/۱۹ برای LightGBM و ۰/۳ برای XGBoost رسیده است. مابقی پارامترها هم تنظیم شده‌اند.

در این قسمت خروجی هر یک از مدل‌ها را با شاخص‌ها ارزیابی دقت، صحت، پوشش و F1 به صورت مستقل ارائه و تحلیل می‌شود. شاخص دقت[2]، حاصل تقسیم موارد درست پیش‌بینی شده به تمام موارد می‌باشد. شاخص صحت[3]، حاصل تقسیم مواردی مثبت درست تشخیص داده شده به موارد مثبتی است که درست یا غلط تشخیص داده شده است. در آخر شاخص پوشش[4] عبارت است از موارد مثبت درست تشخیص داده شده تقسیم بر حاصل جمع داده‌های مثبت و منفی غلط تشخیص داده شده. امتیاز F1 که به صورت جدول (۹) محاسبه می‌شود، میانگینی از شاخص‌های صحت و پوشش می‌باشد:

جدول ۹. ماتریس اغتشاش[5]

پیش‌بینی توسط الگوریتم			
برچسب واقعی		بلی	خیر
	بلی	True Positive (TP)	False Negative (FN)
	خیر	False Positive (FP)	True Negative (TN)

منبع: یافته‌های پژوهش

1. Tune Hyperparameters
2. Accuracy
3. Precision
4. Recall
5. Confusion Matrix

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad -۱$$

$$precision = \frac{TP}{TP + FP} \quad -۲$$

$$Recall = \frac{TP}{TP + FN} \quad -۳$$

$$f1\ score = \frac{2 \times (precision \times Recall)}{precision + Recall} \quad -۴$$

شاخص‌های فوق می‌توانند جهت مقایسه عملکرد پیش‌بینی مدل‌ها و مقایسه آن‌ها کمک کنند. در واقع هر کدام از آن‌ها با رویکرد مربوط به خود به دنبال سنجش صحت و دقت پیش‌بینی می‌باشند که هر چقدر به عدد ۱ نزدیک‌تر باشند، نشان از دقت بیشتر دارد. با توجه به اینکه شاخص f1 ترکیبی از شاخص‌های دقت و صحت می‌باشد، در این پژوهش جهت مقایسه عملکرد دو مدل بیشتر به این شاخص توجه می‌شود.

خروجی‌های مدل‌ها

خروجی مدل‌های یادگیری حافظه کوتاه و بلندمدت، شبکه عصبی بازگشتی، جنگل تصادفی و یادگیری جمعی در بازه تست (مرحله) سوم در جداول زیر ارائه می‌شود. لازم به ذکر است بازه زمانی تمامی خروجی‌ها با هم مشابه است تا بتوان عملکرد مدل‌ها را با یکدیگر مقایسه نمود. بنابراین، به میزان خروجی نهایی (داده تست) مدل یادگیری جمعی، خروجی‌های مدل‌های RNN، FR و LSTM مرتب شده تا بتوان نتایج را مقایسه نمود.

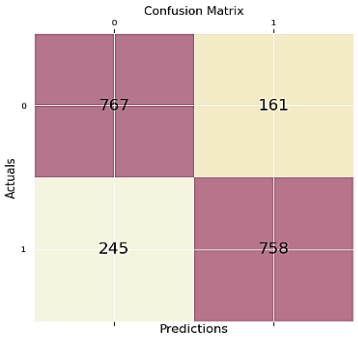
مدل RNN با پارامترهای ورودی مشخص شده در جدول (۳) و (۴) با متغیر هدف وضعیت سقف و کف و متغیرهای ویژگی معرفی شده نتایج زیر را داشته است:

جدول ۱۰. شاخص‌های ارزیابی خروجی مدل RNN

مدل شبکه عصبی بازگشتی – دقت ۷۹٪				
هدف	صحت	پوشش	F1	Support
۰	۰/۸۳	۰/۷۶	۰/۷۹	۱۰۱۲
۱	۰/۷۶	۰/۸۲	۰/۷۹	۹۱۹
دقت	-	-	۰/۷۹	۱۹۳۱
میانگین ماکرو	۰/۷۹	۰/۷۹	۰/۷۹	۱۹۳۱
میانگین وزنی	۰/۷۹	۰/۷۹	۰/۷۹	۱۹۳۱

