

پیش بینی شاخص بورس اوراق بهادار تهران بوسیله روشهای هوشمند یادگیری عمیق

آرش خسروی^۱، محمدرضا نوروزی^۲
^۱ دانشکده مهندسی، مرکز آموزش عالی محلات، محلات، ایران

چکیده

شاخص بازار سرمایه به عنوان دماسنج اقتصادی در هر کشور است، در گزارشهای خبری هنگامیکه از بورس سخن به میان می آید، معمولاً این شاخص است که عنوان می شود. از این رو، پیش بینی این متغیر، بینش مهمی از وضعیت اقتصادی و به دست آوردن استراتژی های مناسب برای سرمایه گذاری را ارائه می دهد. پژوهش در پیش بینی پذیری بازار سهام دارای سابقه های طولانی در اقتصادهای مالی است. در حالی که عقاید متفاوتی پیرامون کارایی بازار وجود دارد، مطالعات تجربی گسترده نشان می دهد که بازارهای مالی تا حدی قابل پیش بینی هستند. در میان روشهای پیش بینی بازده سهام، روشهای آماری یا اقتصادسنجی بر پایه تحلیل حرکات گذشته بازار از همه ی روشها پذیرفته شده تر هستند. اخیراً و با رشد روزافزون توان محاسباتی کامپیوترها، استفاده از روش هوش مصنوعی نیز با اقبال زیادی مواجه شده است. هدف پژوهش حاضر این است که با استفاده از مدل های هوش مصنوعی به پیش بینی شاخص کل بورس اوراق بهادار تهران و شاخص ۵۰ شرکت برتر بپردازد. برای این منظور اطلاعات مربوط به این شاخصها در بازه های زمانی یک روزه برای ۱۰ سال استخراج شده است. این اطلاعات با مدل یادگیری عمیق که یکی از مدل های نوین هوش مصنوعی است، تجزیه و تحلیل شده است.

کلمات کلیدی: بورس اوراق بهادار، وضعیت اقتصادی، شاخص بازار سرمایه، پیش بینی شاخص، شبکه عصبی مصنوعی

Stock Price Prediction with Deep Learning

Arash Khosravi *¹, Mohammad Reza Noruzi²

^{1,2} Faculty of Engineering, Mahallat Institute of Higher Education, Mahallat, Iran.

Abstract

The stock market is an economic thermometer in every country, and news reports usually mention the financial index regarding the stock market. Therefore, predicting this variable provides bright insight into the economic situation and obtaining appropriate investment strategies. Research on stock market predictability has a long history in financial economics. While there are different opinions about market efficiency, extensive empirical studies show that financial markets are somewhat predictable. Therefore, among the methods of predicting stock returns, statistical or econometric methods based on analysing past market movements are more accepted than other methods. Artificial intelligence has also been prevalent with the increasing computing power of computers. The current research uses artificial intelligence models to predict the total index of the Tehran Stock Exchange and the index of the top 50 companies. Moreover, for this purpose, the information related to these indices has been extracted in one-day intervals for ten years. We will analyse this information with the deep learning model, one of the new artificial intelligence models.

Keywords: Tehran Stock Exchange, Stock Market Index, Prediction, Artificial neural network

تاریخچه مقاله:

تاریخ ارسال: ۱۴۰۱/۰۲/۰۱

تاریخ اصلاحات: ۱۴۰۱/۰۴/۲۰

تاریخ پذیرش: ۱۴۰۱/۰۵/۲۰

تاریخ انتشار: ۱۴۰۱/۰۶/۳۰

Keywords:

Tehran Stock Exchange Stock
Market Index
Prediction
Artificial neural network

* ایمیل نویسنده مسئول:

Khosravi.280@gmail.com

۱ - مقدمه

پیش‌بینی شاخص جهت ایجاد دید کلی از وضعیت اقتصادی و اخذ استراتژی‌های سرمایه‌گذاری، یکی از مسائل مهم به شمار می‌رود. مطالعات تجربی زیادی نشان می‌دهند که به رغم وجود نظریه بازار کارا، بازارهای مالی تا حدی قابل پیش‌بینی هستند. در میان روش‌های پیش‌بینی بازده، روش‌های اقتصادسنجی مبتنی بر تحلیل تغییرات گذشته سهام بیشترین اقبال را داشته‌اند. این نگرش از روش‌های خطی و غیرخطی متنوعی استفاده می‌کند. روش‌هایی چون خودهمبستگی، خودهمبستگی میانگین متحرک واریانس ناهمسانی شرطی اتورگرسیو، واریانس ناهمسانی شرطی اتورگرسیو تعمیم یافته بیشتر مورد توجه محققین هستند [۱]. اخیراً مدل‌های مبتنی بر شبکه عصبی با توجه به جامعیت این روش و عدم وجود برخی از پیش‌فرضها در خصوص داده‌ها، گسترش زیادی نسبت به روش‌های آماری یافته است. در کشورهایی که بازار سرمایه آنها توسعه یافته است، و تعداد شرکتهای پذیرفته شده در این بازار نسبت به تعداد شرکتهای فعال در کشور زیاد است، بازار سرمایه آنها نماگر کاملی از وضعیت اقتصادی آن کشور می‌باشد، و شاخص بورس اوراق بهادار به عنوان یک نماگر اقتصادی به شمار می‌رود که با پیش‌بینی این نماگر می‌توان نسبت به وضعیت و مدیریت نقدینگی کشور و همچنین اخذ استراتژی‌های سرمایه‌گذاری مناسب اقدام نمود. به عنوان مثال در دوره‌هایی که شاخص برای یک دوره طولانی مدت روند صعودی به خود می‌گیرد، موجب می‌گردد، که نقدینگی از سایر بازارهای موازی از قبیل ارز، سکه و مسکن به این بازار گسیل شود. در این حالت معمولاً سرمایه‌گذاران استراتژی‌های خرید و نگهداری را اخذ می‌نمایند و اقدام به سرمایه‌گذاری با دید بلندمدت می‌نمایند. در نقطه مقابل زمانی که شاخص برای یک دوره طولانی مدت روند نزولی به خود می‌گیرد، معمولاً نقدینگی به سمت بازارهای موازی بازار سرمایه انتقال می‌یابد و سرمایه‌گذاران در بازار سرمایه اقدام به اخذ استراتژی‌های سرمایه‌گذاری کوتاه مدت می‌نمایند.

