

**موسسه آموزش عالی زند شیراز**

دانشكده مهندسـي – نرم افزار کامپیوتر

**طرح پیشنهادی پايان‌نامه كارشناسي ارشد**

***ترکیب تحلیل احساسات و رویکرد مبتنی بر شبکه های عصبی بازگشتی برای تخمین و تحلیل رمزارزها***

***The combination of sentiment analysis and an approach based on recurrent neural networks for the estimation and analysis of cryptocurrencies***

**توسط :**

**علی منصورآبادی 400430307**

استاد / اساتید راهنما:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **نام و نام خانوادگی** | **دانشکده** | **رتبه علمی** | **رشته – گرایش** | **امضاء** |
| امیرحسین راسخ | مهندسی | استادیار | مهندسی کامپیوتر - نرم افزار |  |

استاد / اساتید مشاور:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **نام و نام خانوادگی** | **دانشکده** | **رتبه علمی** | **رشته – گرایش** | **امضاء** |
| علی قرمزیان | مهندسی | استادیار | علوم اعصاب شناختی - رایانش و هوش مصنوعی |  |

**تاریخ:**

**26/02/1403**

|  |
| --- |
| 1. **مقدمه (تشریح و بیان مسئله):** |

ارزهای رمزنگاری شده[[1]](#footnote-1)، ارزهای مجازی[[2]](#footnote-2) هستند که بر اساس فناوری زنجیره بلوکی[[3]](#footnote-3) کار می‌کنند. از زمان معرفی بیت‌کوین[[4]](#footnote-4) در سال ۲۰۰۹، ارزهای دیجیتال به‌صورت گسترده‌ای در بازارهای مالی دیده‌ می‌شوند. افراد در زمینه‌های متفاوتی از جمله تجارت ارزهای رمزنگاری و سرمایه‌گذاری در صندوق‌ها و شرکت‌های مربوطه فعالیت می‌کنند؛ طبق گزارش کوین‌مارکت‌کپ[[5]](#footnote-5)، ارزش بازار جهانی ارزهای رمزنگاری شده در سپتامبر ۲۰۲۲، به مقدار تخمینی ۹۳۲.۴۹ میلیارد دلار رسیده است. اگرچه سرمایه‌گذاری‌ها به بازدهی پرسود انجامیده‌اند، اما نوسانات قیمتی پراکنده در اکثر ارزهای رمزنگاری، چالش‌ها و ریسک‌هایی را برای این بازار ایجاد کرده‌اند. به عنوان مثال، قیمت بیت‌کوین از زمان عرضه در بازار بسیار نوسانات داشته است و در سال‌های ۲۰۱۶ و ۲۰۱۷ اوج‌های +۱۲۲٪ و +۱۳۶۰٪ را به دست آورده است. اتریوم[[6]](#footnote-6)، ایکس‌آرپی[[7]](#footnote-7) و لایت‌کوین[[8]](#footnote-8) نیز در سال ۲۰۱۷ نوسانات مشابهی را تجربه کرده‌اند [1].

یک نکته مهم درباره ارزهای رمزنگاری شده این است که آن‌ها مستقل از بانک‌های سنتی هستند و توسط هیچ مرکزی کنترل نمی‌شوند، که این امر آن‌ها را از ارزهای سنتی متمایز می‌کند. از آنجایی که وجود زنجیره بلوکی برای ارزهای رمزنگاری لازم است، ادعا می‌شود که این ارزها همه ویژگی‌های زنجیره بلوکی را به ارث می‌برند. به عنوان مثال، بیت‌کوین به مردم راهی امن برای انجام معاملات دیجیتال به طور نیمه‌ناشناس[[9]](#footnote-9) فراهم می‌کند. فناوری زنجیره بلوکی توجه دولت‌های سراسر جهان را به خود جلب کرده است و انگیزه پشت این توجه، از نگرانی‌های مرتبط با جرائم نشات می‌گیرد. با این حال، شبکه بیت‌کوین بر روی الگوریتم‌های اثبات کار[[10]](#footnote-10) و اثبات سهم[[11]](#footnote-11) کار می‌کند که مصرف انرژی بالایی در فرآیندهای محاسباتی خود برای امن‌سازی شبکه می‌طلبند [2].

اثبات‌کار الگوریتم توافقی است که در برخی از سیستم‌های زنجیره بلوکی مانند بیت‌کوین استفاده می‌شود که از کاربران می‌خواهد مقداری مشخص از کار محاسباتی را انجام دهند تا معاملات را تأیید کرده و آن‌ها را به زنجیره بلوکی اضافه کنند. این کار، که به عنوان استخراج شناخته می‌شود، معمولاً با استفاده از سخت‌افزارهای ویژه مانند مدارهای یکپارچه خاص برنامه‌ریزی شده[[12]](#footnote-12) انجام می‌شود و مقدار قابل توجهی از انرژی را مصرف می‌کند. یک نقطه ضعف این سیستم‌ها این است که غالباً به خاطر مصرف انرژی بالا و متمرکز کردن قدرت استخراج مورد انتقاد قرار می‌گیرند. مکانیزم‌های جایگزین مانند اثبات‌سهم نیز هستند که نیازی به انجام کار محاسباتی برای تأیید معاملات ندارند و به جای این کار، کاربران مجبور هستند مقدار معینی از دارایی مورد نظر را نگه دارند. مزیت این‌گونه سیستم‌ها این است که با مصرف انرژی کمتر نسبت به سیستم‌های اثبات‌کار، گزینه‌ موثرتری خواهند‌بود [3].

