



فصلنامه‌ی اقتصاد مقداری

صفحه ی اصلی وب سایت مجله:

www.jqe.scu.ac.ir

شاپا الکترونیکی: 2717-4271

شاپا چاپی: 2008-5850



انتخاب مدل بهینه برای پیش بینی بازدهی دارایی: مقایسه مدل‌های ARIMA- MIDAS و GARCH

کامبیز هژبر کیانی*، سعید صدرزاده مقدم**

* استاد گروه اقتصاد، گروه اقتصاد، دانشکده مدیریت و اقتصاد، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران
** دانشجوی دکتری اقتصاد، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران (نویسنده مسئول)

اطلاعات مقاله	طبقه‌بندی JEL: C89, C22, G17
تاریخ دریافت:	واژگان کلیدی:
تاریخ بازنگری:	مدل میداس، خودرگرسیون میانگین متحرک، واریانس ناهمسان
تاریخ پذیرش:	شرطی، بیتکوین، داده‌های پرتواتر
انتشار آنلاین از تاریخ:	
ارتباط با نویسنده (گان) مسئول:	ادرس پستی:
ایمیل: saeed.sadrzadeh@srbiau.ac.ir	تهران- انتهای بزرگراه شهید ستاری- میدان دانشگاه- بلوار شهدای
0000-0001-9930-0385	حصارک- دانشگاه آزاد اسلامی واحد علوم و تحقیقات . صندوق پستی
14515/775	

اطلاعات تکمیلی:

این مقاله برگرفته از پایان نامه‌ی دکتری سعید صدرزاده مقدم در رشته اقتصاد به راهنمایی آقای دکتر کامبیز هژبر کیانی در دانشگاه آزاد واحد علوم تحقیقات تهران می‌باشد.

فردانی: از تمامی افراد و موسساتی که در انجام این تحقیق مولف را مساعدت نمودند، قدر دانی می‌شود.
تضاد منافع: نویسنده مقاله اعلام می‌کند که در انتشار مقاله ارائه شده تضاد منفعلی وجود ندارد.
منابع مالی: نویسنده (ها) هیچگونه حمایت مالی برای تحقیق، تألیف و انتشار این مقاله دریافت نکرده اند.

چکیده

در این مطالعه به منظور بررسی و مقایسه دو رویکرد مدل‌سازی در اقتصادسنجی از نظر دقت پیش‌بینی، سری زمانی قیمت بیتکوین برای بازه زمانی 2015 تا 2023 میلادی انتخاب شده است. بیتکوین بعنوان محبوبترین و شناخته شدهترین رمز ارز، در سالیان اخیر تبدیل به یک دارایی مهم برای بسیاری از سرمایه‌گذاران بازارهای مالی و حتی مردم عادی گردیده. لذا مطالعات بسیاری برای کشف ساختار قیمتی و تغییرات این دارایی در طول زمان در حال انجام می‌باشد. نوسانات بسیار بالای این رمز ارز پیچیدگی‌های زیادی در این مطالعات ایجاد کرده است. رویکرد مدل‌سازی برای داده‌های با تواترهای زمانی متفاوت (MIDAS) بعنوان یک رویکرد جدید در مدل‌سازی در سالیان گذشته مطرح و با توجه به حذف محدودیت‌های مهم در مدل‌سازی اقتصادسنجی به صورت روزافزون در حال بررسی و مطالعه است. در این مطالعه میزان دقت پیش‌بینی مدل خودرگرسیون جمعی میانگین متحرک همراه با واریانس ناهمسان شرطی و مدل‌سازی با رویکرد داده‌های ترکیبی با تواتر زمانی متفاوت مقایسه شده است. مدل‌سازی انجام شده در این تحقیق براساس داده‌های درون روزانه‌ای قیمت بیتکوین (با تواتر بالا - 5 دقیقه‌ای) بر روی متغیرهای بازدهی روزانه و هفتگی بیتکوین صورت گرفته است. ابتدا یک مدل خودرگرسیون جمعی میانگین متحرک با وجود اثرات ناهمسانی در واریانس شرطی و با در نظر گرفتن انواع مدل‌های گارچ و توزیع‌های متفاوت در هر یک و سپس مدل داده‌های با تواتر زمانی متفاوت با توابع وزنی مختلف برآورد گردیده است. در نهایت با تقسیم بازه نمونه‌گیری به دو بخش، پیش‌بینی بازدهی سری زمانی بیتکوین با استفاده از

هر دو رویکرد انجام و با اطلاعات واقعی مقایسه شده است. نتایج نشان می‌دهد مدل سازی با رویکرد داده‌های با تواتر زمانی متفاوت نتایج بهتری در خصوص پیش‌بینی بازدهی هر دو سری زمانی بازدهی روزانه و هفتگی بیتکوین ارائه می‌نماید. محاسبات و برآوردها توسط زبان برنامه نویسی R انجام شده است.

ارجاع به مقاله:

هژبر کیانی، کامبیز، صدرزاده مقدم، سعید. (سال انتشار). انتخاب مدل بهینه برای پیش بینی بازدهی دارایی : مقایسه مدل های ARIMA-GARCH و MIDAS. اقتصاد مقداری، دوره (شماره)، ص-ص.



© 2021 Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0 license) (<http://creativecommons.org/licenses/by-nc/4.0/>)

گیر فایل
انتشار

1- مقدمه

اهمیت پیش‌بینی دقیق بازدهی دارایی‌ها، به‌ویژه در مورد دارایی‌های نوظهور مانند رمزارزها، یکی از موضوعات اصلی و با اهمیت در این حوزه است. بیت‌کوین، که به سرعت به یکی از مهم‌ترین دارایی‌های مالی در سطح جهانی تبدیل شده است، نقش محوری در تصمیم‌گیری‌های سرمایه‌گذاری و استراتژی‌های مدیریت ریسک دارد. پیش‌بینی‌های دقیق بازدهی بیت‌کوین می‌تواند به سرمایه‌گذاران کمک کند تا فرصت‌های سرمایه‌گذاری را شناسایی کرده و از نوسانات قیمتی به نفع خود استفاده کنند. علاوه بر این، درک بهتری از روند بازار می‌تواند به کاهش ریسک‌های مالی و افزایش ثبات در بازارهای مالی کمک کند. بنابراین، توسعه مدل‌های پیش‌بینی که قادر به مقابله با پیچیدگی‌های بازارهای رمزارز هستند، از اهمیت بالایی برخوردار است و می‌تواند تأثیرات مثبتی بر اقتصاد جهانی داشته باشد. پیش‌بینی مقادیر آینده یک سری زمانی اقتصادی یا مالی همواره مورد توجه محققین بوده و روش‌های متنوعی نیز در این حوزه بکار گرفته شده و با در طول سال‌های گذشته ابداع گردیده است. مفهوم پیش‌بینی متناظر بر آینده و رویدادهای دوره زمانی آتی است، که غیرقطعی بودن و نقص در اطلاعات از ویژگی‌های اصلی آن است. رویکردهای پیش‌بینی به دو گروه رویکردهای کمی و کیفی تقسیم می‌شوند. این تقسیم‌بندی بر اساس میزان وابستگی رویکرد پیش‌بینی، به الگوها و روش‌های ریاضیاتی و آماری است. روش‌های کمی پیش‌بینی که بر اساس محاسبات صرفاً ریاضی صورت می‌گیرند به دو گروه روش‌های کمی رگرسیونی و روش‌های کمی غیررگرسیونی تقسیم می‌شوند. مواردی مانند میانگین ساده، تعدیل نمایی و ... در گروه روش‌های ریاضیاتی غیررگرسیونی قرار می‌گیرند. در گروه مقابل و در حوزه روش اقتصادسنجی، الگوهای رگرسیونی علی و غیرعلی وجود دارند که مدل‌های خودرگرسیون با واریانس ناهمسان شرطی، خودرگرسیون جمعی میانگین‌متحرک از روش‌های علی در حوزه پیش‌بینی متغیرهای اقتصادی هستند (Brockwell et al, 2002). یک مدل خودرگرسیون میانگین‌متحرک متشکل از دو فرآیند خودرگرسیون و میانگین‌متحرک است که با تجلیل روند تغییرات گذشته و حال یک سری زمانی، ساختار درونی آن را کشف و امکان پیش‌بینی مقادیر آینده را برای محقق فراهم می‌نماید. لذا یکی از الگوهای مهم و کاربردی در اقتصادسنجی، مدل‌سازی سری‌های زمانی با استفاده از مقادیر گذشته و با وقفه سری‌زمانی می‌باشد که در این مطالعه، به‌عنوان یک روش سنتی مدل‌سازی و پیش‌بینی مقادیر آتی سری‌زمانی معرفی می‌شود. در فرآیند توسعه مدل‌های پیش‌بینی بازارهای مالی، از زمان‌های اولیه، که مدل‌های ساده رگرسیونی برای تحلیل داده‌های مالی استفاده می‌شدند، تا دوران کنونی که مدل‌های پیچیده‌تر و دقیق‌تری بکار گرفته می‌شود، طراحی مدل مناسب با چالش‌های خاص خود همراه بوده است. مدل‌های خودرگرسیون با واریانس ناهمسان شرطی و خودرگرسیون جمعی میانگین‌متحرک، که در دهه‌های اخیر مورد استفاده قرار گرفته‌اند، نمونه‌هایی از تلاش‌های علمی برای فهم بهتر و پیش‌بینی رفتار بازارهای مالی هستند. با این حال، این مدل‌ها نیز با محدودیت‌هایی، به‌ویژه در برخورد با داده‌های با تواترهای زمانی متفاوت مواجه بودند. بعبارت دیگر یکی از محدودیت‌های مهم در مدل‌سازی که تا سالین اخیر، محققین این حوزه ملزم به رعایت و پیروی از آن بودند، لزوم تواتر یکسان در متغیرهای مدل بود. محققین به‌منظور مدل‌سازی متغیرهای سری‌زمانی با تواترهای زمانی متفاوت، بایستی ابتدا تمام متغیرها را هم‌تواتر کرده و سپس وارد مدل نمایند. برای اینکار دو راهکار وجود دارد. یکسان‌سازی تواتر متغیرها به سمت متغیر با تواتر زمانی بالاتر و یا یکسان‌سازی به سمت متغیر با تواتر زمانی پایین‌تر. که روش اول هم‌فزونی زمانی¹ و روش دوم جداسازی زمانی² نام‌گذاری شده است (Bayat & Nofaresti, 2015). هم‌فزونی زمانی عبارت است از کاهش تواتر یک متغیر از طریق میانگین‌گیری یا مجموع‌یابی آن در طول زمان که موجب برآورد مدل در پایین‌ترین تواتر زمانی خواهد شد. لذا با هم‌فزونی زمانی مجموعه اطلاعاتی در دسترس کوچک شده و سری زمانی هم‌فزون شده نسبت به سری زمانی ابتدایی هموارتر می‌شود. این اثر در تبدیل به تواترهای زمانی بزرگتر (مانند ماه به سال نسبت به فصل به سال) شدیدتر

¹ Temporal Aggregation

² Temporal Disaggregation

رخ داده و به همین دلیل در زمان مواجهه با حجم نمونه کم استفاده از هم‌فزونی زمانی توصیه نمی‌شود. در روش جداسازی مدل در بالاترین تواتر زمانی تخمین زده می‌شود لذا فرض می‌شود که سری زمانی در بستر با تواتر بالا تولید می‌شود در حالیکه بصورت سیستماتیک متغیرها تنها در برخی از مقاطع زمانی مشاهده شده‌اند. عبارات دیگر مشاهدات مفقود شده در این روش بازیابی می‌شوند. مزیت این روش نسبت به روش هم‌فزونی عدم از دست‌دادن اطلاعات مربوط به متغیر با تواتر زمانی بالا است (Silvestrini & erdas, 2008). در اکثر موارد به‌کارگیری این روش‌ها موجب کاهش کارایی مدل برآورد شده و از دست رفتن اطلاعات مفید می‌گردد. تا سال‌ها یافتن الگویی که بتواند از متغیرهای با تواتر بالا، به‌منظور پیش‌بینی متغیرهایی با تواتر پایین استفاده نموده و پیش‌بینی مناسبی از متغیر کم تواتر را ممکن نماید، چالشی اساسی بود. این محدودیت‌ها منجر به توسعه مدل‌های جدیدی گردید که قادر به ادغام داده‌های با تواتر بالا و پایین برای ارائه پیش‌بینی‌های دقیق‌تر باشند. روش داده‌های ترکیبی با تواتر زمانی متفاوت به‌عنوان یک رویکرد جدید و نوآورانه در مدل‌سازی متغیرهای اقتصادی توسط کلاین و سوچو³ (1989) ارائه و بعدها توسط گیسلز (2004) توسعه یافت (Ghysels, 2004). مدل‌های داده‌های با تواتر زمانی متفاوت به مدل‌هایی گفته می‌شود که در آنها متغیرها با تواترهای زمانی مختلف وارد مدل می‌شوند. این مدل‌ها به دو دسته الهام گرفته شده از هم‌فزونی زمانی و جداسازی زمانی تقسیم می‌شوند. مدل‌های معرفی شده توسط گیسلز نمونه تکامل یافته روش هم‌فزونی زمانی و مدل‌های حالت-فضا⁴ نیز حالت توسعه یافته روش جداسازی زمانی است. در یک مدل حالت-فضا تمام متغیرهای با تواتر زمانی بالا مدل‌سازی شده و از طریق یک فیلتر کالمن⁵ پارامترهای مدل برآورد می‌شود. لذا در این مدل‌ها هرچه تعداد متغیرهای مدا بیشتر باشد و یا اختلاف میان تواتر زمانی متغیرها بالاتر باشد پیچیدگی‌های مدل افزایش می‌یابد. مزیت اصلی مدل‌های رگرسیون با داده‌های تواتر زمانی متفاوت نسبت به مدل‌های حالت-فضا استفاده از پارامترهای کمتر و در نتیجه فرم خلاصه‌تر و پیچیدگی‌های محاسباتی کمتر است. مطالعات اندرو و همکاران (2012) نشان داد که دقت پیش‌بینی به استفاده از مدل‌های رگرسیون با داده‌های زمانی متفاوت نسبت به مدل‌های حالت-فضا افزایش می‌یابد (Andreou et al, 2012). در این مطالعه با استفاده از اطلاعات یک سری زمانی، تلاش شده است رویکردهای سنتی در پیش‌بینی در مقابل رویکرد نوین معرفی‌شده مورد آزمون قرار گیرد. این سؤال اصلی مطرح می‌شود که آیا روش‌های سنتی اقتصادسنجی نسبت به رویکردهای نوین، از نظر خطای کمتر و دقت بالاتر در پیش‌بینی، مدل‌های بهتری را ارائه می‌دهد یا خیر؟ به شکل خاص در این مطالعه مدل سنتی خودرگرسیون جمعی میانگین‌متحرک در مقابل رویکرد مدل‌سازی داده‌های ترکیبی با تواتر زمانی متفاوت مقایسه شده است. در این مقایسه از داده‌های تواتر زمانی بالا سری زمانی بیت کوین (داده‌های میان روزانه 5 دقیقه‌ای) استفاده شده و بر اساس معیارهای نمایش‌دهنده میزان خطای پیش‌بینی و بررسی برتری مدل‌ها نسبت به یکدیگر، مقایسه میان دو رویکرد انجام شده است.