لذا همواره پیش‌بینی شاخص به عنوان یک نماگر اقتصادی مورد توجه سرمایه‌گذاران و مسئولان اقتصادی یک کشور است [۲-۵]. روش‌های پیش‌بینی در دو دسته کلی روش‌های ساختاری و سری زمانی طبقه‌بندی می‌گردد. در روش‌های ساختاری مقدار یک متغیر (متغیر وابسته) با توجه به سایر متغیرهای مستقل تعیین می‌گردد، در حالی که در مدل‌های سری زمانی، وقفه‌های زمانی متغیر وابسته به عنوان متغیر مستقل محسوب می‌گردند. در روش‌های ساختاری علاوه بر پیش‌بینی متغیر وابسته، نیاز به پیش‌بینی مقدار متغیر یا متغیرهای مستقل نیز وجود دارد، در حالی که در روش سری زمانی، وقفه‌های یک متغیر برای پیش‌بینی مقدار آتی همان متغیر به کار می‌رود. با توجه به سادگی مدل‌های سری زمانی، این روش مخصوصاً در حوزه‌های مالی و اقتصادی، گسترش زیادی یافته است. مدل‌های پیش‌بینی سری

زمانی نیز به دو دسته کلی مدل‌های آماری و مدل‌های هوش مصنوعی تقسیم‌بندی می‌شوند [۳]. مدل‌های آماری چون واریانس ناهمسانی شرطی اتورگرسیو، واریانس ناهمسانی شرطی اتورگرسیو تعمیم یافته که لازمه این مدل‌ها وجود رابطه خطی به استثنای (GARCH)، توزیع نرمال و مانا بودن داده‌هاست، که معمولاً این شرایط در بازارهای مالی مهیا نیست. مدل‌های هوش مصنوعی نیز شامل شبکه عصبی، سیستم فازی، مدل ژنتیک و ... هستند که در بین آنها شبکه عصبی با توجه به ویژگی‌های آن، کاربرد گسترده‌ای یافته است [۴-۷]. وجود برخی از مزیتها سبب برتری مدل‌های هوش مصنوعی نسبت به مدل‌های آماری شده است. مزیتی که شبکه عصبی نسبت به مدل‌های اقتصادسنجی دارد عدم نیاز به مانا بودن سری زمانی و رابطه خطی است. به عبارت دیگر، شبکه عصبی به دلیل جامعیت توان تعیین هرگونه رابطه خطی و غیرخطی را دارا است. در این پژوهش در تلاش هستیم تا با کمک شبکه عصبی LSTM، شاخص بورس و شاخص ۵۰ شرکت برتر را پیش‌بینی کنیم [۸].

۲- پیشینه پژوهش

۲-۱- انواع سرمایه‌گذاری

روش‌های مختلفی برای تحلیل سرمایه‌گذاری وجود دارد که از جمله آنها می‌تواند به روش‌های تحلیل سرمایه‌گذاری تکنیکال، بنیادی، پیش‌بینی سری‌های زمانی و یادگیری ماشینی اشاره کرد [۹]. تحلیل بنیادی با استفاده از داده صورت‌های مالی شرکتها به تجزیه و تحلیل می‌پردازد. تحلیل تکنیکال بر مبنای نظریه گشت تصادفی بیان می‌کند که تاریخ تکرار شونده است و به دنبال پیدا کردن الگوهای موجود است تا بدین وسیله بتواند روند حرکتی قیمت را پیش‌بینی کند. روش پیش‌بینی سری‌های زمانی کلاسیک ۱ به تجزیه و تحلیل داده‌ها پرداخته و عمدتاً به دنبال آن است که به وسیله یک رابطه خطی بین داده‌های تاریخی بتواند ارزش آینده را پیش‌بینی نماید [۱۰-۱۲]. روش یادگیری ماشینی، رویکردی نوین است و این روش در مواقعی که تعداد متغیرهای مستقل و وابسته زیادند و روابط خطی بین آنها وجود ندارد، بهترین روش است. پژوهش‌ها نشان می‌دهد از آنجا که قیمت سهام دارای رفتار غیر خطی است، این روش موفق به کاهش خطا در پیش‌بینی شده است.