سرمایه‌گذاران به یک رویکرد پیش‌بینی نیاز دارند تا بتوانند نوسانات قیمت ارزهای رمزنگاری شده را به‌طور موثری تحلیل کرده و ریسک را کاهش و سود را افزایش دهند. علاوه بر این، ممکن است از پیش‌بینی‌های نوسانات برای تخمین تغییرات قیمت آنها استفاده شود که برای توسعه و تحلیل استراتژی‌های مالی مفید است. با این حال، همانند پیش‌بینی قیمت سهام، که بازار آن پویا[[13]](#footnote-13) و پیچیده است، پیش‌بینی قیمت ارزهای رمزنگاری شده نیز در حال حاضر به عنوان یکی از دشوارترین وظایف پیش‌بینی در حوزه مالی   
محسوب می‌شود [4].

اکثر پژوهشگران، مسئله پیش‌بینی ارزهای دیجیتال را به عنوان یک مثال از پیش‌بینی سری زمانی[[14]](#footnote-14) مطرح می‌کنند، زیرا که ایده اصلی این است که از داده‌های قیمت تاریخی[[15]](#footnote-15) استفاده کرد تا قیمت‌های آینده را در یک دوره زمانی خاص در آینده پیش‌بینی کرد. تحلیل سری زمانی همچنین در پیش‌بینی هواشناسی و پیش‌بینی تقاضای خرده‌فروشی و تامین نیز استفاده شده است [1].

رویکردهای یادگیری ماشین[[16]](#footnote-16) می‌توانند الگوهای غیرخطی را استخراج کنند که در پیش‌بینی قیمت ارز‌های دیجیتال بسیار مفید خواهد ‌بود. با این حال، روش‌های سنتی یادگیری ماشین[[17]](#footnote-17)، مانند   
شبکه‌های عصبی چندلایه[[18]](#footnote-18) [5] و ماشین‌های بردار پشتیبان[[19]](#footnote-19) [6] از برخی مشکلات مانند حساسیت به بیش‌برازش[[20]](#footnote-20) رنج می‌برند و در استخراج الگوهای پنهان سطح بالا[[21]](#footnote-21) از داده‌های توالی ارزهای رمزنگاری خوب عمل نمی‌کنند. برای رفع این مشکلات، مدل‌های پیش‌بینی مبتنی بر یادگیری عمیق[[22]](#footnote-22) استفاده شده‌اند که عملکرد بهتری نسبت به روش‌های سنتی یادگیری ماشین را دارند [7]. در کارهای اخیر نیز این موضوع اثبات شده است که شبکه‌های عصبی حافظه طولانی کوتاه مدت[[23]](#footnote-23) و واحدهای بازگشتی دروازه‌ای[[24]](#footnote-24) عملکرد به نسبت بهتری در پیش‌بینی قیمت ارزهای دیجیتال ارائه می‌دهند [1].

در تحقیقات دیگر بر روی روش‌های یادگیری گروهی[[25]](#footnote-25) برای کاهش واریانس[[26]](#footnote-26) و انحراف، با ترکیب مجموعه مدل‌های یادگیری ضعیف[[27]](#footnote-27) مختلف تمرکز داشته‌اند [8]. همچنین محققان به این نتیجه رسیده‌اند که مدل‌های یادگیری گروهی، عملکرد بهتری نسبت به مدل‌های یادگیری ماشین به صورت تکی دارند و اینکه مدل بوستینگ گرادیان[[28]](#footnote-28) دقت بهتری نسبت به روش جنگل تصادفی[[29]](#footnote-29) دارد [9].

|  |
| --- |
| 1. **مروری بر کارهای انجام شده:** |

اسلم و همکارانش در سال 2022 در [10] برای تشخیص، تجزیه و تحلیل احساسات درباره توییت‌های مرتبط با رمزارزها، تحقیقات خود را انجام داده‌اند. تجزیه و تحلیل احساسات در رمزارزها اهمیت زیادی دارد زیرا به طور گسترده برای پیش‌بینی قیمت بازار رمزارز استفاده می‌شود و نیازمند طبقه‌بندی با دقتی بالا هستند. برای تحلیل احساسات، توییت‌ها از توییتر استخراج می‌شوند و مجموعه داده با استفاده از دو ابزار تکست‌بلاب[[30]](#footnote-30) و تکست‌تو‌ایموشن[[31]](#footnote-31) برچسب‌گذاری می‌شود. همچنین چندین مدل یادگیری ماشین و یادگیری عمیق برای طبقه‌بندی احساسات استفاده شده‌اند. این مطالعه از حافظه طولانی کوتاه مدت و واحد بازگشتی دروازه‌ای برای ایجاد یک مدل استفاده می‌کند تا عملکرد طبقه‌بندی را بهبود بخشد. یافته‌ها نشان می‌دهد که ترکیب حافظه طولانی کوتاه مدت و واحد بازگشتی دروازه‌ای در مقایسه با سایر مدل‌ها از نظر پیش‌بینی، عملکرد بهتری دارد.

چن در سال 2023 در مقاله [11] برای پیش‌بینی قیمت بیت‌کوین روز بعد، متغیرهای قیمت بیت‌کوین، ویژگی‌های فنی خاص[[32]](#footnote-32) بیت‌کوین، رمزارزهای دیگر، کالاها[[33]](#footnote-33)، شاخص بازار[[34]](#footnote-34)، ارز خارجی [[35]](#footnote-35)و توجه عموم[[36]](#footnote-36) را به عنوان ورودی استفاده کرده‌است. طبق این مطالعه، رگرسیون جنگل تصادفی[[37]](#footnote-37) دقت پیش‌بینی بهتری نسبت به حافظه طولانی کوتاه مدت دارد. در تحقیقات قبلی، حافظه طولانی کوتاه مدت به عنوان یک الگوریتم با دقت بالا در پیش‌بینی قیمت بیت‌کوین شناخته و به صورت گسترده‌ای استفاده شده بود. در اصل این مقاله از الگوریتم یادگیری ماشین رگرسیون جنگل تصادفی استفاده می‌کند، که تاکنون به طور گسترده توسط سایر پژوهشگران در ادبیات پیشین استفاده نشده است و به این نتیجه رسیده شده که این مدل عملکرد به نسبت بهتری از خود نشان می‌دهد.