منبع: یافته‌های پژوهش

همان‌طور که در جدول (۱۰) مشخص شده است، دقت مدل RNN در پیش‌بینی وضعیت سقف و کف ۷۹ درصد می‌باشد. صحت مدل در تشخیص کف‌ها ۸۳ درصد و در سقف‌ها ۷۶ درصد بوده است. این موضوع نشان‌دهنده عملکرد بهتر این مدل در تشخیص کف‌ها نسبت به سقف‌هاست. البته که بازار نزولی رمز ارزها در بازه تست شاید بتواند دلیل دیگر برای این تفاوت باشد. شاخص پوشش برای وضعیت‌های کف ۷۶ درصد و برای وضعیت های سقف ۸۲ درصد می‌باشد. شاخص F1 که یک شاخص متعادل برای مدل است نیز، عدد ۷۹ درصد را ثبت کرده است. به‌طورکلی، می‌توان گفت ۷۹ درصد از سقف‌ها و کف‌ها به درستی پیش‌بینی شده‌اند. در شکل (۲) ماتریس اغتشاش این مدل برای داده‌های تست مشخص شده است:



شکل ۲. ماتریس اغتشاش RNN روی داده تست

منبع: یافته‌های پژوهش

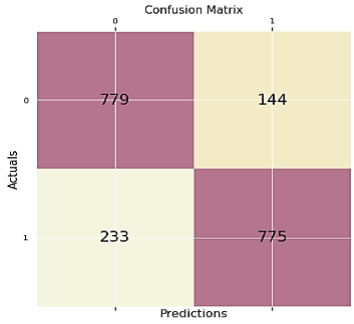
مدل LSTM مدل بعدی است که در جدول (۹) شاخص‌های ارزیابی آن مورد بررسی قرار گرفته است. این مدل نیز بر اساس ساختار و هایپر پارامترهایی که در جدول (۳) و (۴) مشخص شد، اجرا شده است.

جدول ۱۱. شاخص‌های ارزیابی خروجی مدل LSTM

مدل یادگیری حافظه طولانی و کوتاه‌مدت – دقت: ۸۰٪				
هدف	صحت	پوشش	F1	Support
۰	۰/۸۴	۰/۷۷	۰/۸۱	۱۰۱۲
۱	۰/۷۷	۰/۸۴	۰/۸۰	۹۱۹
دقت	-	-	۰/۸۰	۱۹۳۱
میانگین ماکرو	۰/۸۱	۰/۸۱	۰/۸۰	۱۹۳۱
میانگین وزنی	۰/۸۱	۰/۸۰	۰/۸۰	۱۹۳۱

منبع: یافته‌های پژوهش

همان‌طور که در جدول (۱۱) مشخص شده است، دقت مدل LSTM ۸۰ درصد و نسبت به مدل RNN یک درصد بیشتر است. صحت مدل در پیش‌بینی وضعیت کف‌ها ۸۴ و در وضعیت سقف‌ها ۷۷ درصد می‌باشد. شبیه به RNN این مدل هم در تشخیص کف‌ها از سقف‌ها عملکرد بهتری از خود نشان داده است. شاخص پوشش برای وضعیت کف ۷۷ درصد و برای وضعیت سقف ۸۴ درصد می‌باشد. به‌طورکلی، شاخص F1 نیز نشان‌دهنده دقت ۸۱ درصدی برای این مدل می‌باشد. در شکل (۳) ماتریس اغتشاش برای این مدل روی داده تست مشاهده می‌شود.



شکل ۳. ماتریس اغتشاش LSTM روی داده تست

منبع: یافته‌های پژوهش

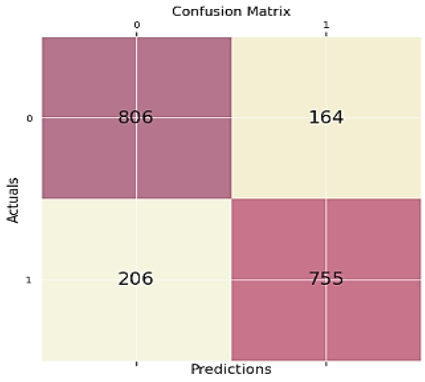
مدل جنگل تصادفی سومین مدلی است که بر اساس ساختار مدل و پارامترهای جدول (۵) روی داده‌ها اجرا شده است. در جدول (۱۲) شاخص‌های ارزیابی مدل مشاهده می‌شود:

جدول ۱۲. شاخص‌های ارزیابی خروجی مدل RF

مدل جنگل تصادفی - دقت: ۸۰.۸۳٪				
هدف	صحت	پوشش	F1	Support
۰	۰/۸۳	۰/۸۰	۰/۸۱	۱۰۱۲
۱	۰/۷۹	۰/۸۲	۰/۸۰	۹۱۹
دقت	-	-	۰/۸۱	۱۹۳۱
میانگین ماکرو	۰/۸۱	۰/۸۱	۰/۸۱	۱۹۳۱
میانگین وزنی	۰/۸۱	۰/۸۱	۰/۸۱	۱۹۳۱

منبع: یافته‌های پژوهش

همانطور که در جدول (۱۲) مشخص شده است، دقت مدل جنگل تصادفی به‌طورکلی، ۸۰/۸۳ درصد می‌باشد. این مدل نسبت به دو مدل LSTM و RNN دقت بیشتری دارد. صحت مدل در پیش‌بینی وضعیت کف ۸۳ درصد و وضعیت سقف ۷۹ درصد می‌باشد. شاخص پوشش برای وضعیت کف ۸۰ درصد و برای سقف ۸۲ درصد می‌باشد. شاخص F1 برای وضعیت کف و سقف به ترتیب ۸۱ و ۸۰ درصد می‌باشد. در شکل ۴ ماتریس اغتشاش برای RF ارائه شده است.