بسیاری از این مدل‌ها بر مبنای کارایی بازار مطرح شده‌اند، حال آنکه در صحت این موضوع ابهاماتی وجود دارد؛ بنابراین در این تحقیق به بررسی کارایی بازار می‌پردازیم. فرضیه قدم زدن تصادفی که به آن گشت تصادفی نیز گفته می‌شود، یک تئوری مالی است که بیان می‌کند که روند تغییرات قیمت‌های بازار سهام به صورت گشت‌های تصادفی بوده و بنابراین قابل پیش‌بینی نیستند. این فرضیه با فرضیه بازار کارا سازگار است. ردپای استفاده از این مفهوم برای اولین بار می‌توان در

¹ Traditional Time Series

اوراق بهادار می‌تواند نقدینگی مورد نیاز خود را به دست آورده و تکاپوی سرمایه خود را بهبود بخشد. وظیفه اصلی بورس اوراق بهادار، فراهم آوردن بازاری شفاف و منصفانه برای دادوستد اوراق بهادار پذیرفته شده و همچنین ارائه سیستمی مناسب برای نظارت بر جریان دادوستد، عملیات بازار و فعالیت اعضای آن است. بورس اوراق بهادار به عنوان یک بازار منسجم و سازمان یافته، مهمترین متولی جذاب و سامان‌دادن صحیح منابع مالی سرگردان است و با جمع‌آوری نقدینگی جامعه و فروش سهام شرکتها، ضمن به حرکت درآوردن چرخ های اقتصاد جامعه از طریق تامین سرمایه‌های مورد نیاز پروژه‌ها، کاهش دخالت دولت در اقتصاد و نیز افزایش درآمدهای مالیاتی، منافع اقتصادی چشمگیری به ارمغان آورد و در کنار اینها، اثرات تورمی ناشی از وجود نقدینگی فراوان در جامعه را نیز از بین می‌برد [۲۲].

۲-۳- کارهای پیشین

گرچه هنوز بیش از ۶۰ سال از آغاز روش‌های محاسبات مبتنی بر شبکه‌های عصبی مصنوعی نمی‌گذرد، با این حال این شبکه‌ها به دلیل ویژگی‌هایی مانند پردازش موازی، هوشمندی و انعطاف‌پذیری جایگاه چشمگیری در مسائل پیچیده از قبیل شناخت الگو، خوشه‌بندی، مدلسازی، تخمین و پیش‌بینی برای خود باز نموده اند. ایده شبکه عصبی در سال ۱۹۴۳ توسط مک کلوچ و پیتز ارایه شد. این دو نفر مدل ساده نرون را ایجاد کردند که این نرون قابلیت ساخت دروازه‌های AND و OR را داشت. سپس این مدل توسط دانشمندان دیگر توسعه پیدا کرد و از حالت تک‌نرون به شبکه‌ای از نرون‌ها توسعه یافت. برای استفاده از یک شبکه عصبی باید ابتدا شبکه را تعمیم داد. در سالهای گذشته ساختارهای متنوعی برای شبکه‌های عصبی ارایه شده است [۲۳].

حبیب زاده و ایزدپور (۱۳۹۹) برای تعیین مهمترین نسبت‌ها و شاخص‌های مالی موثر بر پیش‌بینی سودآوری شرکتها، ۱۲۴ شرکت در بورس اوراق بهادار تهران را در بازه زمانی سالهای ۱۳۸۷ تا ۱۳۹۵ با مبنای قرار دادن اطلاعات بنیادی صورت‌های مالی منتشر شده در کدال ۲ مورد بررسی قرار دادند. آنها برای تعیین متغیرهای معنادار در پیش‌بینی سودآوری از روش درخت تصمیم C5 استفاده نمودند. سپس با انتخاب ۸ متغیر موثر که عبارتند از سود خالص به کل دارایی، فروش، سود عملیاتی به فروش، سود خالص به حقوق صاحبان سهام، سود و زیان انباشته به حقوق صاحبان سهام، سود قبل از بهره و مالیات به کل دارایی، کل بدهی‌ها به کل دارایی‌ها و دارایی جاری دقت پیش‌بینی سه روش شبکه عصبی، درخت تصمیم C5 و ماشین بردار پشتیبان را اندازه‌گیری و مقایسه کردند. نتایج بدست آمده نشان داد که درخت تصمیم C5 دارای بالاترین دقت و کمترین خطا در پیش‌بینی است [۲۴].

آثار معامله‌گر فرانسوی «جولز رنالت» که در سال ۱۸۶۳ کتابی نوشته و بعد از او ریاضیدان فرانسوی «لوئیز بچلیر» یافت که در رساله دکتری خود با عنوان «تئوری سفته بازی در سال ۱۹۰۰ تفسیر و بینش قابل توجهی از این فرضیه را نشان داده است. ایده مشابهی بعدها بوسیله «پائول گوتتر» استاد دانشکده مدیریت اسلوان ام‌آی‌تی در کتابی با عنوان «اسپرت تصادفی قیمت‌های بازار سهام» که در سال ۱۹۶۴ به چاپ رسید، شکل گرفت. این عبارت در کتاب «گام تصادفی در وال‌استریت» نوشته «بورتون مالکیل» استاد اقتصاد دانشگاه پرینستون در سال ۱۹۷۳ مورد پسند عامه قرار گرفت و فراگیر شد [۱۳-۱۷]. در حالی که پیشتر در مقاله سال ۱۹۶۵، ایکن فاما با عنوان «گام تصادفی در قیمت‌های بازار سهام» که نسخه‌ای با جزئیات فنی کمتر از رساله دکتری او بود، مورد استفاده قرار گرفته بود. این تئوری که قیمت‌های سهام به صورت تصادفی تغییر می‌کنند پیشتر در مقاله سال ۱۹۵۳ «موريس کندال» با عنوان «تحليل اقتصادي سري‌های زمانی قسمت اول: قیمت‌ها» پیشنهاد داده شده بود [۱۸].