سیبه و همکارانش در سال 2023 طبق مطالعه [2]، از سه نوع تکنیک یادگیری عمیق به نام‌های حافظه طولانی کوتاه مدت، واحد بازگشتی دروازه‌ای و حافظه طولانی کوتاه مدت دوطرفه[[38]](#footnote-38) برای پیش‌بینی قیمت‌های سه ارز دیجیتال بیت‌کوین، اتریوم و لایت‌کوین استفاده کرده‌اند. عملکرد مدل‌ها با استفاده از دو معیار اندازه‌‎گیری ریشه میانگین مربعات خطا[[39]](#footnote-39) و درصد مطلق خطای متوسط [[40]](#footnote-40)​​ارزیابی شده‌است. نتایج مطالعه نشان داد که مدل حافظه طولانی کوتاه مدت دو طرفه پیش‌بینی‌های دقیق‌تری را برای هر سه ارز ارائه می‌دهد و پس از آن مدل واحد بازگشتی دروازه‌ای قرار دارد. ترکیب جریان‌های رو به جلو و رو به عقب در مدل‌ دوطرفه، عملکرد و پیش‌بینی‌ در داده‌های سری زمانی را بهبود می‌بخشد. استنتاج نهایی این مطالعه این است که الگوریتم‌های یادگیری عمیق در پیش‌بینی قیمت‌ رمزارزها موثر هستند و در مطالعات آینده، تأثیر توییت‌ها[[41]](#footnote-41) و احساسات[[42]](#footnote-42) بر قیمت‌های رمزارزها با استفاده از تکنیک‌های یادگیری ماشین بررسی خواهد شد.

شمشاد و همکارانش در سال 2023 در مقاله [12] یک سیستم پیش‌بینی ارزدیجیتال ارائه داده‌اند که شش تکنیک پیشرفته یادگیری ماشین و یادگیری عمیق شامل ماشین‌های بردار پشتیبان، میانگین متحرک خودهمبسته یکپارچه[[43]](#footnote-43)، پرافت فیس‌بوک[[44]](#footnote-44)، حافظه طولانی کوتاه مدت، حافظه طولانی کوتاه مدت دو طرفه و حافظه طولانی کوتاه مدت به صورت پشته‌ای[[45]](#footnote-45) پیاده‌سازی شده‌اند. برای ارزیابی کامل هر الگوریتم، از سه مجموعه داده ارزدیجیتال کاردانو[[46]](#footnote-46)، بی‌ان‌بی[[47]](#footnote-47) و اتریوم استفاده شده است. هر مدل بر روی هر مجموعه داده آموزش داده شده است تا قیمت 24 ساعته ده روز آینده را پیش‌بینی کند. برای بهبود ارزیابی، نمونه‌های آموزش[[48]](#footnote-48) و آزمون[[49]](#footnote-49) هر ارز برای همه مدل‌ها ثابت نگه داشته شده است و عملکرد آن‌ها با استفاده از تکنیک‌های ارزیابی مختلف مانند میانگین مطلق خطا[[50]](#footnote-50)، ریشه میانگین مربعات خطا، درصد مطلق خطای متوسط، ضریب تعیین[[51]](#footnote-51) و میانگین مربعات خطا[[52]](#footnote-52) ارزیابی شده است. نتایج این مطالعه نشان می‌دهد که مدل میانگین متحرک خودهمبسته یکپارچه در پیش‌بینی ارزها برای ده روز آینده، عملکرد بهتری نسبت به مدل‌های دیگر ارائه می‌دهد. از طرف دیگر، پرافت فیس‌بوک در تغییرات و نواسانات کم، عملکرد نسبتا خوبی ارائه داده‌است. این مدل با ‌دقت خوبی روندهای کاردانو و بی‌ان‌بی را پیش‌بینی کرده، اما روندهای اتریوم را به‌طور نادرستی پیش‌بینی کرده است.

اویدله و همکارانش در سال 2023 در مطالعه [13] مدل‌های یادگیری عمیق و روش‌های مبتنی‌بر درخت افزایشی[[53]](#footnote-53) را مقایسه کرده و با استفاده از پیکربندی‌ بهینه مدل‌ها، قیمت‌های پایانی روزانه شش رمزارز مختلف را پیش‌بینی کرده‌اند. براساس یافته‌های حاصل از این مطالعه، می‌توان به این نتیجه رسید که تکنیک‌های یادگیری عمیق نسبتاً درصد بالاتری از واریانس[[54]](#footnote-54) کلی در داده‌ها را دریافت می‌کنند (بین 88٪ و 98٪). با این حال روش‌های مبتنی‌بر درخت افزایشی، دچار مشکلاتی در پیش‌بینی قیمت‌های پایانی روزانه برای رمزارزهای اتریوم، بی‌ان‌بی و دوج[[55]](#footnote-55) هستند. مدل شبکه عصبی پیچشی[[56]](#footnote-56) نتایج دقیق‌تری نسبت به سایر تکنیک‌های عمیق یادگیری ارائه می‌دهد و بعد از این مدل، واحد بازگشتی دروازه‌ای و سپس ماشین افزایشی گرادیان[[57]](#footnote-57) قرار دارند.