شکل ۴. ماتریس اغتشاش RF روی داده تست

منبع: یافته‌های پژوهش

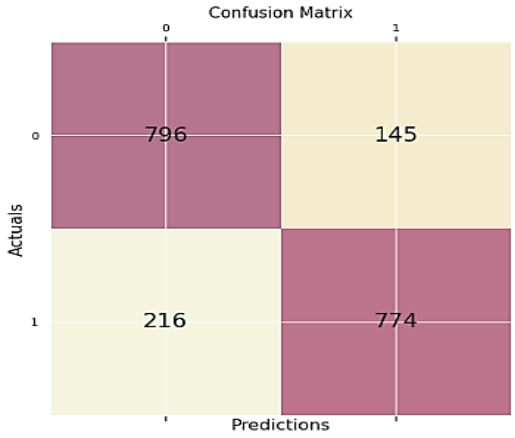
مدل یادگیری جمعی روی مدل‌های XGBoost و LightGBM اجرا شده است و نتایج شاخص‌های ارزیابی این ساختار در جدول (۱۳) ارائه شده است:

جدول ۱۳. شاخص‌های ارزیابی خروجی مدل یادگیری جمعی

مدل یادگیری جمعی روی خروجی XGBoost و LightGBM - دقت: ۸۱٫۳۰٪				
هدف	صحت	پوشش	F1	Support
۰	۰/۸۵	۰/۷۹	۰/۸۲	۱۰۱۲
۱	۰/۷۸	۰/۸۴	۰/۸۱	۹۱۹
دقت	-	-	۰/۸۱	۱۹۳۱
میانگین ماکرو	۰/۸۱	۰/۸۱	۰/۸۱	۱۹۳۱
میانگین وزنی	۰/۸۲	۰/۸۱	۰/۸۱	۱۹۳۱

منبع: یافته‌های پژوهش

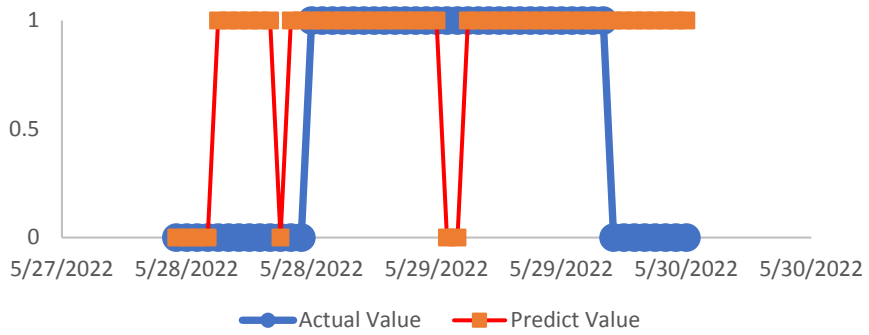
همانطور که در جدول (۱۳) مشخص شده است، دقت مدل ۸۱/۳۰ درصد می‌باشد که نسبتاً به تمامی مدل‌ها (RF، LSTM، RNN) بیشتر است. صحت مدل در پیش‌بینی کف‌ها ۸۵ درصد و در سقف‌ها ۷۸ درصد می‌باشد. شاخص پوشش ۷۹ درصد برای کف‌ها و ۸۴ درصد سقف‌ها می‌باشد. شاخص F1 نیز ۸۲ درصد برای وضعیت کف و ۸۱ درصد برای وضعیت سقف می‌باشد. این اعداد نشان‌دهنده برتری مدل طراحی شده این پژوهش می‌باشد. ماتریس اغتشاش برای مدل یادگیری جمعی در شکل ۵ مشخص شده است.