اقتصاددانان و افرادی که به نظریه گشت تصادفی اعتقاد دارند، این مسئله را برای بازار سهام بکار می‌برند. عدم وجود همبستگی بین قیمت قبلی و قیمت کنونی به راحتی قابل تشخیص است. اگر روزی قیمت سهامی بالا برود، هیچ فعال بازار سرمایه نمی‌تواند دقیقاً پیش‌بینی کند که قیمت سهم روز بعد نیز بالا خواهد رفت. سهامی که به نظر می‌رسد در روند صعودی قرار دارد، ممکن است در هر لحظه سقوط کند، که این نشان‌دهنده تصادفی بودن تغییر قیمت سهام است [۱۹-۲۰].

۲-۲- آشنایی با بازار بورس و اوراق بهادار تهران

بورس اوراق بهادار تهران در بهمن ماه سال ۱۳۴۶ بر پایه قانون مصوب اردیبهشت‌ماه ۱۳۴۵ تأسیس شد. بورس اوراق بهادار به معنی یک بازار مشکل و رسمی سرمایه است که در آن خرید و فروش سهام شرکتها با اوراق قرضه دولتی با مؤسسات معتبر خصوصی، تحت ضوابط و قوانین و مقررات خاصی انجام می‌شود. مشخصه مهم بورس اوراق بهادار، حمایت قانون از صاحبان پس‌اندازها و سرمایه‌های راکد و الزامات قانونی برای متقاضیان سرمایه است [۲۱]. بورس اوراق بهادار که به عنوان نبض اقتصاد کشور مورد توجه تحلیل‌گران اقتصادی است، از سویی مرکز جمع‌آوری پس‌اندازها و از سوی دیگر، مرجع رسمی و مطمئنی است که دارندگان پس‌اندازهای راکد، می‌توانند محل نسبتاً مناسب و ایمن سرمایه‌گذاری را جستجو کرده و وجوه مازاد خود را برای سرمایه‌گذاری در شرکتها به کار انداخته و یا با خرید اوراق قرضه دولتها و شرکتها معتبر، از سود معین و تضمین شده‌ای برخوردار شوند. بازار اوراق بهادار در واقع مکمل بخش بانکی در تأمین نیازهای مالی بنگاه‌های خصوصی و دولتی است. وجود یک بازار اوراق بهادار که در چارچوب مکانیزم بازار عمل کند، می‌تواند در تخصیص منابع مالی، بهینه و کارا تر عمل نماید. سیستم بانکی نیز با استفاده از مکانیزم

۳- روش پیشنهادی

۳-۱- جامعه و نمونه آماری

در این تحقیق شاخص کل بازار و شاخص پانچا شرکت برتر به عنوان جامعه آماری در نظر گرفته شده‌اند. نمونه آماری این تحقیق داده‌های این شاخص‌ها از ابتدای فروردین ۱۳۸۸ تا انتهای دی ۱۳۹۶ انتخاب شده است؛ در ضمن علت انتخاب این جامعه آماری در دسترس بودن اعداد مربوط به شاخص در بازه‌های روزانه و عدم نیاز به تعدیل قیمتی بوده است.

همچنین از ارزش معاملات (روزانه) به عنوان متغیر کمکی برای مدل استفاده می‌کنیم.

مطابق با پژوهش بک (۲۰۱۸)، به منظور اطمینان از عدم وجود داده‌های پرت، داده‌ها را توسط رابطه زیر نرمال می‌کنیم و سپس از مقادیر رابطه زیر در مدل خود استفاده می‌کنیم:

$$\bar{X} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

بطوریکه X داده اصلی در نقطه مورد نظر، X_{max} و X_{min} به ترتیب حداکثر و حداقل مقدار ممکن یک متغیر خاص در همان نقطه مورد نظر می‌باشند [۲۱].

۳-۲- آمار توصیفی

در جدول شماره ۱ مقادیر آمار توصیفی شاخص‌ها به همراه ارزش معاملات و بازده آنها آورده شده است. اطلاعات این جدول مربوط به داده‌های شاخص بورس تهران از ابتدا فروردین ۱۳۸۸ می‌باشد. بیشترین بازدهی شاخص در یک روز ۵/۴ درصد و کمترین آن منفی ۵/۵ درصد است. میانگین بازده برای هر روز نیز ۰/۴ درصد برآورد شده است. در این جدول آزمون توزیع نرمال متغیرها نیز آورده شده است. همانگونه که ملاحظه می‌شود هیچیک از متغیرها دارای توزیع نرمال نمی‌باشند.

همانطور که در جدول شماره ۱ مشاهده می‌شود داده‌های بازدهی این سه شاخص دارای ضریب کشیدگی مثبت می‌باشد و همچنین دارای چولگی به راست می‌باشند. همچنین میانگین و انحراف بدست آمده نشان دهنده وجود نوسان در طی دوره بررسی است. در نهایت با توجه به خروجی آزمون کولموگروف-اسمیرنوف هم می‌توان ادعا کرد که شاخص‌ها و بازده آنها نرمال نمی‌باشند.

زارعی و همکاران (۱۳۹۷) به بررسی و مقایسه روشهای شبکه عصبی فازی با شبکه عصبی موجک فازی برای پیش‌بینی قیمت سهام بانکهای پذیرفته شده در بورس اوراق بهادار تهران پرداختند. داده‌های این پژوهش از سال ۱۳۹۰ تا ۱۳۹۵ است و در این پژوهش، از سیستم منطق فازی به همراه سیستم شبکه عصبی چندلایه با ساختار بهینه‌سازی پس‌انتشار خطا و ماکزیم همپوشانی تبدیل موجک گسسته برای متغیرهای نرخ ارز، نفت اوپک، طلا، شاخص کل سهام و همچنین ارزش معاملات برای پیش‌بینی قیمت سهام استفاده شده است. نتایج این پژوهش نشان می‌دهد قابلیت اطمینان پیش‌بینی با شبکه عصبی موجک فازی بیش از ۹۰ درصد و شبکه عصبی فازی بیش از ۸۰ درصد است، در نتیجه شبکه عصبی موجک فازی از درجه اطمینان بالاتری نسبت به شبکه عصبی فازی برخوردار است [۲۵].