2- پیشینه تحقیق

2-1- مطالعات انجام شده با مدل‌های ARIMA و GARCH

مطالعات مربوط به پیش‌بینی در علم اقتصاد در طول دهه‌های آغازین قرن بیستم بسیار مورد توجه محققین قرار گرفت. در سال‌های اخیر الگوهای سری‌زمانی به منظور پیش‌بینی تناسب بیشتری نشان داده‌اند، چرا که متغیرهای اقتصادی سری‌زمانی، تمام اطلاعات مربوط به خود را در بردارند و می‌توان قوی‌ترین منبع برای توضیح هر متغیر را خود متغیر دانست (Moshiri, 2001). ارائه الگوهای متفاوت در این حوزه نیز دال بر اهمیت پیش‌بینی‌های سری‌زمانی است. البته در مورد برخی از سری‌ها مانند قیمت سهام، این توجه بیشتر مشهود است. به عنوان مثال وو و لو (1993)، قیمت سهام در آمریکا را برای بازه‌های زمانی متفاوت با استفاده از مدل‌های اقتصادسنجی پیش‌بینی کردند. نتایج مطالعه‌ی آنها نشان

³ Klien & Sojo (1989).

⁴ Stat-Space

⁵ Kalman Filter

داد که در پیش‌بینی کوتاه‌مدت قیمت سهام، شبکه‌ی عصبی مصنوعی⁶ در مقایسه با روش ARIMA پیش‌بینی‌های دقیق‌تری را ارائه می‌کند، اما در بلندمدت فرآیند ARIMA توانایی بالاتری برای پیش‌بینی دارد (Wu and Lu, 1993). رینگ وود و همکاران (1993)، مصرف هفتگی برق در کشور انگلستان را با استفاده از مدل‌هایی مختلف پیش‌بینی کردند. نتایج این مطالعه نشان داد که پیش‌بینی سری زمانی مصرف انرژی براساس یک فرآیند AR دقیق‌تر از مدل‌سازی این سری زمانی بر اساس یک فرآیند ARIMA است، اما در مقابل فرآیند ARIMA خطای کمتری در مقایسه با مدل AR دارد (Ringwood et al, 1993). غفاری و فرهادی (1394) در مقاله‌ای به مقایسه دقت پیش‌بینی چهار مدل ARIMA، GARCH، ARIMA-GARCH و مدل حالت-فضا در تخمین و پیش‌بینی شاخص قیمت بورس اوراق بهادار تهران (تیبکس) پرداختند. آنها در این مطالعه از داده‌های روزانه برای پیش‌بینی درون و برون‌نمونه‌ای استفاده کردند. همچنین برای بررسی بیشتر و افزایش دقت پیش‌بینی مدل‌های بکار گرفته شده در تحقیق، شبیه‌سازی با روش مونت کارلو⁸ برای دو دوره زمانی میان‌مدت و کوتاه‌مدت با استفاده از برون‌داده و نیز برای یک مقایسه کلی با درون‌داده انجام دادند. دقت پیش‌بینی نتایج مدل‌ها با استفاده از معیار RMSE ارزیابی شد. نتایج این مطالعه نشان داد که مدل GARCH برای پیش‌بینی‌های برون‌نمونه‌ای و در سه دوره زمانی (بلندمدت، میان‌مدت و کوتاه‌مدت) از دقت پیش‌بینی بیشتری نسبت به سایر مدل‌ها برخوردار است، اما در هنگام مقایسه با پیش‌بینی‌های درون نمونه‌ای، مدل ARIMA مدل مناسب‌تری است (Ghafari & Farhadi, 2015). پهلوانی و روشن (1394) در مقاله‌ای به مقایسه کارایی پیش‌بینی مدل ARIMA و مدل ترکیبی ARIMA-GARCH با استفاده از سری زمانی نرخ ارز در ایران پرداختند. در این مطالعه از داده‌های نرخ دلار منتشر شده توسط بانک مرکزی ج.ا.ا برای بازه زمانی یکسال استفاده شد. آنها با محاسبه بازدهی دلار و تقسیم بازه زمانی داده‌های جمع‌آوری شده به دو بخش، پیش‌بینی‌های درون و برون‌نمونه‌ای را انجام و قدرت پیش‌بینی مدل‌های مذکور را بر اساس معیارهای MAE، RMSE، مقایسه قرار گرفت. نتایج نشان داد که مدل ترکیبی EGARCH(2,1) - ARIMA((7,2),(12)) دارای کمترین میزان خطا در پیش‌بینی مقادیر درون و برون نمونه‌ای سری زمانی بازدهی دلار است (Pahlevani & Roshan, 2015). صالحی فر (1398) در مطالعه‌ای به منظور بررسی بازدهی و ریسک رمزارز بیت‌کوین در مقایسه با بازارهای رقیب (طلا، ارز و بورس) از مدل‌های GJR-GARCH و GARCH آستانه استفاده کرد. وی با استفاده از داده‌های روزانه بیت‌کوین و شاخص‌های مرتبط با بازارهای رقیب در یک بازه زمانی 5 ساله، بازدهی لگاریتمی را برای هر یک از شاخص‌ها محاسبه و نوسانات بازدهی را مدل‌سازی کرد. یافته‌های مطالعه نشان داد اگرچه بازده و ریسک بیت‌کوین نسبت به سایر فرصت‌های سرمایه‌گذاری مانند ارز، طلا، سکه و بورس در داخل کشور به طور قابل ملاحظه‌ای بیشتر است، اما نمی‌توان رفتار آن را از نظر ریسک و بازدهی با بازارهای رقیب مرتبط دانست. همچنین برخلاف سایر دارایی‌ها، در معاملات بیت‌کوین اثر اخبار مثبت بیشتر از اخبار منفی است. در نهایت فرضیه دایربرگ⁹ مبنی بر اینکه بیت‌کوین چیزی بین طلا و ارز است، مورد تایید قرار نگرفت (Salehifar, 2019). باریوربا (2017) در مطالعه‌ای نوسانات قیمت بیت‌کوین را با استفاده از مدل‌های خانواده GARCH و در نظر گرفتن توزیع خطای نرمال مدل‌سازی کرد و نتیجه گرفت مدل AR(1) - CGARCH(1, 1) بهترین عملکرد برای پیش‌بینی نوسان قیمتی بیت‌کوین دارد (Bariviera, 2017). کاتانیا، گراسی و راوازلو (2019) توانایی چندین مدل تک‌متغیره و دو متغیره را برای پیش‌بینی برخی رمزارزها مانند بیت‌کوین، لایت‌کوین، ریپل و اتریوم مقایسه کردند. نتایج نشان داد که ترکیب مدل‌های تک‌متغیره، بهبود قابل توجهی در پیش‌بینی‌های نقطه‌ای، برای دو رمزارز بیت‌کوین و اتریوم بوجود می‌آورد (Catania, Grassi & Ravazzolo, 2019). نیجینگ هووا (2020) در مطالعه‌ای

⁶ Artificial Neural Networks (ANN).

⁷ Generalized Autoregressive Conditional Heteroscedasticity (GARCH).

⁸ Monte Carlo Simulation

⁹ Dyrhberg, 2016

دقت مدل‌های ARIMA و مدل ترکیبی شبکه عصبی مصنوعی با حافظه کوتاهمدت طولانی¹⁰ را در پیش‌بینی قیمت رمزارزهای بیت‌کوین و اتریوم مقایسه نمود. در این مطالعه با جمع‌آوری 10 هزار داده مربوط به دو رمز ارز بیت‌کوین و اتریوم، مدل $ARIMA(1,1,0)$ برای هر دو سری زمانی انتخاب و پیش‌بینی‌های به‌دست‌آمده از این مدل را با نتایج شبیه‌سازی در شبکه عصبی مقایسه شد. نتایج نشان داد که هر دو مدل دقت کافی در پیش‌بینی قیمت بیت‌کوین و اتریوم را دارند. با اینکه مدل شبکه عصبی زمان طولانی‌تر و منابع بیشتر (از نظر قدرت پردازش سیستم) برای آموزش و شبیه‌سازی صرف می‌کند، اما کارایی بالاتری در پیش‌بینی نسبت به مدل ARIMA نشان می‌دهد. بر اساس نتایج مدل شبکه عصبی با تعداد داده‌های کمتر، پیش‌بینی‌های بهتری ارائه می‌کند و مدل ARIMA برای بازه‌های زمانی طولانی‌تر نسبت به بازه‌های نزدیک‌تر پیش‌بینی‌های بهتری ارائه می‌نماید (Yiqing Hua, 2020). چگنی و گردعزیز (1399) به مقایسه کارایی دو روش پیش‌بینی ANN و روش ARIMA، در پیش‌بینی قیمت سهام در بازار سهام ایران پرداختند. بدین منظور از داده‌های دو شرکت دارویی البرز دارو و جام دارو استفاده و مدل‌های ARIMA و ANN برای هر دو شرکت برآورد شد. نتایج نشان داد که مدل ANN پیش‌بینی بهتری از قیمت سهام در مقایسه با روش ARIMA دارد (Chegeni & Gord, 2020). عباسی (1400) به منظور معرفی بهترین مدل GARCH با عملکرد بهینه در افق‌های زمانی، از سری زمانی قیمت نفت خام وست تگزاس اینترمدیت¹¹ استفاده نمود. در این مطالعه نوسانات قیمت روزانه و هفتگی نفت خام بر اساس الگوهای GARCH تک‌رژیمی $GJR-GARCH(1,1)$ و مدل‌های تغییر رژیم (HYGARCH, FIGARCH, EGARCH, GARCH و MRS-GARCH) و پیش‌بینی و مدل‌سازی شدند. نتایج حاکی از دقت بالای مدل MRS-GARCH بر حسب داده‌های هفتگی برای پیش‌بینی‌های درون‌نمونه‌ای و برتری مدل‌های GARCH تک‌رژیمی برای پیش‌بینی‌های برون‌نمونه‌ای بود (Abasi, 2021).

2-2- مطالعات انجام شده با مدل‌های MIDAS

کلمنتس و گالوانو (2008) برای پیش‌بینی رشد تولید کل در ایالت متحده، از رویکرد MIDAS استفاده کردند. آن‌ها نتیجه گرفتند که استفاده از رویکرد مدل‌سازی MIDAS نسبت به مدل AR و خودرگرسیون با وقفه‌های توزیعی¹² منجر به کاهش خطای مجذور میانگین مربعات می‌شود و با در نظر گرفتن شاخص‌های ماهیانه، این رویکرد برای پیش‌بینی‌های کوتاهمدت کاربرد دارد (Clements & Galvao, 2008). برکچیان و رضایی (1395) در مطالعه‌ای دقت پیش‌بینی مدل ARDL و مدل MIDAS را برای پیش‌بینی نرخ تورم فصلی ایران مورد مقایسه قرار دادند. نتایج مطالعه نشان داد که به‌کارگیری مدل MIDAS با استفاده از مشاهدات ماهیانه نرخ تورم برای پیش‌بینی تورم فصلی، اغلب منجر به نتایج دقیق‌تری نسبت به مدل‌های ARDL می‌شوند. این بهبود به‌ویژه در پیش‌بینی‌های یک گام به جلو خود را بهتر نشان می‌دهد (Barakchian & Rezaei, 2016). پیش بهار و همکاران (1398) ضمن بررسی عوامل موثر بر رشد ارزش افزوده بخش کشاورزی، به پیش‌بینی مقادیر آبی با استفاده از مدل MIDAS پرداخته و برآوردها را با نتایج به‌دست‌آمده از پیش‌بینی الگوی ARIMA مقایسه نمودند. مقایسه نتایج به‌دست‌آمده برای رشد ارزش افزوده بخش کشاورزی در مقایسه با داده‌های واقعی، حاکی از قدرت پیش‌بینی الگوی MIDAS نسبت به مدل‌سازی با الگوی ARIMA بود (Pishbahar et al, 2019). دشتیان و همکاران (1398) با به‌کارگیری یک الگوی تلاطم تصادفی با اثرات اهرمی و داده‌های ترکیبی با تواتر متفاوت، به بررسی اثر ناطمینانی نرخ ارز بر جریان گردشگری در ایران پرداختند. آن‌ها با استفاده از داده‌های فصلی و سالیانه برای بازه زمانی 1368 الی 1394، ابتدا ناطمینانی نرخ ارز را با یک الگوی تلاطم تصادفی با اثر اهرمی استخراج و سپس با استفاده از روش داده‌های ترکیبی با تواتر زمانی متفاوت اثر تلاطم نرخ ارز

¹⁰ Long-Short Term Memory (LSTM)

¹¹ West Texas Intermediate

¹² Autoregressive Distributed Lag (ARDL).