شکل ۵. ماتریس اغتشاش Ensemble روی داده تست

منبع: یافته‌های پژوهش

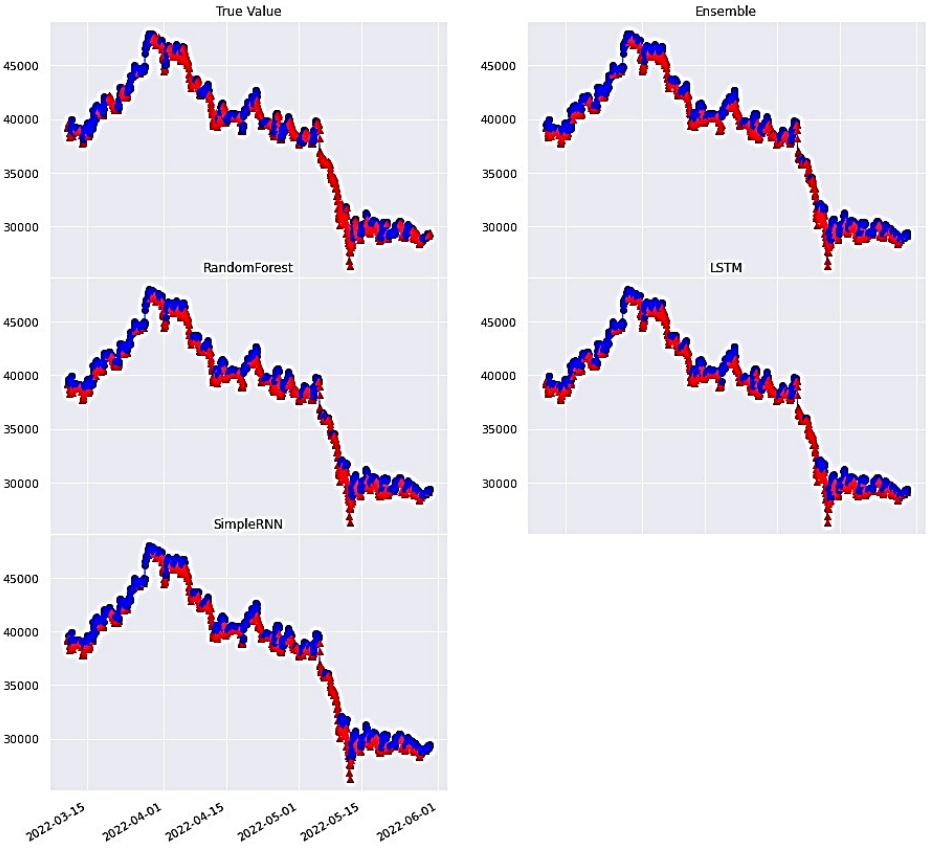
در نمودار (۷) مقادیر واقعی وضعیت سقف و کف بیت‌کوین برای ۱۰۰ داده آخر در کنار مقادیر پیش‌بینی شده مدل یادگیری جمعی ارائه شده است. خط مشکی رنگ مقادیر پیش‌بینی شده و خط قرمز رنگ مقادیر واقعی می‌باشد. همانطوری که مشخص است بسیار از وضعیت ها (سقف یا کف) به صورت درست پیش‌بینی شده‌اند.



نمودار ۶. نمودار مقادیر واقعی سقف و کف در مقایسه با مقادیر پیش‌بینی شده توسط مدل یادگیری جمعی

منبع: یافته‌های پژوهش

در نمودار (۷) می‌توان قدرت پیش‌بینی هر یک از الگوریتم ها در مقابل وضعیت های سقف و کف واقعی به صورت چشمی مورد مقایسه قرار داد.



نمودار ۷. مقایسه عملکرد پیش‌بینی مدل‌های یادگیری جمعی، جنگل تصادفی، حافظه کوتاه و بلند مدت و شبکه عصبی بازگشتی در پیش‌بینی وضعیت سقف و کف واقعی

منبع: یافته‌های پژوهش

در نمودار (۸)، نمودار با عنوان True Value قیمت به بیت‌کوین به همراه نشانه گذاری وضعیت سقف و کف واقعی ارائه شده است. وضعیت سقف با رنگ آبی و وضعیت کف با رنگ قرمز مشخص شده است. مدل‌های پیش‌بینی سعی در پیش‌بینی این وضعیت ها (سقف‌ها و کف‌ها) دارند. تمامی مدل‌ها با عنوآن‌های Ensemble، Random Forest، LSTM و SimpleRNN در نمودار (۸) با وضعیت های پیش‌بینی شده (سقف آبی و کف قرمز) مشخص شده است. این نمودار نشان‌دهنده عملکرد مناسبی تمامی مدل‌ها در پیش‌بینی سقف و کف می‌باشد.

در جدول (۱۴) شاخص‌های ارزیابی تمامی مدل‌ها به صورت تجمیع شده در کنار هم ارائه شده است.