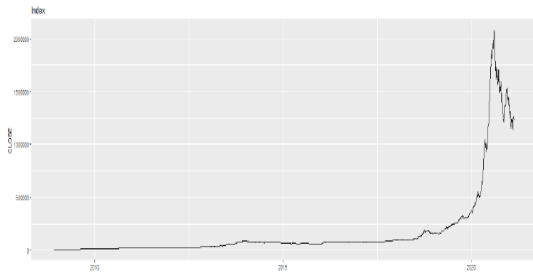
خنجر پناه و همکاران (۱۳۹۷) بیان کردند که تحلیل تکنیکال توانایی پیش‌بینی در کوتاه مدت را دارد. به همین دلیل مدلهایی با استفاده از ابزارهای شبکه عصبی، و مقدار حدی را برای پیش‌بینی روند حرکتی قیمت سهم در روز آینده ارائه کردند. داده‌های این تحقیق شامل برخی شاخص‌های تکنیکال سهم ایران خودرو در بورس اوراق بهادار تهران از سال ۹۲ تا ۹۷ است. نتایج حاکی از آن است که از لحاظ آماری در آزمون ناپارامتری برابری نسبت‌ها، مدل‌های ارائه شده تفاوت معناداری با هم نداشته‌اند، اما معیارهای سنجش خطا نشان می‌دهند که مدل شبکه عصبی، خطای کمتری در پیش‌بینی سهام در بازار بورس تهران دارد [۲۳].

فلاح‌پور و علیپور (۱۳۹۳) به پیش‌بینی شاخص بازار بورس به وسیله شبکه عصبی موجکی پرداختند. دلیل استفاده از این نوع شبکه به این علت است که قیمت سهام تحت تاثیر عوامل اقتصادی و سیاسی مختلفی می‌باشد به همین سبب پیش‌بینی شاخص سهام با استفاده از این داده‌ها همراه با نویز است و باعث می‌شود پیش‌بینی با خطاهای بزرگی همراه باشد. در واقع موجک سبب نویز زدایی می‌شود زیرا موجک‌ها تابع ریاضی هستند که داده‌ها را به فرکانس‌های مختلفی تبدیل می‌کنند. در نتیجه استفاده از مدل سبب انتشار موجک گسسته می‌شود و خطای پیش‌بینی را حداقل می‌کند و به مقدار واقعی نزدیکتر می‌شود [۲۶].

زارع و کردلوئی (۱۳۸۸) به پیش‌بینی قیمت سهام روز بعد در بورس اوراق بهادار تهران به وسیله شبکه پرسپترون چند لایه پرداختند. در این تحقیق هدف اصلی کم کردن خطای پیش‌بینی با روشهای مختلف بود در نتیجه بر صحت و دقت بیشتر مدل دلالت دارد. در این پژوهش از قیمت سهام شرکت سرمایه‌گذاری غدیر استفاده شده است؛ زیرا این شرکت سبب سرمایه‌گذاری است که عملکرد شرکتهای زیر مجموعه آن بر روی شرکت غدیر تاثیر می‌گذارد [۲۷].

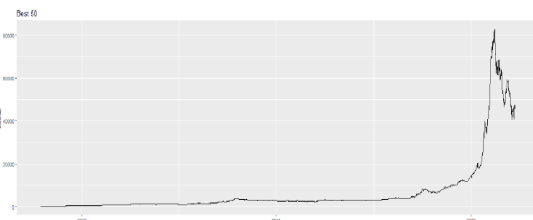
^۳ کمترین بازدهی شاخص مربوط به بیست و ششم اسفند ۱۳۹۳ می‌باشد که شاخص از ۶۵۱۲۲ واحد تا ۶۱۵۳۲۱ واحد کاهش یافته است.

یک میلیون و دویست هزار واحد افت داشته و در این بازه به نوسان می‌پردازد.



(شکل-۲): نمودار سری زمانی شاخص کل هم وزن.

در شکل ۲ نیز همانند شکل ۱ مشاهده می‌شود که نمودار سری زمانی تا اواخر سال ۲۰۱۹ (معادل بهمن ماه ۱۳۹۸)، بدون نوسان شدیدی در پایین عدد هشتاد هزار قرار دارد. اما از ابتدای سال ۲۰۲۰، این شاخص شروع به رشد با شیب ملایم‌تری نسبت به شاخص کل کرده و نهایتاً طی مدتی کوتاه تا پانصد و سی هزار واحد رشد می‌کند. پس از این رشد، این شاخص مجدداً تا سیصد و شصت هزار واحد افت داشته و در این بازه به نوسان می‌پردازد.



(شکل-۳): نمودار سری زمانی شاخص ۵۰ شرکت برتر.

شکل ۳ نشان دهنده سری زمانی شاخص ۵۰ شرکت برتر می‌باشد. این شاخص نیز مانند دو شاخص قبلی ابتدا یک رشد با شیب بسیار زیاد تا سطح هشتاد هزار واحد داشته و مجدداً تا چهل هزار واحد نزول کرده است.