واقعی را بر جریان گردشگری در ایران برآورد نمودند. نتایج نشان داد تلاطم نرخ ارز واقعی و تولید ناخالص داخلی بر ورود گردشگران خارجی به ایران، به‌ترتیب اثرات منفی و مثبت داشته است (Dashtban et al, 2019). فنگ‌ما و همکاران (2020) با رویکرد MIDAS و استفاده از الگوهای تغییر جهت مارکوف در برآورد ضرایب، به مدل‌سازی نوسانات رمازرها پرداختند. آن‌ها با استفاده از داده‌های میان‌روزانه و تجزیه واریانس تحقق‌یافته¹³ رمازرها، چهار مدل ترتیبی (با اضافه کردن مرحله‌ای متغیر و شرایط به مدل) معرفی و بر اساس میزان خطا در پیش‌بینی‌های برون‌نمونه‌ای، مدل بهینه برای پیش‌بینی واریانس تحقق‌یافته را انتخاب نمودند. نتایج این مطالعه نشان‌دهنده بهبود قابل‌توجه قدرت پیش‌بینی برای بازه زمانی دو هفته و یک‌ماهه، در صورت استفاده از مدل تکمیلی بود (Feng Ma et al, 2020). نوفرستی و سزاوار (2021) به منظور پیش‌بینی متغیرهای کلان اقتصادی جهت ارزیابی دقیق سیاست‌های اقتصادی و لزوم واکنش سریع به شوک‌های وارده و برنامه‌ریزی متناسب، یک الگوی کلان‌سنجی داده‌های ترکیبی با تواتر زمانی متفاوت را براساس چارچوب الگوی عرضه کل- تقاضای کل و متشکل از چهار بازار عمده بازار کالاها و خدمات، پول، ارز و کار طراحی نمودند. نتایج مطالعه نشان داد که استفاده از مشاهدات جدید در متغیرهای با تواتر بالا در الگو، منجر به بهبود دقت نتایج در پیش‌بینی متغیرهای درون‌زای الگو خواهد شد. همچنین نتایج مطالعه میزان تاثیر سیاست‌های مالی اجرا شده توسط دولت را بر بخش‌های مختلف اقتصاد را نشان می‌دهد (Noferești & Sezavar, 2021).

همانگونه که مشاهده می‌شود اغلب مطالعات انجام شده در خصوص پیش‌بینی سری‌های زمانی مبتنی بر مدل‌سازی‌های اقتصادسنجی براساس الگوهای ARIMA و مدل‌های مبتنی بر ناهمسانی واریانس شرطی (GARCH، ARCH) انجام شده و در برخی از موارد مقایسه مدل‌های مذکور با الگوریتم‌های شبیه‌سازی شبکه عصبی صورت گرفته است. در اغلب این مطالعات یک مدل پایه با چند توزیع خطا و با چند مدل از خانواده مدل‌های ARIMA-GARCH با یک توزیع خطا مورد بررسی قرار گرفته است. همچنین مطالعات انجام شده در حوزه مدل‌سازی MIDAS مخصوصاً در ادبیات داخلی، براساس داده‌های با تواترهای زمانی متفاوت سالیانه، فصلی و نهایتاً ماهیانه است. حال آنکه داده‌های موجود در بازارهای مالی مانند بازار سهام و رمازرها بصورت میان‌روزانه نیز در دسترس است. مدل‌های MIDAS امکان بررسی میزان اهمیت و همچنین تاثیر داده‌های میان‌روزانه در دقت پیش‌بینی‌های روزانه و هفتگی را فراهم می‌نماید. بنابراین بررسی این موضوع که آیا اطلاعات اضافی بدست آمده بعلت تواتر زمانی بالای داده‌های میان‌روزانه با مدل‌سازی MIDAS می‌تواند موجب افزایش محسوس میزان دقت پیش‌بینی‌ها نسبت به رویکردهای سنتی مدل‌سازی اقتصادسنجی شود چالش مهمی است.

لذا در این مطالعه ضمن ترکیب مدل‌سازی خودرگرسیون جمعی میانگین متحرک با فرض ناهمسانی واریانس شرطی که حالت‌های مختلف مدل‌های GARCH شامل (sGARCH، eGARCH، gjrGARCH) و توزیع‌های مختلف (توزیع نرمال¹⁴، توزیع چوله نرمال¹⁵، توزیع تی‌استیودنت¹⁶، توزیع چوله تی‌استیودنت¹⁷ و توزیع خطای تعمیم یافته¹⁸) در نظر گرفته شده، از رویکرد جدید مدل‌سازی MIDAS نیز استفاده شده است. همچنین به منظور بررسی جامع در بخش مدل‌سازی MIDAS، سه مدل با توابع وزن‌دهی آلمون¹⁹، آلمون نمایی²⁰ و بتا²¹ مورد برآورد قرار گرفته است.

¹³ Realized Variance (RV)

¹⁴ Normal Distribution

¹⁵ Skew Normal Distribution

¹⁶ t-Student distribution

¹⁷ Skew t-Student distribution

¹⁸ Generalized error distribution (GED)

¹⁹ Almon weight function

²⁰ Exponential Almon weight function

²¹ Beta weight function

داده‌های مورد استفاده در این مطالعه، داده‌های پرتواتر میان‌روزانه (سری زمانی قیمت بیت کوین بصورت 5 دقیقه‌ای) است که بازدهی روزانه و هفتگی، از طریق این داده‌ها محاسبه و مدل‌سازی داده‌های با تواتر متفاوت بر این اساس انجام شده است. به عبارت دیگر مقایسه انجام‌شده در این مطالعه مقایسه‌ای میان روش‌های سنتی پیش‌بینی در اقتصادسنجی و رویکردهای نوین در این حوزه است با این تفاوت که در استفاده از روش‌های سنتی، جمعیتی از الگوهای استفاده‌شده تاکنون بکار گرفته‌شده است.

3- روش شناسی تحقیق

در دهه‌های اخیر، علاقه محققین برای استفاده از مدل‌های اقتصادسنجی به منظور پیش‌بینی متغیرها در انواع بازارهای مالی و غیر مالی افزایش یافته است. برای هر بازار بخصوص، باید مدل مناسب با وضعیت آن بازار انتخاب شود، در این راستا اقتصاددانان و محققان زیادی دست به مقایسه مدل‌های مناسب‌تر برای بازارهای مختلف زدند. اغلب این مدل‌ها با تحلیل الگوها و روند داده‌های گذشته تلاش می‌کنند رفتارهای آینده را پیش‌بینی و اطلاعات مورد نیاز برای تصمیم‌گیری‌های موثر را فراهم نمایند. در این راستا یکی از عوامل موثر در توسعه مدل‌های پیش‌بینی مواجهه با خصوصیات منحصر به فرد داده‌ها مانند بازه زمانی جمع‌آوری یا ویژگی‌های ذاتی مانند نوسانات بالا است که موجب تلاش برای طراحی مدل‌های جدید شده است. در این بخش ابتدا رویکردهای پایه‌ای تجزیه تحلیل سری‌های زمانی ارائه و سپس یک رویکرد جدید در مدل‌سازی سری‌های زمانی برای داده‌های با تواتر زمانی متفاوت مورد بررسی قرار گرفته است.

3-1- مدل‌های خودرگرسیون میانگین متحرک هم‌انباشته (ARIMA)

باکس و جنکینز (1976) در کتاب خود با عنوان: تجزیه تحلیل سری‌های زمانی، پیش‌بینی و کنترل²²، شکل جدیدی از مدل‌سازی برای پیش‌بینی متغیرهای اقتصادی را معرفی نمودند. در این رویکرد با استخراج ساختار حاکم بر گذشته یک سری زمانی، می‌توان نسبت به توضیح و پیش‌بینی مقادیر آینده اقدام نمود (Box & Jenkins, 1976). در حالت کلی مدل‌سازی سری‌های زمانی، فرض مانایی²³ به‌عنوان اولین گام مطرح می‌شود. در صورت نامانایی²⁴ بودن سری زمانی، برای مدل‌سازی ابتدا باید سری زمانی را با رویکردهایی مانند تفاضل‌گیری مانا کرد. اگر یک سری زمانی با d مرتبه تفاضل‌گیری تبدیل به سری زمانی مانا شود، آن را سری زمانی هم‌انباشته از مرتبه d نامیده و مدل‌سازی از طریق مدل خودرگرسیون میانگین متحرک انباشته $ARIMA(p,d,q)$ که شکل کلی آن به‌صورت روابط

$$\Delta^d y_t = \alpha + \beta_1 \Delta^d y_{t-1} + \dots + \beta_p \Delta^d y_{t-p} + \varepsilon_t + \gamma_1 \varepsilon_{t-1} + \dots + \gamma_q \varepsilon_{t-q} + v_t$$

$$\beta(L) \Delta^d y_t = \alpha + \gamma(L) \varepsilon_t + v_t \quad (1)$$

(2) است، انجام خواهد شد.

$$\Delta^d y_t = \alpha + \beta_1 \Delta^d y_{t-1} + \dots + \beta_p \Delta^d y_{t-p} + \varepsilon_t + \gamma_1 \varepsilon_{t-1} + \dots + \gamma_q \varepsilon_{t-q} + v_t \quad (1)$$

$$\beta(L) \Delta^d y_t = \alpha + \gamma(L) \varepsilon_t + v_t \quad (2)$$

که در اینجا y_t سری زمانی مورد بررسی در زمان t که با وقفه‌های خود (y_{t-1}) ، جمله اختلال (ε_t) و وقفه‌های جمله اختلال (ε_{t-1}) ارتباط دارد. مانایی یک سری زمانی از طریق آزمون‌های دیکی

²² Time Series Analysis: Forecasting and Control

²³ Stationary

²⁴ Non-Stationary

فولر²⁵ و یا فیلیپس پرون²⁶ بررسی و فرضیه وجود ریشه واحد²⁷ در سری زمانی آزمون می‌شود. پس از شناسایی ساختار مانایی سری زمانی، با استفاده از روش سه مرحله‌ای باکس-جنکینز، مدل بهینه برای پیش‌بینی مقادیر آتی انتخاب می‌گردد. باکس-جنکینز یک روش استاندارد سه مرحله‌ای برای ساخت یک مدل ARIMA معرفی کردند.

(1) شناسایی و انتخاب نوع مدل : مرتبه‌های p و q برای مدل‌های AR ، MA یا ARMA تعیین می‌شود.
(2) برآورد ضرایب : ضرایب مدل‌های انتخاب شده برآورد می‌شود.
(3) بررسی مدل : باقیمانده‌های مدل برآورد شده، تجزیه و تحلیل می‌شود.
مرحله برآورد شامل یک فرآیند بهینه‌سازی با استفاده از معیارهای خطا مانند معیار اطلاعات آکائیک (AIC)، بیزین (BIC) و یا معیار حنان‌کوین (HQIC) است. در این پژوهش از معیار AIC استفاده شده است.

$$AIC = -2\ln(\hat{L}) + 2k \quad (3)$$

در اینجا $\ln(\hat{L})$ مقدار تابع درست‌نمایی و K درجه آزادی است. مدلی که کمترین مقدار AIC را داشته باشد بعنوان مدل بهتر در میان مدل‌های مورد بررسی انتخاب می‌شود. روش‌های متفاوتی برای محاسبه تابع درست‌نمایی وجود دارد که در این پژوهش از برآوردگر حداکثر درست‌نمایی²⁸ استفاده شده که با وجود افزایش زمان محاسبات، نتایج دقیق‌تری را ارائه می‌دهد. در نهایت در مرحله بررسی مدل و بررسی باقیمانده‌های آن، در صورتیکه نتایج نشان دهد باقیمانده‌ها حالت استاندارد ندارند، مراحل قبلی تا زمان رسیدن به مدل بهینه تکرار می‌شود.