جدول ۱۴. مقایسه شاخص ارزیابی مدل‌ها به صورت تجمیعی

ردیف	نام مدل	سقف / کف	دقت	صحت	پوشش	F1-Score
۱	RNN	۰	٪۷۹	۰/۸۳	۰/۷۶	۰/۷۹
		۱		۰/۷۶	۰/۸۲	۰/۷۹
۲	LSTM	۰	٪۸۰	۰/۸۴	۰/۷۷	۰/۸۱
		۱		۰/۷۷	۰/۸۴	۰/۸۰
۳	RF	۰	۸۰,۸۳ ٪	۰/۸۳	۰/۸۰	۰/۸۱
		۱		۰/۷۹	۰/۸۲	۰/۸۰
۴	Ensemble	۰	۸۱,۳۰ ٪	۰/۸۵	۰/۷۹	۰/۸۲
		۱		۰/۷۸	۰/۸۴	۰/۸۱

منبع: یافته‌های پژوهش

جدول (۱۴) نشان‌دهنده تقویت مدل‌ها در مدلی ۴ است که در این پژوهش ارائه شده است. Ensemble توانسته با شاخص F1 ۸۲ درصد برای وضعیت کف و ۸۱ درصد برای وضعیت سقف از تمامی مدل‌های دیگر عملکرد بهتری را ارائه کند. بنابراین، به عنوان نتیجه می‌توان اعلام کرد که مدل Ensemble در این پژوهش عملکرد (دقت، صحت، پوشش و F1) بهتری نسبت به مدل‌های RF، RNN و LSTM در پیش‌بینی وضعیت سقف و کف بیت‌کوین دارد.

سیستم معاملاتی مبتنی بر مدل‌های پیش‌بینی و داده واقعی با توجه به اینکه مقادیر ۰ و ۱ را اتخاذ می‌کنند، بر اساس همین مقادیر معامله نیز خواهند کرد. به این معنی که وقتی عدد ۰ به ۱ تبدیل می‌شود سیگنال خرید و وقتی عدد۱ به ۰ تبدیل می‌شود سیگنال فروش می‌باشد. مقدار سرمایه اولیه هزار دلار در نظر گرفته شده است. همچنین، برای هر معامله عدد ثابت ۱۰ درصد از سرمایه در نظر گرفته شده است. نتایج عملکرد هر یک از سیستم های معاملاتی در ادامه ارائه شده است.

در نمودار (۹) بازه تستی که در آن سیستم های معاملاتی تست شده‌اند مشخص شده است:

نمودار ۸. مودار بیت‌کوین به تتر در بازه زمانی تست

منبع: یافته‌های پژوهش

همانطوری که در نمودار (۹) مشخص است بیت‌کوین پس از ۲۶ درصد رشد از قیمت ۳۸۱۵۰ دلار به ۴۸۲۵۳ دلار رسیده و پس از آن محدوده به محدوده ۳۷۵۰۰ دلار رسیده است.

عملکرد هر یک از سیستم های معاملاتی در جدول زیر بر اساس شاخص‌های عملکردی را می‌توان مشاهده کرد:

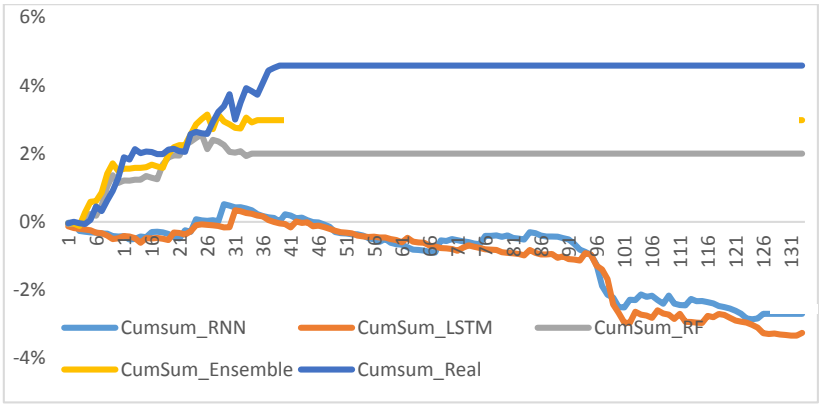
جدول ۱۵. مقایسه شاخص‌های عملکردی سیستم های معاملاتی مبتنی بر مدل‌های پیش‌بینی و داده‌های واقع

ردیف	نام سیستم معاملاتی	تعداد معامله	بازدهی (٪)	افت سرمایه (٪)	درصد موفقیت
۱	مبتنی بر مدل RNN	۱۲۶	-٪۲/۷۰	٪۳/۳۶	٪۲۳
۲	مبتنی بر مدل LSTM	۱۳۲	-٪۳/۲۵	٪۳/۶۶	٪۲۱/۸
۳	مبتنی بر مدل RF	۳۴	٪۲	٪۰/۶۲	٪۵۸
۴	مبتنی بر مدل یادگیری جمعی پیشنهاد شده	۳۵	٪۲/۹۸	٪۰/۴۲	٪۶۰
۵	مبتنی بر داده‌های واقعی سقف و کف	۳۹	٪۴/۵۹	٪۰/۷۱	٪۵۹