۳-۴- نمودار توابع خودهمبستگی و خودهمبستگی جزئی

در این قسمت نمودار توابع خودهمبستگی و خودهمبستگی جزئی برای بازده سری زمانی‌های شاخص کل، شاخص کل هم‌وزن و شاخص ۵۰ شرکت بزرگتر آورده شده است. همانگونه که ملاحظه می‌شود تابع خود همبستگی بین وقفه‌های (وقفه‌های ۸ تا ۱۱ ام برای این شاخص‌ها، بالاتر از خط فاصله اطمینان قرار می‌گیرد و لذا خودهمبستگی بین بازدهی در اینجا کاملاً مشهود می‌شود (نمودار ACF). در ادامه برای تعیین متغیرهای مستقل در مدل‌های LSTM و RNN نیز از همین وقفه‌ها استفاده خواهیم کرد.

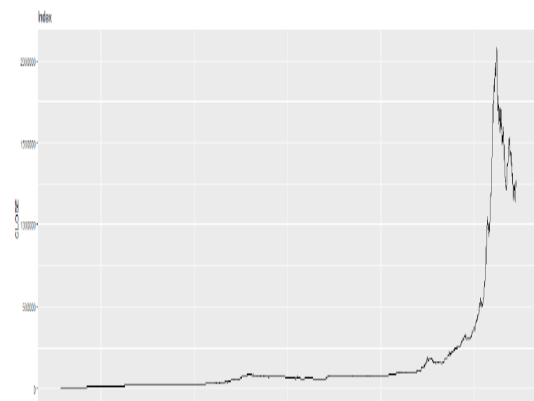
همچنین با توجه به نمودارهای خود همبستگی جزئی که در ردیف دوم نمودار ۴ قرار دارند، متوجه می‌شویم که میانگین متحرک در

(جدول-۱): آمار توصیفی شاخص‌ها.

شرح	میانگین	انحراف	کشیدگی	چولگی
شاخص کل	۱۷۶۳۹۷	۳۶۰۳۱۲	۳/۳۱	۱۳/۳
ارزش معاملات	۱۱۲۴۷/۲	۲۹۹۳۶/۶۱	۳/۹۵	۱۷/۴۹
بازده	۰/۰۰۴	۰/۰۱	۰/۳۲	۶/۸۶
۵۰ شرکت	۶۹۷۴	۱۳۹۴۳	۳/۳۹	۱۳/۹۶
ارزش معاملات	۴۲۸/۰۲۹	۵۷۸/۱۹	۳/۳۸	۱۴/۷۸
بازده	۰	۰/۰۱	۰/۲۵	۶/۴۳
شاخص هم وزن	۸۷۹۳۸	۱۴۱۴۳۸	۱/۹۱	۵/۱۱
ارزش معاملات	۲۰۹۹۳/۷۳	۳۸۳۸۹/۹۱	۲/۶۶	۷/۳۸
بازده	۰	۰/۰۱	۰/۴۲	۴/۷۴
معنی داری	۰	۰	۰	۰

۳-۳- نمودار سری زمانی شاخص‌ها

یک سری زمانی مجموعه مشاهدات تصادفی‌ای است که بر اساس زمان مرتب شده باشند. به این ترتیب نمودار سری زمانی مربوط به شاخص کل، شاخص کل هم وزن و شاخص ۵۰ شرکت برتر که نشان دهنده مقدار این شاخص‌ها در طول زمان می‌باشد. شکل ۱ نمودار سری زمانی شاخص کل را نشان می‌دهد.



(شکل-۱): نمودار سری زمانی شاخص کل.

همانطور که در شکل ۱ مشاهده می‌کنید، نمودار سری زمانی فوق تا اواخر سال ۲۰۱۹ (معادل بهمن ماه ۱۳۹۸)، بدون نوسان شدیدی در پایین عدد پانصد هزار قرار دارد. اما از ابتدای سال ۲۰۲۰، این شاخص شروع به رشد با شیب بسیار زیادی کرده و نهایتاً طی مدتی کوتاه تا دو میلیون واحد رشد می‌کند، پس از این رشد، این شاخص مجدداً تا

پیش‌بینی بازدهی این شاخص‌ها نیز تاثیرگذار است؛ چرا که در نمودار PACF نیز وقفه‌ها خطوط فاصله اطمینان را رد کرده‌اند.

که در رابطه فوق h تابع فعالسازی، u وزن ورودی‌ها، x تابع ورودی و W هم وزن خروجی مرحله قبل می‌باشد.

۳-۶- بررسی وجود حافظه بلندمدت

قبل از تخمین داده‌ها، ابتدا لازم است اثر وجود حافظه بلندمدت را بر بازده در جزء میانگین شرطی داده‌ها بررسی کنیم. برای بررسی وجود حافظه بلندمدت از دو آزمون R/S و GPH استفاده شده است که نتایج حاصل از آن را در جدول زیر نشان داده‌ایم.