2-3- ناهمسانی واریانس مشروط و مدل‌های (GARCH)

در اغلب سری‌های زمانی اقتصادی و مالی، دلایلی برای وجود واریانس ناهمسان مشاهده می‌شود. بروز ناهمسانی واریانس، به همراه یک همبستگی خطی میان واریانس‌های شرطی در سری زمانی، باعث شد محققانی نظیر انگل چگونگی مدل‌سازی نوسانات شرطی (واریانس شرطی) را بررسی نمایند. در تحلیل و کشف رابطه ساختاری یک سری زمانی، همواره جزئی به‌عنوان خطا مورد نظر قرار می‌گیرد که مطابق با فروض اولیه در مدل‌سازی، متغیری با واریانس ثابت در طول زمان تعریف می‌شود. حال‌آنکه سری‌های زمانی معمولاً در طول زمان نوسانات شدیدی را تجربه می‌کنند. این موضوع فرض واریانس ثابت یا عبارتی واریانس همسان را مورد تردید قرار می‌دهد. انگل، نشان داد که می‌توان به‌صورت هم‌زمان میانگین و واریانس شرطی یک سری زمانی را مدل‌سازی کرد. هدف اصلی مدل‌سازی نوسانات، استفاده از یک فرآیند ARMA به‌منظور شناسایی هرگونه روند پایدار، در تغییرات جزء خطا است. در صورتیکه یک سری زمانی خودرگرسیون مرتبه اول ($p=1$) را در نظر بگیریم:

$$var(y_t) = E[(y_{t-1} - \alpha_0 - \alpha_1 y_t)^2] = E_t(\varepsilon_{t+1})^2 \quad (4)$$

²⁵ Dicky-Fuller

²⁶ Phillips-Perron

²⁷ Unit Root

²⁸ Maximum Likelihood Estimator

در اینجا $E_t(\varepsilon_{t+1})^2$ یک فرآیند خودرگرسیون از مرتبه p در نظر گرفته می‌شود و براساس رابطه (5) زیر برآورد خواهد شد.

$$\hat{\varepsilon}_t^2 = \beta_0 + \beta_1 \hat{\varepsilon}_{t-1}^2 + \dots + \beta_p \hat{\varepsilon}_{t-p}^2 + u_t \quad (5)$$

در صورتیکه $\beta_1 = \beta_2 = \dots = \beta_p = 0$ باشد، واریانس جمله اختلال ثابت و برابر با β_0 خواهد بود. در غیر این صورت واریانس شرطی γ_t مطابق با فرآیند خودهمبسته معرفی شده در (4) تغییر می‌کند. به این الگو واریانس ناهمسان شرطی خودرگرسیونی گفته می‌شود. جملات اختلال می‌توانند از الگوهای مختلفی پیروی نمایند. لذا شکل‌های متفاوتی از الگوی ARCH وجود دارد. الگوی معرفی شده توسط انگل برای واریانس شرطی توسط بولرسلف (1986) توسعه داده شد (Balerslov, 1986). در این مدل الگوی واریانس شرطی از اجزای ARCH (ε_{t-i}^2) و نیز GARCH (σ_{t-i}^2) به صورت ترکیبی، با مرتبه‌های متفاوت بصورت رابطه $\sigma_t^2 = \alpha_0 + \sum_1^q \alpha_i \varepsilon_{t-i}^2 + \sum_1^p \beta_i \sigma_{t-i}^2$ (6) تعیین می‌شود.

$$\sigma_t^2 = \alpha_0 + \sum_1^q \alpha_i \varepsilon_{t-i}^2 + \sum_1^p \beta_i \sigma_{t-i}^2 \quad (6)$$

که در آن σ_t واریانس شرطی ε_t است. گلاستن، جگنزان و رانکل (1993) فرض عدم تقارن اثر شوک‌های مثبت و منفی بر واریانس‌های شرطی را مطرح کردند (Glosten, Jaganathan & Runkle, 1993). مدل‌های GARCH با این ویژگی به مدل‌های GJR-GARCH مشهور است. در این الگو $\varepsilon_{t-1} = 0$ یک آستانه²⁹ در نظر گرفته می‌شود که شوک‌های بزرگتر از شوک آستانه، تاثیرات متفاوتی نسبت به شوک‌های کوچکتر از شوک آستانه دارند. مدل GARCH در این حالت به صورت رابطه $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 d_{t-1} \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ (7) خواهد بود.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 d_{t-1} \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (7)$$

در اینجا d_{t-1} یک متغیر مجازی است که با ازای $\varepsilon_{t-1} < 0$ برابر با یک و به ازای $\varepsilon_{t-1} \geq 0$ برابر با صفر می‌باشد (Enders, 1948). از دیگر مشکلات مدل‌های GARCH، تضمین مثبت بودن تمامی ضرایب است. نلسون (1991) به منظور بررسی اثر عدم تقارن در اطلاعات یا چولگی‌هایی که اغلب در بازده‌های مالی مشاهده می‌شوند، مدل GARCH نمایی³⁰ را معرفی کرد که شرط نامنفی بودن پارامترها حذف و معادله واریانس شرطی به شکل رابطه $\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i}) + \sum_{j=1}^q \beta_j \ln(\sigma_{t-j}^2)$ (8) معرفی می‌شود (Nelson, 1991).

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i}) + \sum_{j=1}^q \beta_j \ln(\sigma_{t-j}^2) \quad (8)$$

²⁹ Threshold

³⁰ Exponential GARCH (eGARCH).

از دیگر خصوصیات مدل‌های GARCH خصوصیت لپتوکورتوسیس³¹ توزیع تجربی بازده‌های مالی است. محققان اغلب به‌منظور ایجاد توزیع‌های دنباله‌دار، از توزیع خطای تعمیم‌یافته (GED) یا توزیع t استیودنت استفاده می‌کنند که علاوه بر فرض گوسی سنتی، خطاهای ε_t بر اساس توزیع GED یا توزیع t استیودنت توزیع می‌شود. تابع چگالی احتمال توزیع t استیودنت با در نظر گرفتن درجه آزادی ν و با این فرض که برای وجود گشتاور مرتبه دوم، باید درجه آزادی بزرگتر از دو باشد، به‌صورت رابطه $f(\varepsilon_t) =$

$$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)} (\nu - 2)^{-\frac{1}{2}} (h_t)^{-\frac{1}{2}} \left[1 + \frac{\varepsilon_t^2}{h_t(\nu-2)}\right]^{-\frac{(\nu+1)}{2}} \quad (9) \text{ است.}$$

$$f(\varepsilon_t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)} (\nu - 2)^{-\frac{1}{2}} (h_t)^{-\frac{1}{2}} \left[1 + \frac{\varepsilon_t^2}{h_t(\nu-2)}\right]^{-\frac{(\nu+1)}{2}} \quad (9)$$

$$f(\varepsilon_t) = \frac{\nu \exp\left[-\left(\frac{1}{2}\right) \left| \frac{\varepsilon_t}{\lambda h_t^{\frac{1}{2}}}\right|^{\frac{\nu}{\lambda}}\right]}{h_t^{1/2} \lambda^2 \left(1 + \frac{1}{\nu}\right) \Gamma\left(\frac{1}{\nu}\right)}$$

تابع توزیع GED نیز به‌صورت رابطه (10) تعریف می‌شود.

$$f(\varepsilon_t) = \frac{\nu \exp\left[-\left(\frac{1}{2}\right) \left| \frac{\varepsilon_t}{\lambda h_t^{\frac{1}{2}}}\right|^{\frac{\nu}{\lambda}}\right]}{h_t^{1/2} \lambda^2 \left(1 + \frac{1}{\nu}\right) \Gamma\left(\frac{1}{\nu}\right)} \quad (10)$$

در اینجا ν پارامتری است که ضخامت دنباله توزیع را در مقایسه با توزیع نرمال نشان می‌دهد. زمانی که $\nu = 2$ باشد توزیع GED تبدیل به توزیع نرمال استاندارد خواهد شد. برای $\nu < 2$ و $\nu > 2$ توزیع GED به ترتیب دنباله ضخیم‌تر و نازک‌تر از توزیع نرمال خواهد داشت.

3-3- مدل‌سازی داده‌های ترکیبی با توأثر زمانی متفاوت (MIDAS)

در روش‌های سنتی مدل‌سازی سری‌های زمانی برای پیش‌بینی متغیرهای اقتصادی، تمام متغیرهای درگیر در الگو لزوماً از توأثر زمانی یکسانی برخوردارند، به‌عنوان‌مثال چنانچه متغیر وابسته فصلی باشد، متغیرهای توضیح دهنده نیز باید فصلی باشند. حال چنانچه در یک رابطه رگرسیونی متغیرهایی وجود داشته باشند، که توأثرهای زمانی یکسانی نداشته و به‌صورت سالیانه، فصلی، ماهیانه و ... مشاهده و جمع‌آوری شده باشند، بر اساس الگوهای سنتی امکان برآورد ضرایب در این رگرسیون وجود نخواهد داشت، مگر آنکه کلیه متغیرهای مدل با استفاده از روش‌هایی، در یک توأثر زمانی تجمیع و سپس ضرایب رگرسیون برآورد شود. به‌عنوان‌مثال داده‌های فصلی و یا ماهیانه با استفاده از تجمیع زمانی به داده‌های سالیانه تبدیل شوند. ایده برآورد مدل رگرسیون داده‌های با توأثرهای زمانی متفاوت، اولین بار توسط کلاین و سوگو (1989) ارائه

³¹ Leptokurtosis

و بعدها توسط گیسلز (2004) توسعه یافت. این رویکرد امکان انجام برآورد رگرسیون، بدون توجه به توأتر متفاوت در متغیرها را فراهم می‌نماید. ساخت الگویی با این رویکرد دو مزیت عمده را در مدل‌سازی اقتصادی‌سنجی ایجاد می‌نماید. نخست آنکه قرار گرفتن متغیرهای با توأتر زمانی بالا (پُرتوآتر به‌عنوان متغیرهای توضیح‌دهنده) در کنار متغیرهای با توأتر زمانی پایین (کم توأتر به‌عنوان متغیرهای وابسته) در یک مدل و در کنار همدیگر، این امکان را فراهم می‌کند تا پیش‌بینی‌های دقیق‌تری برای متغیر وابسته در آینده‌ای نزدیک ارائه شود. دومین مزیت در زمان ارائه اطلاعات جدید برای متغیر با توأتر زمانی بالا فراهم می‌شود. گیسلز، سینکو و والکانو (2006) رگرسیون ساده MIDAS را به شکل روابط $y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta) x_{t-h}^{(m)} + \varepsilon_t$

$$(11) \beta_1 B(L^{1/m}; \theta) x_{t-h}^{(m)} + \varepsilon_t$$

تعریف

$$(13) x_{t-1-\frac{s}{m}}^{(m)}$$

کردند (Ghysels, Sinko & Valkanov, 2006).

$$y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta) x_{t-h}^{(m)} + \varepsilon_t$$

$$B(L^{1/m}; \theta) = \sum_{j=1}^K b(k; \theta) L^{(j-1)/m} \quad (12)$$

$$L^{s/m} x_{t-1}^{(m)} = x_{t-1-\frac{s}{m}}^{(m)} \quad (13)$$

که t واحد زمانی مرتبط به متغیر وابسته (کم‌توآتر) y_t است و m توأتر زمانی مربوط به متغیر توضیح‌دهنده $x_t^{(m)}$ (پُرتوآتر) است که ارتباط میان این دو متغیر را از طریق تعداد مشاهده متغیر پُرتوآتر در بازه زمانی متغیر کم‌توآتر مشخص می‌کند. به‌عنوان مثال اگر متغیر وابسته (کم‌توآتر) بصورت سالیانه و متغیر توضیح‌دهنده (پُرتوآتر) بصورت فصلی جمع‌آوری شود، مقدار m برابر با 4 خواهد بود. یعنی برای هر مقدار در هر سال از متغیر کم‌توآتر، 4 مشاهده (4 فصل) برای متغیر پُرتوآتر وجود دارد. k تعداد وقفه بهینه برای متغیر پُرتوآتر است که در مدل وارد و مشخص می‌شود که متغیر وابسته با چند وقفه از متغیر پُرتوآتر در ارتباط است. بصورت بنیادی اگر به ازای هر مقدار از متغیر کم‌توآتر، کلیه مقادیر متغیر پُرتوآتر در مدل رگرسیون وارد شود، تعداد ضرایب موردنیاز برای برآورد افزایش خواهد یافت. به‌عنوان مثال به ازای هر مقدار از y_t باید 4 مقدار از متغیر x_t در نظر گرفته شده و 4 ضریب مشخص برآورد شود. با در نظر گرفتن یک بازه 10 ساله برای برآورد مدل مشاهده می‌شود که نیاز به برآورد 40 ضریب برای متغیر توضیح‌دهنده خواهد بود. این موضوع مدل رگرسیون را از نظر فرم و درجه آزادی با مشکلات جدی مواجه می‌نماید. لذا در مدل MIDAS به‌منظور حل این موضوع و ارائه فرم خلاصه‌شده رگرسیون از توابع وزن‌دهی استفاده شده است. $b(k; \theta)$ تابع وزن‌دهی چندجمله‌ای، برای اعمال وزن‌های خاص به وقفه‌های x_t است، که به تعداد وقفه k و بردار θ وابسته می‌باشد و در حالت کلی بصورت رابطه

$$b(k; \theta) =$$

$$f\left(\frac{k}{k^{max}}, \theta\right) / \sum_{i=1}^{k^{max}} f\left(\frac{k}{k^{max}}, \theta\right) \quad (14)$$

تعریف می

شود.

$$b(k; \theta) = f\left(\frac{k}{k^{max}}, \theta\right) / \sum_{i=1}^{k^{max}} f\left(\frac{k}{k^{max}}, \theta\right) \quad (14)$$

گیسلز ضمن معرفی فرم کلی برای تابع وزن‌دهی مطابق با معادله شماره 13، نمونه‌هایی از توابع وزن‌دهی مانند تابع وزن‌دهی المون، تابع وزن‌دهی المون نمایی، تابع وزن‌دهی بتا، به ترتیب روابط

$$\beta \cdot b(k; \theta) = \sum_{k=0}^{k^{max}} \sum_{p=0}^P \theta_p \cdot k^p \quad (15)$$

$$b\left(\frac{k}{m}; \theta_1; \theta_2\right) = \frac{F\left(\frac{k}{m}; \theta_1, \theta_2\right)}{\sum_{k=1}^{k^{max}} F\left(\frac{k}{m}; \theta_1, \theta_2\right)} \quad (17)$$