مشارکت اصلی رائو و همکارانش در سال 2023 طبق مطالعه [14]، بررسی استفاده از روش‌های یادگیری عمیق نوین برای پیش‌بینی ارزش و رفتار آینده رمزارزها است. در این مطالعه، مدل‌های یادگیری عمیق ترکیبی مبتنی بر لایه‌های حافظه طولانی کوتاه مدت، حافظه طولانی کوتاه مدت دو طرفه و شبکه عصبی پیچشی استفاده شده‌اند. بر اساس یافته‌های این مقاله می‌توان به این نتیجه رسید که ترکیب لایه های پیچشی و حافظه طولانی کوتاه مدت دوطرفه توانسته بهترین عملکرد را از خودش نشان دهد. در آینده قصد نویسندگان مقاله، ایجاد یک سیستم پشتیبان تصمیم‌گیری[[58]](#footnote-58) برای پیش‌بینی رمزارزها است که دقیق و قابل اعتماد باشد. همچنین این ایده نیز بررسی شده‌ است که می‌توان یک سیستم هوش مصنوعی مبتنی بر چارچوب تشخیص ناهنجاری[[59]](#footnote-59) ساخت تا از الگوریتم‌های بدون نظارت[[60]](#footnote-60) برای یافتن نقاط ناهنجار یا سیگنال‌های نادر استفاده کند.

موری و همکارانش در سال 2023 در مقاله [1] مقایسه‌ای بین مدل‌های یادگیری عمیق، یادگیری ماشین و آماری برای پیش‌بینی قیمت‌های روزانه ارزهای رمزنگاری شده ارائه می‌دهند. نتایج نشان می‌دهد که به طور کلی رویکردهای عمیق بازگشتی[[61]](#footnote-61)، بهترین مدل‌ها برای این کار یعنی پیش‌بینی ارزهای دیجیتال هستند. به طور خاص، حافظه طولانی کوتاه مدت بهترین عملکرد را دارد. همچنین لازم به ذکر است که مدل k نزدیک ترین همسایه[[62]](#footnote-62) نیز تعادل خوبی بین دقت و هزینه محاسباتی ارائه می‌دهد. در آینده نیز نویسنگان مقاله قصد دارند مدل‌هایی را ارائه دهند که قیمت‌ را به صورت ساعتی پیش‌بینی‌ کند و همچنین احساسات بازار و نوسانات رمزارزهای دیگر را نیز مورد بررسی قرار دهند.

تریپاتی و شارما در سال 2023 در مقاله [15] معماری‌های شبکه عصبی عمیق را برای پیش‌بینی معامله کوتاه‌مدت بیت‌کوین ارائه می‌دهند. رویکرد موردنظر از ورودی یک متغیره یعنی قیمت‌ بسته‌شده[[63]](#footnote-63) گذشته و ورودی‌های چند متغیره یعنی استفاده از شاخص‌های مالی و فنی استفاده می‌کند. مدل‌های شبکه عصبی مصنوعی عمیق[[64]](#footnote-64)، حافظه طولانی کوتاه مدت، حافظه طولانی کوتاه مدت دوطرفه و ترکیب شبکه عصبی پیچشی و حافظه طولانی کوتاه مدت در این مقاله پیاده‌سازی شده‌اند. در این مقاله داده‌های پرت ابتدا کامل مدیریت شده‌اند و سپس داده‌ها وارد مدل‌ها شده‌اند. همچنین، تأثیر طول پنجره[[65]](#footnote-65) نیز بررسی شده‌‎است تا معامله‌ای بهینه انتخاب شود. لازم به ذکر است که در این مقاله از بهینه‌سازی بیزین[[66]](#footnote-66) برای یافتن هایپرپارامترهای[[67]](#footnote-67) بهینه برای طراحی معماری شبکه‌های عصبی استفاده شده‌است. طبق نتایج گرفته شده، شبکه عصبی مصنوعی عمیق در مقایسه با سایر مدل‌ها عملکرد بهتری دارد. یافته‌ها نشان می‌دهد که هیچ شاخص فنی نمی‌تواند توان پیش‌بینی یکسانی را در تمام مراحل بازار داشته باشد. در آینده، نویسندگان مقاله قصد دارند که روش‌های بهتری برای شناسایی و مدیریت داده‌های پرت پیاده‌سازی نمایند؛ به طور مثال می‌توان از تکنیک‌های مقاوم‌تری مانند شبکه‌های عصبی رمزگذار خودکار[[68]](#footnote-68) استفاده کرد.

بوتسکا و همکارانش در سال 2024 در مطالعه [16]، مدل‌های یادگیری گروهی و یادگیری عمیق برای پیش‌بینی قیمت‌ ارزهای رمزنگاری شده ارائه داده‌اند. آزمایش‌های زیادی با استفاده از داده‌های سری زمانی تاریخی از چهار ارز رمزنگاری شده انجام شده‌است. نتایج نشان می‌دهد که عملکرد بالایی برای روش‌های پیچیده یادگیری ماشین در هر چهار ارز رمزنگاری شده وجود دارد. مدل ماشین افزایش گرادیان سبک[[69]](#footnote-69) ممکن است استراتژی‌های معاملاتی بسیار سودآوری را برای سرمایه‌گذاران در بازارهای بیت‌کوین، اتریوم و لایت‌کوین فراهم کند. برای بازار ریپل، مدل شبکه عصبی بازگشتی ساده[[70]](#footnote-70) به عنوان بهترین مدل پیش‌بینی برای سرمایه‌گذاران توصیه می‌شود. به طور کلی، این یافته‌ها نشان می‌دهند که مدل‌های یادگیری گروهی و یادگیری عمیق می‌توانند به طور موثر سرمایه‌گذاران را در تصمیمات معاملاتی خود راهنمایی کنند. در آینده، قصد نویسندگان این مقاله این است که داده‌های زمان واقعی[[71]](#footnote-71) بیشتری را بررسی کنند تا از مزایای مدل‌های یادگیری عمیق بیشتر بهره‌مند شوند.