(جدول-۲): نتایج آزمون‌ها

آزمون	آماره آزمون
R/S	۱/۰۸
GPH	۰/۳۹۱

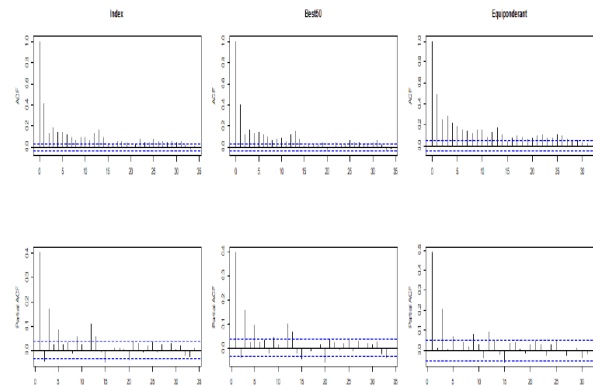
در آزمون‌های فوق فرض صفر، عدم وجود حافظه بلند مدت در بازار بورس اوراق بهادار تهران می‌باشد. با توجه به نتایج آزمون‌های فوق که در جدول ۲ مشاهده می‌شود، در سطح اطمینان ۹۵٪ فرض صفر را رد می‌کنیم و در نتیجه بازار بورس اوراق بهادار تهران را دارای حافظه بلند مدت می‌دانیم.

۳-۷- بررسی شاخص کل

در این بخش با استفاده از ۲۵۰۰ داده‌ی ابتدایی، مدل‌ها را برازش داده و توسط سه مدل RNN ، $ARIMA$ و $LSTM$ و CNN که با خط زرد نشان داده شده‌اند، روز بعد را پیش‌بینی می‌کنیم. روش کار بدین صورت است که هر روز یک داده به مدل اضافه می‌شود و مدل دوباره برازش داده می‌شود و یک پیش‌بینی برای روز بعد ارائه می‌دهد و اینکار همواره ادامه پیدا می‌کند تا اینکه همه داده‌ها پوشش داده شوند.

برازش مدل $ARIMA$ بر روی داده‌های شاخص کل

در نمودار ۵ پیش‌بینی سری شاخص با استفاده از مدل $ARIMA$ آورده شده است. شیوه‌ی کار بدین صورت است که برای هر محاسبه از میان مدل‌های برازش شده، مدلی که بهترین AIC را دارد، انتخاب می‌شود و بر اساس آن پیش‌بینی انجام می‌گردد. در حقیقت تعداد ضرایب موجود در مدل با استفاده از این معیار انتخاب شده و برآورد می‌شوند.



(شکل-۴): نمودار توابع خود همبستگی و خود همبستگی جزئی.

۳-۵- بررسی روشهای CNN ، $LSTM$ ، RNN ، $ARIMA$

حرکت سری‌های زمانی مالی ناشی از حرکت متغیرهای توضیحی مربوطه، از جمله سری زمانی خود متغیر پاسخ است. در بیشتر مطالعاتی که روی پیش‌بینی سری زمانی مالی انجام شده است سری زمانی متغیرهای هدف با وقفه‌های قبلی به عنوان یک متغیر توضیحی ضروری تلقی می‌شود و برای مدلسازی از آنها استفاده می‌گردد. در سیستم‌های مختلف پیش‌بینی مبتنی بر شبکه عصبی، این شبکه با شاخص‌های فنی آموزش داده می‌شوند که براساس مقادیر قبلی متغیرهای هدف محاسبه می‌گردند.

برای پیش‌بینی شاخص بورس سهام، چون سری زمانی شاخص هدف منبع مستقیم و مهمی از حرکت آن است، می‌توان آن را همانطور که در شکل ۲ نشان داده شده است، مدلسازی کرد.

با این حال، تعداد نقاط داده برای آموزش کمتر از ۲۰۰۰ مشاهده است، که کافی نیست. آموزش RNN ها با داده‌های محدود ممکن است دقت پیش‌بینی را کمتر کند و در نتیجه باعث پیش برآورد شود. در ادامه به پیش‌بینی سری زمانی‌های شاخص کل و ۵۰ شرکت برتر می‌پردازیم. همچنین با استفاده از مدل $ARIMA$ نیز به پیش‌بینی این شاخص‌ها می‌پردازیم و سپس نتایج این پیش‌بینی را با نتایج پیش‌بینی توسط شبکه‌های بازگشتی $LSTM$ و $ARIMA$ مقایسه می‌کنیم.

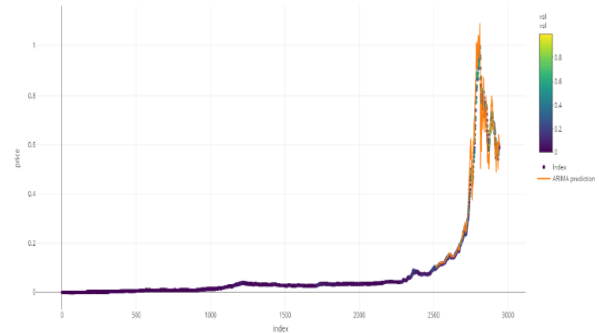
در این مدل، متغیر بازده شاخص به عنوان متغیر وابسته و وقفه‌های ارزش و مجدداً خود شاخص، به عنوان متغیرهای کنترلی استفاده می‌شوند. در این قسمت نیز ترتیب استفاده از داده‌ها بدین شکل است با این تفاوت که مدل مورد استفاده در این بخش شبکه عصبی بازگشتی از مجموعه شبکه‌های یادگیری عمیق می‌باشد. معادله ریاضی مدل RNN بصورت زیر نوشته می‌شود:

$$h_t = u_t x(t) + wh(t-1)$$

می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل اتکایی را برای شاخص ارائه نمی‌دهد.