معرفی نمود. در تابع وزن‌دهی المون ضرایب و وزن‌ها بصورت یک پارامتر مشترک برآورد می‌شود.

$$\beta \cdot b(k; \theta) = \sum_{k=0}^{k^{max}} \sum_{p=0}^P \theta_p \cdot k^p \quad (15)$$

$$b(k; \theta) = \frac{\exp(\theta_1 \cdot k + \theta_2 \cdot k^2)}{\sum_{k=1}^{k^{max}} \exp(\theta_1 \cdot k + \theta_2 \cdot k^2)} \quad (16)$$

$$b\left(\frac{k}{m}; \theta_1; \theta_2\right) = \frac{F\left(\frac{k}{m}; \theta_1, \theta_2\right)}{\sum_{k=1}^{k^{max}} F\left(\frac{k}{m}; \theta_1, \theta_2\right)} \quad (17)$$

$$b\left(\frac{k}{m}; \theta_1; \theta_2\right) = \frac{F\left(\frac{k}{m}; \theta_1, \theta_2\right)}{\sum_{k=1}^{k^{max}} F\left(\frac{k}{m}; \theta_1, \theta_2\right)}$$

که در رابطه

$$F(x, a, b) = \frac{x^{a-1}(1-x)^{b-1}\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \quad (17)$$

تابع F به صورت رابطه

(18) تعریف می‌شود.

$$F(x, a, b) = \frac{x^{a-1}(1-x)^{b-1}\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \quad (18)$$

که $\Gamma(a)$ و $\Gamma(b)$ و $\Gamma(a+b)$ توابع گاما بوده و بصورت رابطه $\Gamma(a) = \int_0^{\infty} e^{-x} \cdot x^{a-1}$ (19) تعریف می‌شوند.

$$\Gamma(a) = \int_0^{\infty} e^{-x} \cdot x^{a-1} \quad (19)$$

شکل تابع وزن دهی و حداکثر تعداد وقفه های بهینه متغیر پرتواتر در هر مدل رگرسیون می تواند متفاوت باشد؛ اما به صورت کلی توابع وزن دهی معرفی شده بر اساس پارامترهای تعریف شده، وزن های غیرمنفی ایجاد می کنند و برای تعیین مقدار ضریب متغیر پرتواتر و وقفه های آن (β)، از فرض برابر یک بودن مجموع وزن های ایجاد شده در مدل رگرسیون استفاده می شود. از نظر پارامتریک الگوی معرفی شده MIDAS یک الگوی خطی است که با وارد کردن تابع وزن دهی و اعمال وزن به وقفه های متغیر پرتواتر، از حالت خطی به غیرخطی تبدیل می شود. بر اساس مطالعه گیسلز و همکارانش لازم است از روش های برآورد غیرخطی³² و حداقل نمودن مجموع مربعات جملات اختلال برای تخمین ضرایب مدل معرفی شده استفاده کرد.

3-4- معیارهای ارزیابی پیش بینی مدل های رگرسیون

پس از برآورد مدل های رگرسیون، به منظور مقایسه قدرت پیش بینی جهت انتخاب بهترین مدل برای پیش بینی یک سری زمانی می توان از معیارهای مختلفی استفاده نمود. معیارهای مورد استفاده جهت ارزیابی مدل برآورد شده و دقت پیش بینی، در این مطالعه به صورت روابط

$$\text{Sum of Square Error : } SSE = \sum_{t=1}^T (y_t - \hat{y}_t)^2 \quad (20) \quad \text{تا}$$

$$\text{(26) Mean Absolute Percentage Error : MAPE} = \frac{100 \sum_{t=T+1}^{T+m} \left| \frac{y_t^f - y_t}{y_t} \right|}{m}$$

می باشند.

$$\text{Sum of Square Error : } SSE = \sum_{t=1}^T (y_t - \hat{y}_t)^2 \quad (20) \quad ($$

$$\text{Relative Square Error : } RSE = \frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{\sum_{t=1}^T (y_t - \bar{y}_t)^2} \quad (21) \quad ($$

$$\text{Root Relative Square Error : } RSE = \sqrt{\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{\sum_{t=1}^T (y_t - \bar{y}_t)^2}} \quad (22) \quad ($$

$$\text{Mean Squared Error : } MSE = \frac{\sum_{t=T+1}^{T+m} (y_t^f - y_t)^2}{m} \quad (23) \quad ($$

$$\text{Root Mean Square Error : } RMSE = \sqrt{\frac{\sum_{t=T+1}^{T+m} (y_t^f - y_t)^2}{m}} \quad (24) \quad ($$

$$\text{Mean Absolute Error : } MAE = \frac{\sum_{t=T+1}^{T+m} |y_t^f - y_t|}{m} \quad (25) \quad ($$

$$\text{Mean Absolute Percentage Error : } MAPE = \frac{100 \sum_{t=T+1}^{T+m} \left| \frac{y_t^f - y_t}{y_t} \right|}{m} \quad (26) \quad ($$

m طول دوره پیش بینی، y_t مقادیر واقعی و y_t^f مقادیر پیش بینی شده است. روابط

$$\text{Sum of Square Error : } SSE = \sum_{t=1}^T (y_t - \hat{y}_t)^2$$

$$\text{Root Relative Square Error : } RSE = \sqrt{\frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{\sum_{t=1}^T (y_t - \bar{y}_t)^2}} \quad \text{تا (20)}$$

$$\text{Mean Squared Error : } MSE = \text{خطای مدل برآورد شده و روابط (22)}$$

$$\text{تا (23)} \quad \frac{\sum_{t=T+1}^{T+m} (y_t^f - y_t)^2}{m}$$

$$(26) \text{Mean Absolute Percentage Error: } MAPE = \frac{100 \sum_{t=T+1}^{T+m} \left| \frac{y_t^f - y_t}{y_t} \right|}{m}$$

میزان خطای پیش‌بینی را مشخص می‌نماید. همچنین به منظور مقایسه صحت پیش‌بینی از آزمون برابری توانایی پیش‌بینی که توسط دیبولد و ماریانو (1995) ارائه شده، استفاده می‌شود. این آزمون بر اساس قدرت پیش‌بینی را تحت فرض برابری قدرت پیش‌بینی دو مدل بررسی می‌کند (Diebold-Mariano, 1995).

$$DM = d / \sqrt{v_d / F} \quad \text{آماره دیبولد-ماریانو بر اساس رابطه (27) بدست می‌آید:}$$

$$DM = d / \sqrt{v_d / F} \quad (27)$$

$$d = (\sum_{t=1}^F d_t) / F \quad (28)$$

$$v_d = \gamma_0 + 2 \sum_{t=1}^{\infty} \gamma_t \quad (29)$$

$$\gamma_t = cov(d_t, d_{t-1}) \quad (30)$$

4- داده‌ها و نتایج تجربی

در سال‌های گذشته، بیت‌کوین به‌عنوان ارز دیجیتال پیشرو در جهان، مورد توجه مصرف‌کنندگان، کسب‌وکارها و سرمایه‌گذاران قرار گرفته است. روند قیمت بیت‌کوین نشان می‌دهد که علاوه بر حجم بالای عرضه و تقاضا، این رمز ارز نوسانات شدید قیمتی را نیز در طول زمان تجربه می‌کند. تاکنون مطالعات بسیاری بر روی تأثیر بیت‌کوین بر بازار و قیمت سایر دارایی‌ها انجام شده است. نتایج این مطالعات می‌تواند در راستای مقابله با آسیب‌پذیری‌های ناشی از نوسانات شدید بیت‌کوین مورد استفاده قرار گیرد. ارزش فعلی بیت‌کوین به‌عنوان یک دارایی و همچنین مقبولیت ایجاد شده برای این رمز ارز، انگیزه‌ای برای بررسی و تحلیل ساختار تغییرات قیمت این رمز ارز و کشف الگوهایی برای پیش‌بینی قیمت آن را ایجاد می‌نماید. در این مطالعه داده‌های میان روزانه‌های قیمت بیت‌کوین با فواصل زمانی 5 دقیقه‌ای و در بازه زمانی ابتدای سال 2015 میلادی تا ششمین ماه سال 2023 میلادی جمع‌آوری شده است. 33 کلیه محاسبات انجام شده در این مطالعه توسط زبان برنامه‌نویسی و محاسباتی R و بسته‌های نرم‌افزاری مرتبط صورت پذیرفته است.

33 اطلاعات جمع‌آوری شده از طریق ایجاد اکانت کاربری در وب سایت (<https://www.cryptodatadownload.com>) و دریافت داده‌های 1 دقیقه‌ای انجام و سپس پردازش لازم با استفاده از نرم‌افزار R صورت پذیرفته است. به منظور قابل استفاده بودن داده‌های روزانه و هفتگی در مدل تواتر متفاوت، به صورت 2015/01/27 الی 2023/05/08 مرتب شده است.



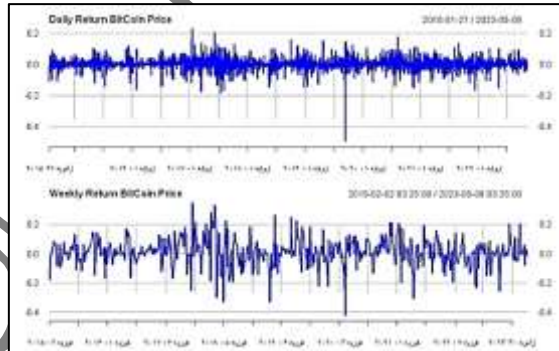
نمودار 1. روند قیمت بیت‌کوین (2015 تا 2023)
 ماخذ: (<https://www.cryptodatadownload.com>)
 Figure 1.Bitcoin Daily Price Trend (2015 to 2023)
 Source: (<https://www.cryptodatadownload.com>)

متفاوتی را در بازه‌های زمانی مختلف تجربه کرده است. بازه قیمت بیت کوین در یک بازه زمانی مشخص از رابطه

$$R_t = \ln(P_t) - \ln(P_{t-1})$$

(31) محاسبه می‌شود.

$$R_t = \ln(P_t) - \ln(P_{t-1}) \quad (31)$$



نمودار 2. بازدهی روزانه و هفتگی بیت‌کوین از سال 2015 تا 2023
 ماخذ: نتایج پژوهش
 Figure 2.Bitcoin Daily and Weekly Returns (2015 to 2023)
 Source: Research Results

که در اینجا R_t بازدهی در زمان t است که می‌تواند بازدهی میان روزانه (اختلاف میان t و $t-1$ یک زمان 5 دقیقه‌ای میان روزانه)، بازدهی روزانه (اختلاف میان t و $t-1$ دو روز متوالی و بازدهی هفتگی (اختلاف میان t و $t-1$ دو روز ابتدایی برای دو هفته متوالی) در نظر گرفته‌شود. در این حالت در هر روز 288 داده 5 دقیقه‌ای برای قیمت بیت‌کوین و 288 داده برای بازدهی میان روزانه بیت‌کوین

محاسبه می‌گردد. **Error! Reference source not found.** بازدهی روزانه (بالا) و بازدهی هفتگی (پایین) را نشان می‌دهد.

جدول 1. خلاصه اطلاعات سری زمانی قیمت روزانه، میان روزانه (5 دقیقه‌ای)، بازدهی روزانه و هفتگی بیت‌کوین
ماخذ: نتایج پژوهش

Table 1. Time Series Information Summary of Daily, Inter-daily (5-minute), Daily and Weekly Returns of Bitcoin
Source: Research Results

	BTC 5-min (intardaily)	BTC (Daily)	Return BTC (Daily)	Return BTC (Weekly)
تعداد مشاهدات	870912	3024	3024	432
کمترین	-0.2026	211.4	-0.4972	-0.4251
بیشترین	0.1167	67527.9	0.2276	0.3477
چارک اول	-0.00086	1040.9	-0.0128	-0.0318
چارک سوم	0.00090	19958.3	0.0177	0.0656
میانگین	5 e-0.06	13892.77	0.0015	0.0107
میانه	0	7936.05	0.0015	0.0118
واریانس	7 e-0.06	260604379	0.0014	0.0102
انحراف معیار	0.00259	16143.24	0.0381	0.1012
کشیدگی	-1.1339	1.399711	-0.8347	-0.3621
چولگی	102.8630	0.9535	12.74	1.6101

خلاصه اطلاعات مربوط به سری زمانی قیمت روزانه، قیمت میان روزانه‌ی 5 دقیقه‌ای، بازدهی روزانه و بازدهی هفتگی در بازه زمانی سال 2015 میلادی تا سال 2023 میلادی در **جدول 1** آورده شده است.

4-1- آزمون مانایی سری زمانی

در اولین گام برای تجزیه و تحلیل سری زمانی قیمت بیت‌کوین، مانایی این سری زمانی مورد بررسی قرار می‌گیرد. نتایج آزمون دیکی فولر تعمیم‌یافته و آزمون فیلیپس پرون وجود ریشه واحد در سری زمانی قیمت روزانه بیت‌کوین و همچنین فرم لگاریتمی آن و نامانایی این دو سری زمانی را تأیید می‌کنند.