|  |
| --- |
| 1. **اهداف پایان نامه و روش انجام کار:** |

همانطورکه بیان شد، ارزهای دیجیتال یکی از محبوب‌ترین بازارهای مالی حال حاظر هستند که روزبه‌روز به تعداد کاربران آن‌ افزوده می‌شود. به خاطر ذات و نوسانات بالا، برای تسلط به این بازار به مهارت، تجربه و زمان زیادی نیاز است. هدف این تحقیق ارائه روشی مبتنی بر یادگیری ماشین و به‌صورت دقیق‌تر یادگیری عمیق است تا در امر پیش‌بینی ارزهای دیجیتال به معامله‌گران کمک کند. لزوم انتخاب رویکرد های یادگیری ماشین این است که در وحله اول، داده موجود باشد و از کیفیت لازم برخوردار باشد. بعد از بررسی وجود داده، باید در داده، الگوهایی نیز وجود داشته باشد زیرا که الگوریتم های یادگیری ماشین در داده‌ها به دنبال الگوهایی برای یادگیری هستند که بتوانند از یادگیری خود، در داده‌های آینده استفاده کنند. به عنوان شرط آخر، الگوریتم‌های یادگیری ماشین در مسائلی باید استفاده شوند که نتوان آن‌ها را به صورت ریاضی فرموله کرد.

طبق موارد گفته شده می‌توان از الگوریتم‌های مبتنی بر یادگیری ماشین در ارزهای دیجیتال استفاده کرد، زیرا که در این بازار هر سه شرط وجود دارد. برای شرط اول، داده‌های مورد نیاز برای این تحقیق را می‌توان از سایت های زیادی نظیر کوین‌مارکت‌کپ[[72]](#footnote-72)، یاهو‌فایننس[[73]](#footnote-73)، کگل[[74]](#footnote-74) و ... بدست آورد. در داده‌های ارزهای دیجیتال همواره الگو های گذشته تکرار می‌شوند، پس در قیمت ها و حرکت بازار الگو وجود دارد. همچنین به دلیل ماهیت بازار و نوسانات ذکر شده، حرکت این بازار را نمی‌توان به صورت ریاضیاتی فرموله کرد. طبق تحقیقات انجام شده، به دلیل پیچیدگی حرکت و الگوهای ارزهای دیجیتال، بهتر است از الگوریتم‌های یادگیری عمیق استفاده شود و از متدهای یادگیری ماشین سنتی استفاده نشود؛ زیرا که روش‌های سنتی یادگیری ماشین قادر به کشف الگوهای پنهان و پیچیده ارزهای دیجیتال نیستند.

برای این تحقیق، داده‌ها از سایت کوین‌مارکت‌کپ دریافت شده‌اند که شامل ستون‌های قیمت بسته‌شدن، قیمت باز‌شدن[[75]](#footnote-75)، بیشترین قیمت[[76]](#footnote-76)، کمترین قیمت[[77]](#footnote-77)، حجم معاملات[[78]](#footnote-78)، ارزش کلی بازار[[79]](#footnote-79)، تاریخ[[80]](#footnote-80)، تاریخ باز‌شدن، قیمت بسته‌شدن، تاریخ کمترین قیمت و تاریخ بیشترین قیمت می‌شود. از میان این ویژگی‌ها و ستون‌ها، در این تحقیق صرفا از سه ستون قیمت بسته‌شدن، تاریخ و حجم معاملات استفاده شده‌است. داده‌ها از اول ماه دوم میلادی سال 2018 تا آخر ماه سوم میلادی سال 2024 جمع‌آوری شده‌اند. لازم به ذکر است که داده‌ها برای این تحقیق به صورت روزانه[[81]](#footnote-81) جمع‌آوری شده‌اند و هدف تحقیق، پیش‌بینی قیمت بسته‌شدن روز آینده با داشتن اطلاعات روز گذشته است.‌

پیش‌پردازش داده‌ها یکی از مهم‌ترین قدم‌ها در پیش‌بینی ارزهای دیجیتال است. یکی از عوامل مهم تاثیرگذار در دقت مدل، مدیریت کردن داده‌های پرت[[82]](#footnote-82) است که در این تحقیق از روش هایی نظیر دامنه میان چارکی[[83]](#footnote-83) استفاده شده‌است تا داده‌های پرت با حفظ ترتیب و توالی، مدیریت شوند.‌ به دلیل متفاوت بودن محدوده ویژگی‌ها باید از روش های نرمال‌سازی[[84]](#footnote-84) استفاده شود تا مدل به یک ویژگی اهمیت بالاتری ندهد. از روش‌های متداول استفاده شده در تحقیقات گذشته برای این امر می‌توان به نرمال‌سازی حداقل-حداکثر[[85]](#footnote-85) طبق فرمول 1 و نرمال‌سازی استاندارد[[86]](#footnote-86) طبق فرمول 2 اشاره کرد. داده‌های‌گم‌شده[[87]](#footnote-87) نیز باید کنترل شوند و بهتر است حذف نشوند تا توالی حفظ شود. برای مدیریت این داده‌ها می‌توان به صورت دستی مقدار اصلی را پیدا کرد و یا میانگین داده قبلی و بعدی را به جای داده گم شده فعلی قرار داد.