منبع: یافته‌های پژوهش

همان‌طور که در جدول ۱۵ مشخص شده است، سیستم معاملاتی مبتنی بر مدل یادگیری جمعی پیشنهاد شده در این پژوهش، توانسته است بازدهی بیشتر از سایر مدل‌ها و نزدیک به مدل مبتنی بر داده‌های واقعی سقف و کف را ایجاد کند. این سیستم معاملاتی توانسته است بازدهی ۲/۹۸ درصد را با شاخص افت سرمایه ۰٫۴۲٪ ایجاد کند. شاخص افت سرمایه این سیستم از استراتژی مبتنی بر داده‌های

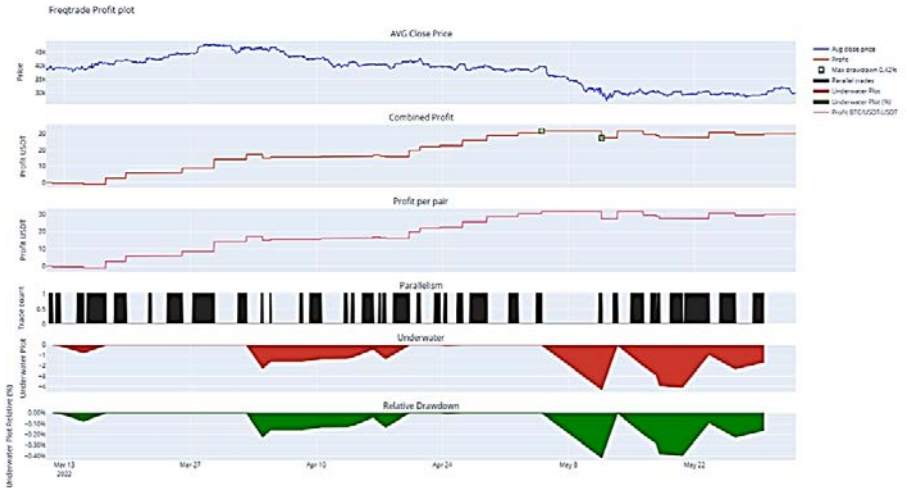
واقعی نیز کمتر است که نشان می‌دهد میزان سرمایه و ریسک هر معامله را برای استراتژی یادگیری جمعی پیشنهادی افزایش داد. در نمودار زیر بازدهی تجمعی هر یک از سیستم‌ها را می‌توان با هم مقایسه نمود:



نمودار ۹. نمودار مقایسه بازدهی تجمعی سیستم‌های معاملاتی مبتنی بر RNN، LSTM، RF و یادگیری جمعی پیشنهادی به همراه داده واقعی سقف و کف

منبع: یافته‌های پژوهش

در نمودار ۱۱ عملکرد سیستم معاملاتی مبتنی بر یادگیری جمعی مشخص شده است:



نمودار ۱۰. عملکرد سیستم معاملاتی مبتنی بر یادگیری جمعی پیشنهادی

منبع: یافته‌های پژوهش

بحث و نتیجه گیری

در بخش قبل، خروجی و شاخص‌های ارزیابی به صورت مستقل ارائه و بررسی شد. این بخش به دنبال پاسخ شفاف به سئوالات پژوهش می‌باشد. با توجه به نقاط قوت و قابل بهبود هر یک از مدل‌های یادگیری ماشین در پیش‌بینی می‌توان گفت در صورتی که مدلی به عنوان مدل ترکیبی خروجی‌های مدل‌ها را تجمیع و تصمیم گیری کند می‌تواند عملکرد بهتری در پیش‌بینی داشته باشد. در این پژوهش، نتایج مدل‌های شبکه عصبی بازگشتی، یادگیری حافظه کوتاه و بلندمدت و جنگل تصادفی به عنوان ورودی به یک ماژول یادگیری جمعی ارائه شد. ماژول یادگیری جمعی با استفاده از مدل‌های یادگیری ماشین XGBoost و LightGBM خروجی‌های مدل‌های اشاره شده را با الگوریتم یادگیری جمعی مبتنی بر رای‌گیری ترکیب می‌کند. فرضیه این پژوهش تقویت دقت پیش‌بینی مدل ارائه شده نسبت به هر یک از زیر الگوریتم‌ها می‌باشد. با با توجه به خروجی‌های ارائه شده در بخش اجرای مدل‌ها، این فرضیه مورد پذیرش قرار گرفت. بر اساس جدول (۱۴) ماژول یادگیری جمعی توانسته است بر اساس شاخص‌های دقت، صحت، پوشش و F1 عملکرد مناسب‌تری نسبت به هر یک از زیر الگوریتم ها در زمان مشابه نشان دهد. بنابراین، سوالات این پژوهش را می‌توان به صورت زیر پاسخ داد:

۱- سیستم معاملاتی هوشمند و الگوریتمی مبتنی بر یادگیری جمعی پیشنهاد شده این پژوهش توانست نزدیک تری بازدهی را به بازدهی سیستم معاملاتی مبتنی بر داده‌ها واقعی سقف و کف را ایجاد کند. همچنین ریسک (شاخص افت سرمایه) آن نسبت به سیستم معاملاتی مبتنی بر داده‌های واقعی کمتر و از ما بقی سیستم ها بهتر بوده است. (جدول ۱۵)

۲- میزان دقت و صحت پیش‌بینی مدل یادگیری جمعی در پیش‌بینی وضعیت سقف و کف قیمت بیت‌کوین به ترتیب ۸۱/۳۱٪ و ۸۲٪ می‌باشد. (جدول ۱۴)

۳- میزان دقت و صحت مدل یادگیری جمعی بر اساس شاخص‌های دقت، صحت، پوشش و F1 نسبت به تمامی زیر الگوریتم ها بیشتر است (جدول ۱۴).

در پژوهش حاضر دقت پیش‌بینی توسط این مدل به ۸۱/۳۱ درصد افزایش یافته است این در حالی است که در پژوهش‌ها دقت این مدل ۵۶ درصد (بشیری و پاریاب، ۱۳۹۹) و ۶۹ درصد (مشاری و همکاران، ۱۳۹۸) بوده است. همچنین، نتیجه این پژوهش با پژوهش باساک[1] و همکاران (۲۰۱۹) مطابقت دارد. زیرا در این پژوهش نیز، مدل RF با مدل‌های XGBoost، ANN، SVM و رگرسیون لاجستیک مورد مقایسه قرار گرفته و نشان داده شده است مدل RF از دقت بیشتری برخوردار است.

از طرفی بهبود بازدهی و ریسک سیستم معاملاتی در بازه ۱ ساعته با در نظر گرفتن نگاه گسسته با پژوهش جاکارت[2] و همکاران مطابقت دارد. اما مدل پیشنهادی این پژوهش جنبه نوآوری این پژوهش هست و کاملا با رویکرد این پژوهش متفاوت است. اما از این نظر که سبب بهبود استراتژی در تایم فریم ۱ ساعته شده است تطابق هایی وجود دارد.

1. Basak
2. Jaquart

پیشنهادهای این پژوهش عبارتند از:

- استفاده از داده چند دارایی به طول همزمان و یکپارچه‌سازی: به این معنی که به جای اینکه مدل صرفا روی داده‌های بیت‌کوین آموزش دیده و پیش‌بینی نماید می‌تواند به طور همزمان چند دارایی در یک بازار (اتریوم، لایت کوین و غیره) یا حتی چند بازار (طلا، ارز، شاخص بورس آمریکا و غیره) را آموزش و سپس، پیش‌بینی نماید.
- استفاده از اندیکاتورهای دیگر برای تشخیص سقف و کف مانند ZigZag و مقایسه آن با نتیجه مدل فعلی که از اندیکاتور AO استفاده کرده است.
- اضافه کردن متغیرهای بنیادی[1] و ادراکی[2] بازار به عنوان متغیرهای ویژگی با هدف افزایش دقت پیش‌بینی می‌تواند کمک بسزایی کند. همچنین مقایسه مدل‌ها و تاثیر اضافه کردن این داده‌ها می‌تواند به نتایج دیگری پژوهش کمک کند.
- مقایسه عملکرد مدل در تایم فریم‌های متفاوت

محدودیت‌های پژوهش عبارتند از:

- شرط توقف همراه با تایم فریم پایین تر به دلیل پیچیدگی و همچنین، نیاز به محاسبات بالا نیاز به سخت افزار با مشخصات نسبتاً بالایی دارد.
- چند دارایی به طور همزمان با توجه به پایگاه داده حجیم و فرخوانی و پردازش آن نیاز به سخت‌افزار با مشخصات نسبتاً بالایی دارد.
- تفاوت جزئی قیمت بیت‌کوین در صرافی‌های متفاوت می‌تواند روی عملکرد و خروجی مدل تاثیرگذار باشد.

ملاحظات اخلاقی

حامی مالی: مقاله حامی مالی ندارد.
مشارکت نویسندگان: تمام نویسندگان در آماده‌سازی مقاله مشارکت داشته‌اند.
تعارض منافع: بنا بر اظهار نویسندگان در این مقاله هیچ‌گونه تعارض منافعی وجود ندارد.
تعهد کپی‌رایت: طبق تعهد نویسندگان حق کپی‌رایت رعایت شده‌است.