جدول 2. آزمون ریشه واحد (سری زمانی قیمت بیت‌کوین، فرم لگاریتمی قیمت، بازدهی روزانه و بازدهی هفتگی)
ماخذ: نتایج پژوهش

Table 2. Unit Root Test
Source: Research Result

آزمون	آماره	قیمت روزانه	لگاریتم قیمت روزانه	بازدهی روزانه	بازدهی هفتگی
دیکی فولر تعمیم یافته	Dickey-Fuller	-2.1647	-1.3952	-13.46	-6.6671
	p-value	0.5086	0.8344	0.01*	0.01*
فیلیپس پرون	Dickey-Fuller Z(alpha)	-8.8562	-3.933	-3293.5	-406.1182
	p-value	0.6158	0.8904	0.01*	0.01*

نتایج این دو آزمون در

جدول 2 نشان می‌دهد که سری‌های زمانی مورد استفاده در این پژوهش، یعنی سری‌های زمانی بازدهی روزانه و سری‌های زمانی بازدهی هفتگی بیت‌کوین مانا هستند.

4-2- انتخاب مدل ARIMA-GARCH

به منظور انتخاب مرتبه‌های p و q در الگوی ARIMA، ضمن بررسی شهودی توابع خودهمبستگی³⁴ و خودهمبستگی جزئی³⁵، یک مجموعه مدل با مرتبه‌های متفاوت از p و q برآورد و معیارهای اطلاعات³⁶ برای آن مجموعه محاسبه می‌شود. بهترین مدل ARIMA در این بخش مدلی است که کمترین مقدار آماره آکائیک را داشته باشد. نرم‌افزار R با استفاده از بسته‌های tseries و forecast امکان برآورد مدل ARIMA با مراتب مختلف را فراهم می‌کند. همچنین تابع auto.arima به صورت پیش فرض بهترین مدل ARIMA را برای یک سری زمانی معرفی می‌نماید. مطابق با نتایج، برای سری زمانی بازدهی روزانه بیت‌کوین، در مدل ARIMA(1,0,2) ضرایب برآورد شده در سطح اطمینان 95 درصد معنادار بوده و این مدل کمترین مقدار معیار آکائیک را دارد. همچنین برای سری زمانی بازدهی هفتگی بیت‌کوین، ضرایب مدل ARIMA(2,0,2) در سطح اطمینان 95 درصد معنادار و مقدار آماره آکائیک، کمترین مقدار است.

جدول 3. مدل ARIMA بر اساس معیار آکائیک
ماخذ: نتایج پژوهش

Table 3. ARIMA Model Based on Akaike Criterion
Source: Research Result

سری زمانی بازدهی روزانه بیت‌کوین ARIMA(1,0,2)				
	Estimate	Std. Error	z value	Pr(> z)
ar1	0.857	0.237	3.60	0.0003***
ma1	-0.896	0.238	-3.76	0.0001***
ma2	0.053	0.018	2.91	0.0035**
intercept	0.0015	0.0007	1.98	0.0469'
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
AIC= -11173.43		log likelihood =5590.715		sigma^2 =0.001451
سری زمانی بازدهی هفتگی بیت‌کوین ARIMA(2,0,2)				
	Estimate	Std. Error	z value	Pr(> z)
ar1	-1.025	0.010	-98.51	2e-16***
ar1	-0.986	0.009	-108.08	2e-16***
ma1	1.007	0.010	91.85	2e-16***
ma2	0.9999	0.01307	76.5047	2e-16***
intercept	0.0111	0.00493	2.2558	0.02408'
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
AIC= -726.69		log likelihood =368.29		sigma^2 =0.01053

نتایج برآورد مدل به همراه معناداری ضرایب برآورد شده و معیار آکائیک در **Error! Reference source not found.** آورده شده است. بررسی شهودی باقیمانده‌های مدل‌های برآورد شده برای سری زمانی بازدهی روزانه و هفتگی بیت‌کوین نشان‌دهنده وجود نوسان در سری‌های زمانی مذکور است. به منظور بررسی دقیق وجود ناهمسان‌واریانس مشروط، از آزمون ARCHTEST در نرم‌افزار R استفاده می‌شود.

جدول 4. آزمون وجود اثرات ARCH
ماخذ: نتایج پژوهش

Table 4. ARCH Effect Test
Source: Research Result

ARCH LM-test; Null hypothesis: no ARCH effects
--

³⁴ Auto Correlation Function

³⁵ Partial Auto Correlation Function

³⁶ Information Criteria

	بازدهی روزانه	بازدهی هفتگی
Chi-squared	67.5498	21.37
p-value	9.17e-10	0.046

نتیجه آزمون در جدول **Error! Reference source not found.** نشان‌دهنده رد فرض صفر، مبنی بر عدم وجود اثر واریانس ناهمسان برای سری‌های زمانی بازدهی روزانه و هفتگی بیت‌کوین است. در مرحله بعدی با در نظر گرفتن مدل پایه $GARCH(1,1)$ ، هر یک از سری‌های زمانی بازدهی روزانه و هفتگی بیت‌کوین در 3 حالت $sGARCH$ ، $eGARCH$ و $gjrgARCH$ برآورد می‌گردند.³⁷ همچنین در هر یک از مدل‌های برآورد شده 5 حالت توزیع نرمال، توزیع نرمال چوله، توزیع استیودنت، توزیع استیودنت چوله و توزیع خطای تعمیم یافته در نظر گرفته شده است که در مجموع 15 مدل برای هر سری زمانی در این مرحله برآورده گردیده است. به منظور آماده‌سازی داده‌ها برای مقایسه دقت پیش‌بینی مدل‌های مختلف، داده‌های روزانه (3024 داده) و هفتگی (432 داده هفتگی) به دو بخش تقسیم شده است. بازدهی روزانه از ابتدای سال 2015 تا انتهای سال 2022 (2528) در برآورد مدل مورد استفاده قرار گرفته و از ابتدای 2023 تا 2023/05/08 (496 داده) به منظور پیش‌بینی مورد استفاده قرار گرفته‌اند. همچنین برای بازدهی هفتگی، اطلاعات هفتگی تا پایان سال 2022 در مدل برآورد شده مورد استفاده قرار گرفته (413 داده هفتگی) و اطلاعات سال 2023 به منظور سنجش دقت پیش‌بینی از مدل خارج شده است.

جدول 5. اطلاعات مدل $ARIMA(1,2)-GARCH(1,1)$ - بازدهی روزانه بیت‌کوین
ماخذ: نتایج پژوهش

Table 5. Data from $ARIMA(1,2)-GARCH(1,1)$ Model - Bitcoin Daily Return
Source: Research Result

gjrgARCH			eGARCH			sGARCH			Log likelihood	AIC
GED	Skwe-t-Student	t-Student	Skew-normal	Skwe-t-Student	t-Student	GED	Skew-normal	t-Student		
5189	5177	5177	4824	4818	5200	4862	4822	5174	4815	-3.80
-4.08	-4.08	-4.08	-3.80	-3.80	-4.10	-3.81	-3.80	-4.08	-3.80	-3.80

جدول 6. اطلاعات مدل $ARIMA(2,2)-GARCH(1,1)$ - بازدهی هفتگی بیت‌کوین
ماخذ: نتایج پژوهش

Table 6. Data from $ARIMA(2,2)-GARCH(1,1)$ Model - Bitcoin Weekly Returns
Source: Research Result

gjrgARCH	eGARCH	sGARCH
----------	--------	--------

³⁷ به منظور تخمین مدل مذکور در نرم‌افزار R از بسته نرم‌افزاری rugarch استفاده می‌شود.

normal	360.27	-1.71
Skew-normal	364.94	-1.72
tStudent	384.90	-1.82
Skwe-TStudent	385.064	-1.82
GED	385.52	-1.82
normal	362.16	-1.71
Skew-normal	366.18	-1.71
tStudent	393.71	-1.86
Skwe-TStudent	403.17	-1.90
GED	389.91	-1.84
normal	362.17	-1.71
Skew-normal	364.29	-1.72
tStudent	391.60	-1.85
Skwe-TStudent	391.71	-1.84
GED	388.44	-1.83

مقادیر معیار آکانیک و لگاریتم درست‌نمایی برای مدل‌های برآورد شده در

جدول 5 و

جدول 6 آورده شده است.³⁸ این نتایج نشان می‌دهد برای بازدهی روزانه بیت‌کوین مدل $ARIMA(1,2)-eGARCH(1,1)$ با توزیع خطای تی‌استیودنت و برای بازدهی هفتگی بیت‌کوین مدل $ARIMA(2,2)-eGARCH(1,1)$ با توزیع خطای تی‌استیودنت چوله دارای کمترین مقدار در معیار آکانیک و بالاترین مقدار لگاریتم درست‌نمایی را دارد لذا این دو مدل در میان مدل‌های خانواده GARCH بعنوان بهترین مدل برای بازدهی روزانه و هفتگی بیت‌کوین انتخاب می‌شوند. نتیجه آزمون لیونگ باکس³⁹ روی باقیمانده‌های استاندارد شده در تمام مدل‌های برآورد شده، نشان‌دهنده عدم وجود خودهمبستگی میان باقیمانده‌ها می‌باشد.

3-4- مدل داده‌های ترکیبی با توأثر زمانی متفاوت

در این بخش با استفاده از رویکرد جدید مدل‌سازی داده‌های ترکیبی با توأثر زمانی متفاوت، سری‌های زمانی بازدهی روزانه و هفتگی بیت‌کوین بررسی و مدل‌سازی می‌شود. بدین منظور از داده‌های بازدهی روزانه و هفتگی که با استفاد از بازدهی‌های میان‌روزانه (5 دقیقه‌ای) ساخته می‌شوند، استفاده شده است. با توجه به انجام معاملات در بازار رمز ارزها به صورت 7 روز هفته، مدل‌سازی در این بخش با در نظر گرفتن متغیر $m=288$ برای متناظر کردن داده‌های پُرتوآثر (بازدهی میان‌روزانه 5 دقیقه‌ای) بعنوان متغیر توضیح دهنده با داده‌های کم‌توآثر (بازدهی روزانه) بعنوان متغیر وابسته و همچنین $m=7$ برای ایجاد تناظر میان داده‌های روزانه (بعنوان متغیر توضیح دهنده) و هفتگی بعنوان متغیر وابسته تعریف می‌شود. توابع وزن‌دهی معمول در الگوی MIDAS توابع آلمون، آلمون نمایی و بتا می‌باشند که هر سه تابع وزن‌دهی بصورت جداگانه در مدل وارد و نتایج برای پیش‌بینی سری زمانی بازدهی روزانه و هفتگی مورد استفاده قرار گرفته است. کلیه برآوردها از طریق بسته نرم‌افزاری `midas_r` که توسط گیسلز و همکاران توسعه داده شده و بصورت اتوماتیک مقادیر اولیه و تعداد وقفه بهینه را انتخاب می‌کند، انجام شده است. با توجه به اینکه هدف نهایی این مطالعه سنجش قدرت پیش‌بینی مدل است، از میان مدل‌های موجود، بهترین مدل با کمترین مقدار آکانیک در هر مرحله بصورت خودکار توسط نرم‌افزار انتخاب می‌شود.

جدول 7 نتایج وقفه بهینه و معیار آکانیک در مدل MIDAS

³⁸ به منظور خلاصه سازی متن، ضرایب و جزئیات مدل‌های برآورد شده در پیوست آورده شده است.

³⁹ Ljung-Box Test

ماخذ : نتایج پژوهش

Table 7 .Results of Optimal Lag and Akaike's Criterion in MIDAS Model
Source : Research Result

سری زمانی بازدهی روزانه			
nbeta	nealmon	almonp	K
۰:۱۵۵	۰:۱۹۹	0:287	
-11295.32	-12560.99	-22979.88	AIC
سری زمانی بازدهی هفتگی			
nbeta	nealmon	almonp	K
0:2	0:2	0:2	
-876.53	-876.53	-876.92	AIC

نتایج تعداد وقفه‌های بهینه مطابق با کمترین مقدار برای معیار آکائیک مربوط به برآورد سری‌های زمانی بازدهی روزانه و هفتگی با استفاده از هر سه تابع وزن‌دهی معرفی شده در جدول 7 آورده شده است.

4-4- پیش‌بینی

هدف اصلی در این مطالعه بررسی دو رویکرد در مدل‌سازی سری‌های زمانی، به منظور تعیین مدل بهینه برای پیش‌بینی است. در این بخش، مدل بهینه بر اساس کمترین میزان خطا و بیشترین مقدار دقت انتخاب می‌شود. به‌منظور انجام پیش‌بینی، محدوده داده‌های جمع‌آوری‌شده به دو بخش آموزش 40 و آزمون 41 تقسیم‌بندی و هر دو الگوی مورد نظر با در نظر گرفتن داده‌های بخش آموزش برآورد می‌گردد. سپس با استفاده از مدل برآورد شده، مقادیر سری زمانی برای بخش آزمون پیش‌بینی و با مقادیر واقعی مقایسه می‌گردد. به‌منظور یکسان بودن بازه زمانی موردنظر جهت پیش‌بینی، کلیه داده‌های مربوط به سال 2023 از نمونه خارج گردیده و هر دو رویکرد برای بازه زمانی ابتدای 2015 تا انتهای 2022 برآورد شده است. پیش‌بینی برای مدل‌های خانواده ARIMA-GARCH در 3 حالت eGARCH، sGARCH و gjrGARCH و هر یک برای 5 حالت توزیع نرمال، توزیع نرمال چوله، توزیع استیودنت، توزیع استیودنت چوله و توزیع خطای تعمیم یافته، با تکنیک پنجره غلطان 42 یک گام رو به جلو انجام شده است. 43 در بخش مدل‌سازی با داده‌های با تواتر متفاوت پیش‌بینی برای سه تابع وزن‌دهی آلمون، آلمون نمایی و بتا، براساس الگوی پنجره غلطان و الگوی بازگشتی 44 انجام و در مجموع 6 مدل در این بخش برآورد و پیش‌بینی شده است. 45

5- نتایج ارزیابی دقت پیش‌بینی

نتایج معیارهای ارزیابی میزان خطای پیش‌بینی برای تمام مدل‌های برآورد شده روی سری زمانی بازدهی روزانه، در Error! Reference source not found. آورده شده است. 46. نتایج این مرحله نشان می‌دهد برای سری زمانی بازدهی روزانه، مدل برآورد شده با الگوی MIDAS و تابع وزن‌دهی آلمون، از نظر دقت در پیش‌بینی برتری کامل نسبت به مدل‌های خانواده ARIMA-GARCH با توزیع‌های مختلف و مدل‌های MIDAS با توابع وزن‌دهی بتا و آلمون نمایی دارد. در دو معیار قدرمطلق میانگین خطا (MAE)

⁴⁰ Train

⁴¹ Test

⁴² Rolling window

⁴³ از تابع ugarchforecast استفاده شده است.