فرمول 1:

فرمول 2:

در فرمول 1، منظور از داده فعلی است، و به ترتیب کوچک‌ترین و بزرگ‌ترین داده هستند و در اصل مقدار نرمال‌شده و یا همان خروجی است. در فرمول 2، منظور از داده فعلی است، میانگین، انحراف معیار و z خروجی است.

همانطور که گفته‌شد اخبار و احساسات مردم در مورد ارزها، روی قیمت آن‌ها تاثیر زیادی دارد؛ از این‌رو داده‌های اخبار نیز باید جمع‌آوری شوند که می‌توان از منابعی مانند گوگل‌نیوز[[88]](#footnote-88)، کوین‌دسک[[89]](#footnote-89)، کریپتونیوز[[90]](#footnote-90) و .. استفاده کرد. داده‌های دریافت‌شده باید پیش‌پردازش شوند و سپس وارد الگوریتم‌های مربوطه برای   
تحلیل احساسات شوند. احساسات بدست‌آمده از اخبار باید با داده‌های تاریخی مربوطه اقدام شوند به‌طوری که احساسات هرروز محاسبه شود. از مدل‌ها و متدهای گوناگونی برای این امر استفاده می‌شود که از جمله آن‌ها می‌توان به برت[[91]](#footnote-91)، فاین‌برت[[92]](#footnote-92)، کریپتوبرت[[93]](#footnote-93) و ... اشاره کرد.

از آن‌جایی که داده‌های ارزدیجیتال به صورت سری‌زمانی هستند، باید از الگوریتم‌هایی استفاده شوند که توانایی استخراج ویژگی‌های زمانی[[94]](#footnote-94) به بهترین نحو را داشته باشند. شبکه‌های عصبی بازگشتی[[95]](#footnote-95) از دسته الگوریتم‌هایی هستند که برای داده‌های سری‌زمانی، به دلیل داشتن حافظه، گزینه بسیار مناسبی هستند. مدل های حافظه طولانی کوتاه مدت و واحد بازگشتی دروازه‌ای از جمله این الگوریتم ها هستند که بر اساس تحقیقات گذشته، در حوزه رمزامرزها عملکرد بسیار خوبی ارائه داده‌اند. بنابراین در این تحقیق، تصمیم بر این شده که از الگوریتم های بازگشتی عمیق برای پیش‌بینی ارزهای دیجیتال استفاده شود با این امید که دقت بالایی کسب شود.

طبق تحقیقات گذشته، برای ارزیابی خطای مسائل رگرسیونی مانند ارزهای دیجیتال، از معیارهای میانگین مطلق خطا طبق فرمول 3، ریشه میانگین مربعات خطا طبق فرمول 4، درصد مطلق خطای متوسط طبق فرمول 5 و میانگین مربعات خطا طبق فرمول 6 استفاده خواهد شد.

فرمول 3:

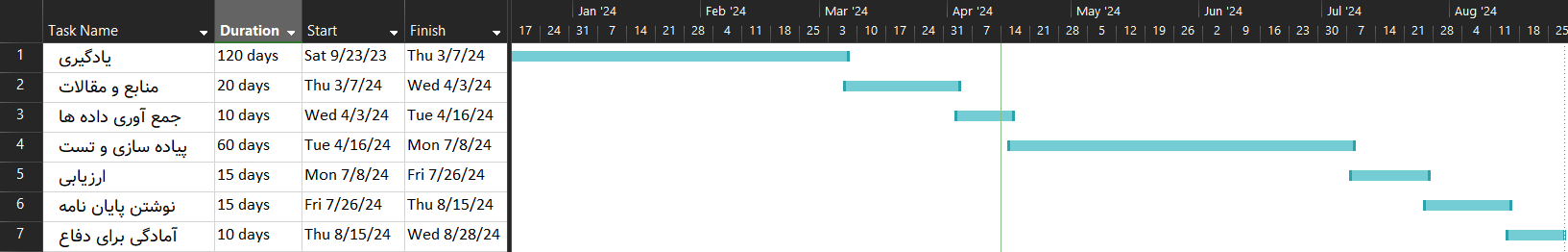
فرمول 4:

فرمول 5:

فرمول 6:

در هر چهار فرمول بالا، مقدار واقعی، مقدار پیش‌بینی شده و تعداد داده‌ها است.

|  |
| --- |
| 1. **زمان بندی:** |



|  |
| --- |
| 1. **مراجع :** |

[1] K. Murray, A. Rossi, D. Carraro, and A. Visentin, "On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles," *Forecasting,* vol. 5, no. 1, pp. 196-209, 2023. [Online]. Available: <https://www.mdpi.com/2571-9394/5/1/10>.

[2] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach," *Fractal and Fractional,* vol. 7, no. 2, p. 203, 2023. [Online]. Available: <https://www.mdpi.com/2504-3110/7/2/203>.

[3] H. Vranken, "Sustainability of bitcoin and blockchains," *Current Opinion in Environmental Sustainability,* vol. 28, pp. 1-9, 2017/10/01/ 2017, doi: <https://doi.org/10.1016/j.cosust.2017.04.011>.

[4] I. E. Livieris, N. Kiriakidou, S. Stavroyiannis, and P. Pintelas, "An Advanced CNN-LSTM Model for Cryptocurrency Forecasting," *Electronics,* vol. 10, no. 3, p. 287, 2021. [Online]. Available: <https://www.mdpi.com/2079-9292/10/3/287>.

[5] W. Kristjanpoller and M. Minutolo, "A hybrid volatility forecasting framework integrating GARCH, Artificial Neural network, Technical Analysis and Principal Components Analysis," *Expert Systems with Applications,* vol. 109, 05/01 2018, doi: 10.1016/j.eswa.2018.05.011.