1. Fundamental
2. Sentimental

References

Aggarwal, D; Chandrasekaran, S; & Annamalai, B. (2020). A complete empirical ensemble mode decomposition and support vector machine-based approach to predict Bitcoin prices. *Journal of Behavioral and Experimental Finance*, 27, Article 100335. https://doi.org/10.1016/j.jbef.2020.100335

Aghakhani, K; & Karimi, A. (2014). Investigating methods for predicting stock prices in the stock market and introducing an optimized method. *In National Conference on New Achievements in Engineering and Basic Sciences* (pp. 1-7). Ardabil. (In Persian)

Ampomah, E; Qin, Z; & Nyame, G. (2020). Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement. *Information*, 11(6), 332. https://doi.org/10.3390/info11060332

Basak, S; Kar, S; Saha, S; Khaidem, L; & Dey, S. (2018). Predicting the direction of stock market prices using tree-based classifiers. The North American *Journal of Economics and Finance*, 47, Article 102-117. https://doi.org/10.1016/j.najef.2018.06.013

Faghihi Nezhad, M; & Minaei Bidgoli, B. (2021). Development of an ensemble learning-based intelligent model for stock market forecasting. Scientia Iranica, 28(1), 395-411. https://doi.org/10.24200/sci.2019.50353.1654

Gupta, A; & Nain, H. (2020). Bitcoin price prediction using time series analysis and machine learning techniques. In A. Das, K. Deep, A. Tiwari, R. Kumar, & R. Bansal (Eds.), Intelligent Communication, Control and Devices (pp. 533-542). Springer. https://doi.org/10.1007/978-981-15-7106-0_54

Jiang, M; Liu, J; Zhang, L; & Liu, C. (2020). An improved stacking framework for stock index prediction by leveraging tree-based ensemble models and deep learning algorithms. Physica A: Statistical Mechanics and its Applications, 541, Article 123390.

Li, Y; & Pan, Y. (2022). A novel ensemble deep learning model for stock prediction based on stock prices and news. *International Journal of Data Science and Analytics*, 13(1), 139-149. https://doi.org/10.1007/s41060-021-00279-9

Mallqui, D. C. A; & Fernandes, R. A. S. (2019). Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. Applied Soft Computing, 75, 596-606. https://doi.org/10.1016/j.asoc.2018.11.038

Mohammadi Almoti, M; Haddadi, M. R; & Nadimi, Y. (2018). Modeling and evaluating different prediction models: Short-term memory, long-term memory, Markov switching, and hyperbolic tangent in predicting crude oil price fluctuations of OPEC. *Financial Engineering and Securities Management Quarterly*, 9(34), 249-272. (In Persian)

Moshari, M; et al. (2019). Designing a combined intelligent model for predicting golden points of stock prices. *Investment Knowledge Quarterly*, 8(29), 45-65. (In Persian)

Mousavi, S. O. (2019). Algorithmic Trading Technology: New Solutions in the Iranian Capital Market. Challenge Publications. (In Persian)

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Decentralized Business Review, 21260.

Nti, I. K; Adekoya, A. F; & Weyori, B. A. (2019). Random forest-based feature selection of macroeconomic variables for stock market prediction. *American Journal of Applied Sciences*, 16(7), 200-212. https://doi.org/10.3844/ajassp.2019.200.212

Nti, I. K; Adekoya, A. F; & Weyori, B. A. (2020). A comprehensive evaluation of ensemble learning for stock-market prediction. *Journal of Big Data*, 7(1), Article 20. https://doi.org/10.1186/s40537-020-00299-5

Shirzoor Aliabadi, Z; & Ramazanzadeh, H. (2020). Bitcoin and its future prediction. Sixth International Conference on Industrial Engineering and Systems, Mashhad. (In Persian)

Tehrani, E; & Esmaili, A. (2012). Investigating the effect of using important technical analysis indicators on short-term investors' returns in the Tehran Stock Exchange. *Financial Knowledge-Based Securities Analysis*, 5(13), 21-33. (In Persian)

Yun, K. K; Yoon, S. W; & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications*, 186, Article 115716. https://doi.org/10.1016/j.eswa.2021.115716

Yang, Y; Wu, Y; Wang, P; & Jiali, X. (2021). Stock price prediction based on XGBoost and LightGBM. E3S Web of Conferences, 275, Article 01040. https://doi.org/10.1051/e3sconf/202127501040

COPYRIGHTS

This license allows others to download the works and share them with others as long as they credit them, but they can't change them in any way or use them commercially.