⁴⁴ Recursive

⁴⁵ از تابع average_forecast در بسته نرم‌افزاری midas_r استفاده شده است.

⁴⁶ به‌منظور محاسبه معیارهای مذکور از بسته Metrics در نرم‌افزار R استفاده شده است.

ماریانو محاسبه می شود. آزمون دیبولد ماریانو ابتدا برای پیش بینی بدست آمده از مدل های خانواده ARIMA-GARCH اجرا شده است.

جدول 9. نتایج آزمون دیبولد ماریانو برای سری زمانی بازدهی روزانه (تمام مدل ها)
ماخذ: نتایج پژوهش

Table 9 .Diebold Mariano Test Results for Daily Return Time Series (all Models)
Source : Research Result

ARIMA(1,2) - GARCH (1,1)					آماره	تابع وزندهی/ توزیع خطا/ الگوی پیش بینی		مدل
Ged	S.Student	Student	S.Normal	Normal				
0.21	0.47	0.44	0.96	---	pvalue	Normal	ARIMA (1,2) - GARCH (1,1)	
0.81	0.08	0.16	-1.72	---	statistic			
0.10	0.19	0.17	---	0.04	pvalue	S.Normal		
1.30	0.86	0.95	---	1.72	statistic			
0.08	0.90	---	0.83	0.56	pvalue	Student		
1.43	-1.26	---	-0.95	-0.16	statistic			
0.06	---	0.10	0.81	0.53	pvalue	S.Student		
1.52	---	1.26	-0.86	-0.08	statistic			
---	0.94	0.92	0.90	0.79	pvalue	Ged		
---	-1.52	-1.43	-1.30	-0.81	statistic			
0.00	0.00	0.00	0.00	0.00	pvalue	Rolling Forecast	MIDAS	
8.42	8.49	8.49	8.54	8.54	statistic			
0.00	0.00	0.00	0.00	0.00	pvalue	Recursive Forecast		
8.42	8.49	8.49	8.54	8.54	statistic			
0.00	0.00	0.00	0.00	0.00	pvalue	Rolling Forecast	beta	
4.45	4.53	4.53	4.55	4.59	statistic			
0.00	0.00	0.00	0.00	0.00	pvalue	Recursive Forecast		
5.02	5.28	5.29	5.39	5.47	statistic			
0.19	0.25	0.24	0.31	0.24	pvalue	Rolling Forecast		
0.88	0.67	0.69	0.48	0.72	statistic			
0.18	0.24	0.24	0.30	0.23	pvalue	Recursive Forecast		
0.91	0.70	0.72	0.51	0.75	statistic			

نتایج

جدول 9 برای سری زمانی بازدهی روزانه نشان می دهد که در مدل سازی با رویکرد -ARIMA GARCH استفاده از توزیع خطای نرمال چوله، نسبت به توزیع خطای نرمال برتری دارد. در مرحله بعدی رویکرد مدل سازی -ARIMA-GARCH با رویکرد MIDAS برای پیش بینی سری زمانی بازدهی روزانه بیت کوین مورد مقایسه قرار گرفته است. نتایج در این بخش این آزمون نشان می دهد پیش بینی های رویکرد مدل سازی MIDAS با توابع وزندهی آلمون و آلمون نمایی و در هر دو حالت پنجره غلتان و بازگشتی، نسبت به تمام مدل های خانواده ARIMA-GARCH بهتر عمل می کنند.

جدول 10. نتایج آزمون دیبولد ماریانو برای سری زمانی بازدهی روزانه (مدل های MIDAS)
ماخذ: نتایج پژوهش

Table 10 .Diebold Mariano Test Results for Daily Return Time Series (MIDAS Models)
Source : Research Result

مدل	تابع وزندهی	الگوی پیش بینی	آماره	MIDAS
-----	-------------	----------------	-------	-------

beta		nealmon					
Rolling Forecast	Rolling Forecast	Rolling Forecast	Rolling Forecast				
0.00	0.00	0.00	0.00	pvalue	Rolling Forecast	almon	MIDAS
8.70	8.71	8.73	9	statistic	Recursive Forecast		
0.00	0.00	0.00	0.00	pvalue	Rolling Forecast	nealmon	
8.71	8.71	8.73	9	statistic	Recursive Forecast		
0.00	0.00	---	---	pvalue	Rolling Forecast	beta	
4.78	4.78	---	---	statistic	Recursive Forecast		
0.00	0.00	---	---	pvalue	Rolling Forecast	beta	
5.49	5.47	---	---	statistic	Recursive Forecast		
---	---	1	1	pvalue	Rolling Forecast	beta	
---	---	-5.47	-4.78	statistic	Recursive Forecast		
---	---	1	1	pvalue	Rolling Forecast	beta	
---	---	-5.49	-4.78	statistic	Recursive Forecast		

با توجه به این نتایج، مجدداً برای سری زمانی بازدهی روزانه بیت‌کوین آزمون دیبولد ماریانو برای مدل‌های MIDAS با توابع وزن‌دهی مختلف اعمال و در نتایج

جدول 9 برای سری زمانی بازدهی روزانه نشان می‌دهد که در مدل‌سازی با رویکرد ARIMA-GARCH استفاده از توزیع خطای نرمال چوله، نسبت به توزیع خطای نرمال برتری دارد. در مرحله بعدی رویکرد مدل‌سازی ARIMA-GARCH با رویکرد MIDAS برای پیش‌بینی سری زمانی بازدهی روزانه بیت‌کوین مورد مقایسه قرار گرفته است. نتایج در این بخش این آزمون نشان می‌دهد پیش‌بینی‌های رویکرد مدل‌سازی MIDAS با توابع وزن‌دهی المون و المون نمایی و در هر دو حالت پنجره غلتان و بازگشتی، نسبت به تمام مدل‌های خانواده ARIMA-GARCH بهتر عمل می‌کنند.

جدول 10 مشخص می‌شود که مدل MIDAS با تابع وزن‌دهی المون نسبت به مدل MIDAS با توابع المون نمایی و بتا بهتر عمل می‌کند این نتایج، عملکرد بهتر مدل MIDAS با تابع وزن‌دهی المون در پیش‌بینی سری زمانی بازدهی روزانه را نشان می‌دهد که نتایج بدست‌آمده در بخش بررسی معیارهای خطای پیش‌بینی را نیز تایید می‌کند.

جدول 11 نتایج آزمون دیبولد ماریانو برای سری زمانی بازدهی هفتگی (تمام مدل‌ها) ماخذ: نتایج پژوهش

Table 11 .Diebold Mariano Test Results for Weekly Return Time Series (all Models)
Source : Research Result

ARIMA (2,2)- GARCH (1,1)					آماره	تابع وزن‌دهی/ توزیع خطای الگوی پیش‌بینی	مدل
Ged	S.Student	Student	S.Normal	Normal			
0.92	0.52	0.97	0.24	---	pvalue	Normal	ARIMA (2,2)- GARCH (1,1)
-1.45	0.073	-1.94	0.69	---	statistic		
0.96	0.91	0.95	---	0.75	pvalue	S.Normal	
-1.87	-1.37	-1.68	---	-0.69	statistic		
0.59	0.15	---	0.04	0.02	pvalue	Student	
-0.24	1.03	---	1.68	1.94	statistic		
0.91	---	0.85	0.08	0.47	pvalue	S.Student	
-1.37	---	-1.03	1.37	0.07	statistic		
---	0.08	0.40	0.03	0.07	pvalue	Ged	

---	1.37	0.24	1.87	1.45	statistic			
0.04	0.03	0.04	0.03	0.03	pvalue	Rolling Forecast	almon	MIDAS
1.68	1.77	1.68	1.82	1.77	statistic	Recursive Forecast		
0.04	0.03	0.04	0.03	0.03	pvalue	Rolling Forecast	nealmon	
1.72	1.81	1.73	1.87	1.81	statistic	Recursive Forecast		
0.03	0.02	0.03	0.02	0.02	pvalue	Rolling Forecast	beta	
1.80	1.89	1.81	1.94	1.89	statistic	Recursive Forecast		
0.03	0.02	0.03	0.02	0.02	pvalue	Rolling Forecast	beta	
1.81	1.90	1.81	1.95	1.90	statistic	Recursive Forecast		
0.63	0.40	0.62	0.28	0.39	pvalue	Rolling Forecast	beta	
-0.35	0.19	-0.32	0.56	0.26	statistic	Recursive Forecast		
0.66	0.43	0.65	0.28	0.40	pvalue	Rolling Forecast	beta	
-0.41	0.17	-0.39	0.55	0.24	statistic	Recursive Forecast		

جدول 12. نتایج آزمون دیبولد ماریانو برای سری زمانی بازدهی هفتگی (مدل های MIDAS)
ماخذ : نتایج پژوهش

Table 12. Diebold Mariano Test Results for Daily Return Time Series (MIDAS Models)

Source : Research Result

MIDAS				آماره	الگوی پیش بینی	تابع وزن دهی	مدل
beta		almon					
Rolling Forecast	Rolling Forecast	Rolling Forecast	Rolling Forecast				
0.03	0.03	---	---	pvalue	Rolling Forecast	almon	MIDAS
1.79	1.78	---	---	statistic	Recursive Forecast		
0.03	0.03	---	---	pvalue	Rolling Forecast	nealmon	
1.84	1.83	---	---	statistic	Recursive Forecast		
0.02	0.02	0.006	0.00	pvalue	Rolling Forecast	beta	
1.92	1.92	2.50	2.96	statistic	Recursive Forecast		
0.02	0.026	0.009	0.008	pvalue	Rolling Forecast	beta	
1.93	1.92	2.33	2.38	statistic	Recursive Forecast		
---	---	0.96	0.96	pvalue	Rolling Forecast	beta	
---	---	-1.83	-1.78	statistic	Recursive Forecast		
---	---	0.96	0.96	pvalue	Rolling Forecast	beta	
---	---	-1.84	-1.79	statistic	Recursive Forecast		

نتایج آزمون دیبولد ماریانو در

جدول 11 برای سری زمانی بازدهی هفتگی نشان می دهد که مدل $GARCH - ARIMA(1,2)$ با توزیع تی استیوودنت نسبت به توزیع های نرمال و نرمال چوله بهتر عمل می کند. همچنین همین مدل با توزیع خطای تعمیم یافته نسبت به توزیع نرمال چوله عملکرد بهتری دارد. مانند مرحله قبلی تکرار آزمون دیبولد ماریانو برای سری زمانی بازدهی هفتگی نشان می دهد که مدل سازی MIDAS با تابع وزن دهی آلمون و آلمون نمایی و در هر دو حالت پنجره غلتان و بازگشتی، نسبت به کلیه مدل های خانواده $ARIMA - GARCH$ معرفی شده در این پژوهش عملکرد بهتری دارند. تفکیک نتایج و اجرای آزمون دیبولد ماریانو برای مدل MIDAS با انواع تابع وزن دهی، نشان می دهد تابع وزن دهی آلمون نمایی برتری کامل نسبت به دیگر توابع وزن دهی داشته و تابع وزن دهی آلمون نسبت به تابع وزن دهی بتا عملکرد بهتری دارد. نتایج جدول 12 نیز، عملکرد مطلوب مدل سازی MIDAS به همراه تابع وزن دهی آلمون نمایی که در بخش بررسی معیارهای خطای پیش بینی برای پیش بینی بازدهی هفتگی مشاهده شد را تایید می کند.

6- نتیجه‌گیری

در این پژوهش دو رویکرد در مدل‌سازی اطلاعات و داده‌های سری زمانی از نظر دقت در پیش‌بینی مورد بررسی قرار گرفت. رویکرد سنتی اقتصادسنجی در مدل‌سازی سری‌های زمانی که سالیان متمادی در حال استفاده است، در مقابل رویکردی نوین که در سالیان اخیر برای حل یکی از مشکلات مهم در مدل‌سازی اقتصادی، یعنی لزوم تواتر زمانی یکسان در یک مدل رگرسیون، طراحی شده است. در این مطالعه تلاش شد تمامی حالات ممکن در مدل‌سازی برای هر دو رویکرد، شامل انواع مدل‌های موجود، توزیع خطاهای مرسوم و توابع وزنی مورد استفاده در مطالعات گذشته، در کنار همدیگر مورد بررسی و آزمون قرار گیرند. همچنین از داده‌های با تواتر زمانی بالا (میان روزانه 5 دقیقه‌ای) برای محاسبه بازدهی و مدل‌سازی استفاده شد. با توجه به معرفی رویکرد مدل‌سازی برای داده‌های با تواتر زمانی متفاوت به نظر می‌رسد توجه به موارد زیر می‌تواند نتایج به‌دست‌آمده در این مطالعه را شفاف‌تر نماید.