[6] P. Hájek, L. Hikkerova, and J.-M. Sahut, "How well do investor sentiment and ensemble learning predict Bitcoin prices?," *Research in International Business and Finance,* vol. 64, p. 101836, 12/01 2022, doi: 10.1016/j.ribaf.2022.101836.

[7] W. Chen, H. Xu, L. Jia, and Y. Gao, "Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants," *International Journal of Forecasting,* vol. 37, 04/01 2020, doi: 10.1016/j.ijforecast.2020.02.008.

[8] D. Aggarwal, S. Chandrasekaran, and B. Annamalai, "A complete empirical ensemble mode decomposition and support vector machine-based approach to predict Bitcoin prices," *Journal of Behavioral and Experimental Finance,* vol. 27, p. 100335, 2020/09/01/ 2020, doi: <https://doi.org/10.1016/j.jbef.2020.100335>.

[9] X. Sun, M. Liu, and Z. Sima, "A novel cryptocurrency price trend forecasting model based on LightGBM," *Finance Research Letters,* vol. 32, p. 101084, 2020/01/01/ 2020, doi: <https://doi.org/10.1016/j.frl.2018.12.032>.

[10] N. Aslam, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, "Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model," *IEEE Access,* vol. 10, pp. 39313-39324, 2022, doi: 10.1109/ACCESS.2022.3165621.

[11] J. Chen, "Analysis of Bitcoin Price Prediction Using Machine Learning," *Journal of Risk and Financial Management,* vol. 16, no. 1, p. 51, 2023. [Online]. Available: <https://www.mdpi.com/1911-8074/16/1/51>.

[12] H. Shamshad, F. Ullah, A. Ullah, V. R. Kebande, S. Ullah, and A. Al-Dhaqm, "Forecasting and Trading of the Stable Cryptocurrencies With Machine Learning and Deep Learning Algorithms for Market Conditions," *IEEE Access,* vol. 11, pp. 122205-122220, 2023, doi: 10.1109/ACCESS.2023.3327440.

[13] A. A. Oyedele, A. O. Ajayi, L. O. Oyedele, S. A. Bello, and K. O. Jimoh, "Performance evaluation of deep learning and boosted trees for cryptocurrency closing price prediction," *Expert Systems with Applications,* vol. 213, p. 119233, 2023.

[14] K. R. Rao, M. L. Prasad, G. R. Kumar, R. Natchadalingam, M. M. Hussain, and P. C. S. Reddy, "Time-Series Cryptocurrency Forecasting Using Ensemble Deep Learning," in *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*, 10-11 Aug. 2023 2023, pp. 1446-1451, doi: 10.1109/ICCPCT58313.2023.10245083.

[15] B. Tripathi and R. K. Sharma, "Modeling bitcoin prices using signal processing methods, bayesian optimization, and deep neural networks," *Computational Economics,* vol. 62, no. 4, pp. 1919-1945, 2023.

[16] A. Bouteska, M. Z. Abedin, P. Hajek, and K. Yuan, "Cryptocurrency price forecasting – A comparative analysis of ensemble learning and deep learning methods," *International Review of Financial Analysis,* vol. 92, p. 103055, 2024/03/01/ 2024, doi: <https://doi.org/10.1016/j.irfa.2023.103055>.

به نام خدا

**تعهدنامه**

این­جانب علی منصورآبادی به شماره­ی دانشجویی 400430307 دانشجوی مقطع کارشناسی ارشد تأیید می­کنم که این پایان­نامه حاصل پژوهش خودم است و در مواردی که از منابع دیگران استفاده شده، نشانی دقیق و مشخصات کامل آن را نوشته­ام. همچنین اظهار می­نمایم که تحقیق و موضوع پایان­نامه­ام تکراری نیست و موارد زیر را نیز تعهد می­کنم:

1- بدون کسب مجوز موسسه آموزش عالی زند و اجازه از استاد راهنما، تمام یا قسمتی از دستاوردهای پایان­نامه­ی خود را در مجامع و رسانه­های علمی اعم از همایش­ها و مجلات داخلی و خارجی به صورت مکتوب یا غیرمکتوب منتشر ننمایم.

2- اسامی افراد خارج از کمیته­ی پایان­نامه را بدون اجازه­ی­ استادراهنما به جمع نویسندگان مقاله­های مستخرج از پایان­نامه نامه اضافه نکنم.

3- از درج نشانی یا وابستگی کاری (affiliation) نویسندگان سازمان­های دیگر (غیر از موسسه آموزش عالی زند) در مقاله­های مستخرج از پایان­نامه بدون تأیید استاد راهنما اجتناب نمایم.

همه­ی حقوق مادی و معنوی این اثر مطابق با آیین­نامه­ی مالکیت فکری، متعلق به موسسه آموزش عالی زند است. چنانچه مبادرت به عملی خلاف این تعهدنامه محرز گردد، موسسه زند در هر زمان و به هر نحو مقتضی حق هرگونه اقدام قانونی را در استیفای حقوق خود دارد.

علی منصورآبادی

امضا و تاریخ:

26/02/1403

1. Cryptocurrencies [↑](#footnote-ref-1)
2. Virtual Currencies [↑](#footnote-ref-2)
3. Blockchain [↑](#footnote-ref-3)
4. Bitcoin (BTC) [↑](#footnote-ref-4)
5. Coinmarketcap [↑](#footnote-ref-5)
6. Ethereum (ETH) [↑](#footnote-ref-6)
7. Ripple (XRP) [↑](#footnote-ref-7)
8. Litecoin (LTC) [↑](#footnote-ref-8)
9. pseudo-anonymously [↑](#footnote-ref-9)
10. Proof of Work (PoW) [↑](#footnote-ref-10)
11. Proof of Stake (PoS) [↑](#footnote-ref-11)
12. Application Specific Integrated Circuits (ASICs) [↑](#footnote-ref-12)
13. Dynamic [↑](#footnote-ref-13)
14. Timeseries [↑](#footnote-ref-14)
15. Historical Data [↑](#footnote-ref-15)
16. Machine Learning (ML) [↑](#footnote-ref-16)
17. Traditional Machine Learning [↑](#footnote-ref-17)
18. Multi-layer perceptron (MLP) [↑](#footnote-ref-18)
19. Support Vector Machine (SVM) [↑](#footnote-ref-19)
20. Overfit [↑](#footnote-ref-20)
21. High Level [↑](#footnote-ref-21)
22. Deep Learning (DL) [↑](#footnote-ref-22)
23. Long Short-Term Memory (LSTM) [↑](#footnote-ref-23)
24. Gated Recurrent Unit (GRU) [↑](#footnote-ref-24)
25. Ensemble Learning [↑](#footnote-ref-25)
26. Variance [↑](#footnote-ref-26)
27. Weak Learners [↑](#footnote-ref-27)
28. Gradient Boosting [↑](#footnote-ref-28)
29. Random Forest [↑](#footnote-ref-29)
30. TextBlob [↑](#footnote-ref-30)
31. TextToEmotion [↑](#footnote-ref-31)
32. Technical Features [↑](#footnote-ref-32)
33. Commodities [↑](#footnote-ref-33)
34. Market Index [↑](#footnote-ref-34)
35. Foreign Exchange [↑](#footnote-ref-35)
36. Public Attention [↑](#footnote-ref-36)
37. Random Forest Regressor [↑](#footnote-ref-37)
38. Bidirectional Long Short-Term Memory (Bi-LSTM) [↑](#footnote-ref-38)
39. Root Mean Squared Error (RMSE) [↑](#footnote-ref-39)
40. Mean Absolute Percentage Error (MAPE) [↑](#footnote-ref-40)
41. Tweets [↑](#footnote-ref-41)
42. Sentiment [↑](#footnote-ref-42)
43. Autoregressive Integrated Moving Average (ARIMA) [↑](#footnote-ref-43)
44. FB Prophet [↑](#footnote-ref-44)
45. Stacked Long Short-Term Memory (Stacked LSTM) [↑](#footnote-ref-45)
46. ADA [↑](#footnote-ref-46)
47. BNB [↑](#footnote-ref-47)
48. Train Set [↑](#footnote-ref-48)
49. Test Set [↑](#footnote-ref-49)
50. Mean Absolute Error (MAE) [↑](#footnote-ref-50)
51. R squared - Coefficient of determination [↑](#footnote-ref-51)
52. Mean Squared Error (MSE) [↑](#footnote-ref-52)
53. Boosted Tree-Based Techniques [↑](#footnote-ref-53)
54. Variance [↑](#footnote-ref-54)
55. Dogecoin [↑](#footnote-ref-55)
56. Convolutional Neural Network (CNN) [↑](#footnote-ref-56)
57. Gradient Boosting Machine (GBM) [↑](#footnote-ref-57)
58. Decision Support System [↑](#footnote-ref-58)
59. Anomaly Detection [↑](#footnote-ref-59)
60. Unsupervised Algorithms [↑](#footnote-ref-60)
61. Recurrent Deep Learning Approaches [↑](#footnote-ref-61)
62. K-Nearest Neighbors (KNN) [↑](#footnote-ref-62)
63. Close Price [↑](#footnote-ref-63)
64. Deep Artificial Neural Network [↑](#footnote-ref-64)
65. Window Size [↑](#footnote-ref-65)
66. Bayesian Optimization [↑](#footnote-ref-66)
67. Hyperparameters [↑](#footnote-ref-67)
68. Autoencoder [↑](#footnote-ref-68)
69. Light Gradient Boosting Machine (LightGBM) [↑](#footnote-ref-69)
70. Simple Recurrent Neural Network (Simple RNN) [↑](#footnote-ref-70)
71. Real Time [↑](#footnote-ref-71)
72. Coinmarketcap [↑](#footnote-ref-72)
73. Yahoo Finance [↑](#footnote-ref-73)
74. Kaggle [↑](#footnote-ref-74)
75. Open Price [↑](#footnote-ref-75)
76. High [↑](#footnote-ref-76)
77. Low [↑](#footnote-ref-77)
78. Volume [↑](#footnote-ref-78)
79. Marketcap [↑](#footnote-ref-79)
80. Timestamp [↑](#footnote-ref-80)
81. Daily [↑](#footnote-ref-81)
82. Outliers [↑](#footnote-ref-82)
83. Interquartile range (IQR) [↑](#footnote-ref-83)
84. Normalization Methods [↑](#footnote-ref-84)
85. Min-Max Normalization [↑](#footnote-ref-85)
86. Standard Scaler [↑](#footnote-ref-86)
87. Missing Data [↑](#footnote-ref-87)
88. Googlenews [↑](#footnote-ref-88)
89. Coindesk [↑](#footnote-ref-89)
90. Cryptonews [↑](#footnote-ref-90)
91. Bidirectional Encoder Representations from Transformers (BERT) [↑](#footnote-ref-91)
92. Financial Bidirectional Encoder Representations from Transformers (FinBERT) [↑](#footnote-ref-92)
93. Cryptocurrency Bidirectional Encoder Representations from Transformers (CryptoBERT) [↑](#footnote-ref-93)
94. Temporal Features [↑](#footnote-ref-94)
95. Recurrent Neural Networks (RNN) [↑](#footnote-ref-95)