- قدرت مدل‌های خانواده ARIMA-GARCH بر پایه تولید خطای میانگین مربعات کوچک است. این موضوع می‌تواند در زمان مواجهه با نوسانات بالا در یک بازه زمانی کوتاه (داده‌های میان روزانه) یا شوک‌های موقتی، باعث ارائه پیش‌بینی‌های با دقت کمتر شود. مطالعات امره و همکاران (2012) و چلیک و ارگین (2013) نیز نشان می‌دهد مدل‌های MIDAS در پیش‌بینی قیمت و بازدهی دارایی‌هایی که نوسانات بالا در بازه‌های زمانی کوتاه دارند، عملکرد بهتری ارائه می‌کنند (Emre et al, 2012 & Celik & Ergin, 2013).

- در صورت وجود داده‌های با تواتر زمانی بالا برای یک دارایی (مانند رمزارزها و یا داده‌های سهام) و همچنین نوسانات بالا در اینگونه داده‌ها در صورتیکه هدف، پیش‌بینی بازدهی باشد، استفاده از رویکرد مدل‌سازی داده‌های ترکیبی با تواترهای زمانی متفاوت (MIDAS) نسبت به رویکرد مدل‌سازی خودرگرسیون میانگین متحرک و مدل‌سازی واریانس ناهمسان مشروط (ARIMA-GARCH) از حیثیت خواهد داشت.

- پیچیدگی اصلی در مدل‌سازی با رویکرد MIDAS انتخاب تعداد وقفه‌های مناسب برای متغیر با تواتر زمانی بالا و انتخاب تابع وزندهی مناسب است که هر دو تاثیر محسوسی در ارائه پیش‌بینی‌ها با دقت بالاتر خواهند داشت.

- براساس نتایج بدست آمده در این مطالعه، برای بازدهی روزانه، مدل‌سازی MIDAS با استفاده از تابع وزن دهی آلمون نتایج بهتری ارائه می‌کند. در حالیکه برای بازدهی هفتگی، استفاده از تابع وزن دهی آلمون نمایی پیشنهاد می‌شود. لذا به نظر می‌رسد تابع آلمون برای وزن‌دهی سری زمانی روزانه در مقابل داده‌های پرتواتر میان روزانه (5 دقیقه‌ای) عملکرد بهتر و تابع وزن دهی آلمون نمایی برای وزن‌دهی سری زمانی هفتگی در مقابل داده‌های روزانه و میان روزانه بهتر عمل می‌کند.

- نتایج این مطالعه نشان می‌دهد استفاده از داده‌های پرتواتر میان روزانه (5 دقیقه‌ای) و مدل‌سازی MIDAS، دقت پیش‌بینی‌ها را نسبت به رویکردهای مدل‌سازی سنتی اقتصادسنجی، افزایش و میزان خطای پیش‌بینی را به شکل محسوسی کاهش می‌دهد. این موضوع هم برای سری‌های زمانی روزانه و هم هفتگی صدق می‌کند.

- با توجه به وجود داده‌های تواتر بالا (میان روزانه)، در سری‌های زمانی مالی، بازارهای سهام و رمزارزها که منجر به وجود اطلاعات باارزش در بازه‌های زمانی بسیار کوتاه شده است، پیشنهاد اصلی این مطالعه استفاده از رویکرد MIDAS در پیش‌بینی سری‌های زمانی مالی است.

- با توجه به خصوصیت ویژه مدل‌سازی MIDAS که امکان استفاده از چندین متغیر توضیح دهنده با تواترهای زمانی کاملاً متفاوت در سمت راست رگرسیون را فراهم می‌نماید، استفاده از این رویکرد بخصوص برای داده‌های بازارهای مالی در داده‌های معاملات روزانه⁴⁷ می‌تواند منجر به تعیین استراتژی‌های بهینه برای سرمایه‌گذاران بازارهای مالی شود. این خصوصیت برای مدل‌های سنتی اقتصادسنجی متصور نیست.

⁴⁷ tick by tick data

- توسعه مدل‌های MIDAS منجر به طراحی مدل‌های ترکیبی GARCH-MIDAS به منظور استخراج اجزاء کوتاه مدت و بلندمدت متغیر توضیح دهنده با توأثر پایین نسبت به متغیر وابسته گردیده که این مطالعه پیشنهاد می‌کند مدل‌های GARCH-MIDAS نیز بر روی سری‌های زمانی مالی اجرا و نتایج پیش‌بینی مورد بررسی قرار گیرد.
- با توجه به خصوصیات منحصر به فرد سری زمانی قیمت بیت‌کوین مانند نوسانات بسیار بالا پیشنهاد می‌شود مطالعه حاضر روی یک سری زمانی دیگر نیز انجام و کارایی دو رویکرد در مقابل همدیگر مورد سنجش قرار گیرند.

Acknowledgments: Acknowledgments may be made to individuals or institutions that have made an important contribution.

Conflict of Interest: The authors declare no conflict of interest.

Funding: The author(s) received no financial support for the research, authorship, and publication of this article.

Reference

- Abasi Nami, H. (2021). Forecasting Crude Oil prices Volatility and Value at Risk: Single and Switching Regime GARCH Models. *Energy Economics review*, 17(68), 141-174. SID: <https://sid.ir/paper/380078/> [In Persian]
- Andreou, E. Ghysels, E. Kourtellos, A. (2010). Regression Models with Mixed Sampling Frequencies. *Journal of Econometrics*. 158(2), Pages 246-261.
- Barakchian, S. M., & Rezaei, M. (2016). Evaluation of Mixed-Frequency Regressions in Forecasting Seasonal Inflation Rate of Iran. *Journal of Economics and Modelling*, 7(25), 25-46. https://ecof.sbu.ac.ir/article_52259_52c91c7c6bfd6ad4801b43f8e03d4da5.pdf [In Persian]
- Bariviera, A.F. (2017). The In-efficiency of Bitcoin Revisited: A Dynamic Approach. *Economics Letters*, 161, 1-4.
- Bayat, M. Noforesti, M. (2015). *Applied Time Series Econometrics: Mixed Frequency Data Sampling Model*. Noor-e-Elm Pub. [In Persian]
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
- Box, G., Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day.
- Brockwell, P. Davis, R. and Calder, M. (2002). *Introduction to Time Series and Forecasting*. Springer, Vol 2.
- Catania, L. Grassi, S. & Ravazzolo, F. (2019). Forecasting Cryptocurrencies Under Model and Parameter Instability. *International Journal of Forecasting*, 35, 485-501.

- Celik & Ergin. (2013). Volatility Forecasting Using High Frequency Data: Evidence from Stock Markets. *Economic Modeling*, 36, 176-190.
- Chegeni, A. & Gord, A. (2020). Stock Price Prediction in Tehran Stock Exchange Using Artificial Neural Network Model and ARIMA Model: A Case Study of Two Active Pharmaceutical Companies in Stock Exchange. *Financial Engineering and Securities Management*, 11(44), 350-371. SID : <https://sid.ir/paper/396733> [In Persian]
- Clements, M. P. and Galvão, A. B. (2008). Macroeconomic Forecasting with Mixed-Frequency Data: Forecasting Output Growth in the United States. *Journal of Business and Economic Statistics*, 26(4): 546-554.
- Dashtban, M. Elyaspour, B. & Dashtban, S. (2019). The Effect of Exchange Rate Uncertainty on the Flow of Tourism in Iran: A Stochastic Volatility Model with Leverage Effects and Mixed Frequency Data Sampling (MIDAS) Approach. *Journal of Tourism Planning and Development*, 8(29), 173-190. SID : <https://sid.ir/paper/375069/> [In Persian]
- Diebold, F & Mariano, R. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3), 253-63.
- Emre, A. Fendoglu & Saltoglu. (2012). MIDAS Volatility Forecast Performance Under Market Stress: Evidence from Emerging Stock Markets. *Economics Letters*, 117(2), 528-532.
- Enders, W. (2008). *Applied Econometric Time Series*. New York: John Wiley & Sons.
- Feng Ma, M.I.M. Wahab, Chao Liang. (2020). Cryptocurrency Volatility Forecasting: A Markov Regime-Switching MIDAS Approach. *Energy*, 212 (2020), 118743.
- Ghafari, F. & Farhadi, A. (2015). ARIMA-GARCH and State Space Models' Forecasting Power Case Study: Tehran Exchange Stock Index (TEPIX). *Iranian Journal of Applied Economics*, 5(-), 33-42. SID: <https://sid.ir/paper/202056/en> [In Persian]
- Ghysels, E. Santa-Clara, P. Volkanov, R.(2002b). The MIDAS Touch: Mixed Data Sampling Regression. Discussion Paper UCLA and UNC available at: <http://www.unc.edu/eghysels>
- Ghysels, E. Santa-Clara, P. Volkanov, R.(2004).The MIDAS Touch: Mixed Data Sampling Regression Models. CIRANO Working Papers 2004s-20, CIRANO.
- Ghysels, E. Sinko, A. Valkanov, R.(2006). Midas Regressions: Further Results and New Directions. *Econometric Reviews*, 26(1), 53-90.

- Glosten, L. Jagannathan, R. & Runkle, D. (1993). On the Relation Between The Expected Value and The Volatility of Nominal Excess Return on stocks. *Journal of Finance*, 48, 1779-1801.
- Hua, Yiqing. (2020). Bitcoin price prediction using ARIMA and LST, E3S Web of Conferences 218, 01050.
- Klien, L.R. Sojo, E. (1989). Combinations of High and Low Frequency Data in Macroeconomic Model. *Economics in Theory and Practice*. Kluwer Academic Publishers, 3-14.
- Moshiri, S. (2001). Forecasting Iranian Inflation Rates Using Structural, Time Series and Artificial Neural Networks Models. *Journal of Economic Research*, (58), 147-184. SID: <https://sid.ir/paper/121112/> [In Persian]
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *The Econometric Society*, 59(2), 347-370.
- Noferesti, M. and M. Sezavar. (2021). Predicting the Effects of New Sanctions and Evaluating Fiscal Policies in the Context of a Macroeconomic Model with Mixed-Frequency Data Sampling for the Iranian Economy Under Sanctions. *Journal of Economic Modeling Research* 11(44): 85-104.
- Pahlevani, M. Roshan, R (2015). The Comparison Among ARIMA and Hybrid ARIMA-GARCH Models in Forecasting the Exchange Rate of Iran. *International Journal of Business and Development Studies*, 7(1), 31-50.
- Pishbahar, E. BODAGH, S. & Dashti, G. (2019). Forecasting Iran's Agricultural Sector Growth: Using Mixed-frequency Data Sampling (MIDAS) Model. *Journal of Sustainable Growth and Development (the Economic Research)*, 19(3), 145-161. SID: <https://sid.ir/paper/371137/>. [In Persian]
- Salehifar, M. (2019). Risk and Return Behavior of Bitcoin in Comparison with Gold, Currency, and Stock Markets by Application of GJR-GARCH and TGARCH Models. *Financial Engineering and Securities Management*, 10(40), 152-168. SID: <https://sid.ir/paper/197625/> [In Persian]
- Silvestrini, A. & Veredas, D. (2008). Temporal Aggregation of Univariate and Multivariate Time Series Models: A survey. *Economic working papers*, 685, Bank of Italy, Economic Research and International Relations Area.
- Wu, S. I. and Lu, R. P. (1993). Combining Artificial Neural Network and Statistics for Stock Market. *Forecasting*, 257-264.

Choosing the Optimal Model for Predicting Asset Returns: A Comparison of ARIMA-GARCH and MIDAS Models

Kambiz Hojabr Kiani *, Saeed Sadrzadeh Moghadam  **

* Prof. Department of Economic, Faculty of Management and Economics, Islamic Azad University, Science and Research Branch, Tehran, Iran. ORCID:

Email: khkiani@yahoo.com
ORCID: 0000-0002-0191-2138

** Corresponding Author, Ph.D. Candidate, Department of Economic, Faculty of Management and Economics, Islamic Azad University, Science and Research Branch, Tehran, Iran. . TEL: 09132024969

Email: Sadrzadeh.saeed@gmail.com
ORCID : 0000-0001-9930-0385.

This study aims to compare and evaluate two modeling approaches in econometrics in terms of forecasting accuracy, using a time series of Bitcoin prices for the period 2015 to 2023 AD. Bitcoin, as the most popular and well-known cryptocurrency, has become an important asset for many investors in financial markets and even ordinary people in recent years, many studies have been conducted to discover the price structure and changes of this asset over time. The high volatility of this cryptocurrency has created many complexities in these studies. Mixed data sampling (MIDAS) is a new approach for modeling data with different time frequencies in recent years and with regard to the elimination of important limitations in econometric modeling, it is increasingly being studied and investigated. In this study, the forecasting accuracy of ARIMA-GARCH and MIDAS models is compared. The models used in this research are based on intraday data of Bitcoin prices (with high frequency-5 minutes) on the variables of daily and weekly returns of Bitcoin. First, a variant of ARIMA-GARCH model with different types of GARCH models and distributions and then the MIDAS model with different weighting functions are estimated. Finally, by dividing the sampling interval into two parts, both approaches are used to forecast the return of the Bitcoin time series and compared with the actual information. The results show that modeling with the mixed data sampling approach provides better results in terms of forecasting the return of both time series of daily and weekly returns of Bitcoin.

Keywords: MIDAS, ARIMA, Conditional Heteroscedasticity, BitCoin, High-Frequency Data

JEL classification: C22, C89, G17