(جدول-۴): نتایج میانگین خطای روش

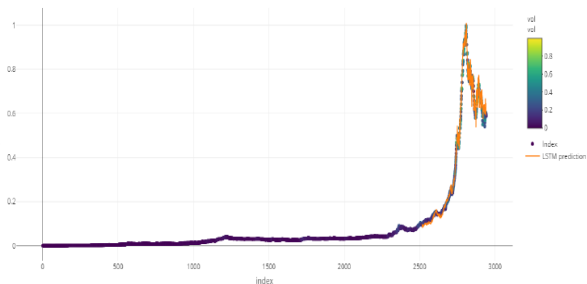
شاخص	MAPE	RMSE
مقدار	۰/۰۶۳۷	۰/۰۵۳۹



(شکل-۵): نمودار پیش‌بینی توسط مدل ARIMA

آموزش شبکه عصبی GARCH-LSTM

در این روش ابتدا وقفه‌های اول تا هشتم شاخص کل و ارزش معاملات را محاسبه کرده و بعنوان متغیر مستقل به شبکه عصبی معرفی می‌کنیم و به آموزش مدل می‌پردازیم. یعنی سری نوساناتی که از مدل GARCH گرفته می‌شود نیز به عنوان متغیر کمکی به مدل اضافه می‌کنیم. روش کار بدین صورت است که هر روز یک داده به مدل اضافه می‌شود و شبکه عصبی دوباره آموزش داده می‌شود و سپس با تلفیق دو مدل GARCH و LSTM پیش‌بینی برای روز بعد ارائه می‌دهد و این کار آنقدر ادامه پیدا می‌کند تا اینکه همه داده‌ها پوشش داده شوند.



(شکل-۷): نمودار پیش‌بینی توسط مدل LSTM و GARCH.

در ادامه مجدداً مقادیر دو معیار MAPE و RMSE محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۵۳۳ و RMSE برابر با ۰/۰۴۳۸ می‌باشد. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۵ درصد می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل اتکایی را برای شاخص ارائه نمی‌دهد.

(جدول-۵): نتایج میانگین خطای روش

شاخص	MAPE	RMSE
مقدار	۰/۰۵۳۳	۰/۰۴۳۸

آموزش شبکه عصبی پیچشی

در گام اول شبکه عصبی پیچشی به گونه ای طراحی می‌شود که از یک مدل پیچش تک بعدی استفاده خواهد شد. برای این منظور در

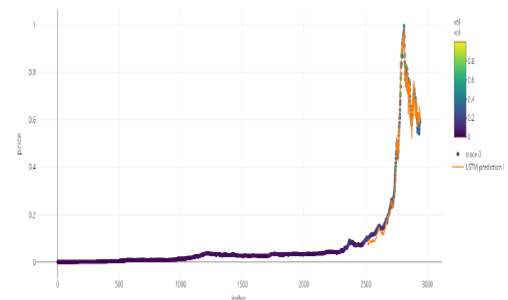
در ادامه مقادیر دو معیار MAPE (میانگین درصد قدر مطلق خطا) و RMSE (جذر میانگین مربعات خطا) محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۵۸۱ و RMSE برابر با ۰/۰۵۲۳ شده است. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۵ درصد می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل اتکایی را برای شاخص ارائه نمی‌دهد.

(جدول-۳): نتایج میانگین خطای روش

شاخص	MAPE	RMSE
مقدار	۰/۰۵۸۱	۰/۰۵۲۳

شبکه عصبی بازگشتی LSTM

در این روش ابتدا وقفه‌های اول تا هشتم شاخص کل و ارزش معاملات را محاسبه کرده و بعنوان متغیر مستقل به شبکه عصبی معرفی می‌کنیم و به آموزش مدل می‌پردازیم. روش کار بدین صورت است که هر روز یک داده به مدل اضافه می‌شود و شبکه عصبی دوباره آموزش داده می‌شود و یک پیش‌بینی برای روز بعد ارائه می‌دهد و اینکار آنقدر ادامه پیدا می‌کند تا اینکه همه داده‌ها پوشش داده شوند.



(شکل-۶): نمودار پیش‌بینی توسط مدل LSTM.

در ادامه مجدداً مقادیر دو معیار MAPE و RMSE محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۶۳۷ و RMSE برابر با ۰/۰۵۳۹ می‌باشد. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۸ درصد

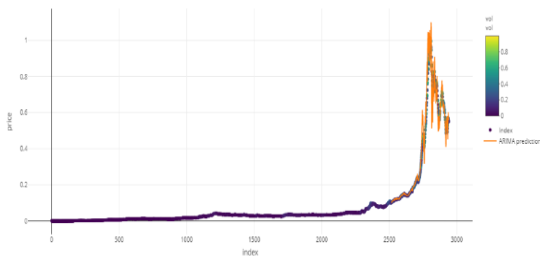
عصبی LSTM و GARCH تلفیقی از همه کمتر شده است که نشان می‌دهد این روش هرچند کم ولی دقت بیشتری برای پیش‌بینی شاخص کل دارد. بعد از آن نیز مدل ARIMA انتخاب شده با استفاده از معیار AIC بیشترین دقت را داشته است. در نهایت هم مدل LSTM و CNN با بیشترین خطا نسبت به دو روش دیگر شاخص کل را پیش‌بینی می‌کند.

(جدول-۷): نتایج میانگین خطای روش‌ها

شاخص	MAPE	RMSE
ARIMA	۰/۰۵۸۱	۰/۰۵۲۳
LSTM	۰/۰۶۳۷	۰/۰۵۳۹
LSTM-GARCH	۰/۰۵۳۳	۰/۰۴۳۸
CNN	۰/۰۷۰۲	۰/۰۵۸۶

۳-۸- بررسی شاخص ۵۰ شرکت برتر

برازش مدل ARIMA بر روی داده‌های شاخص ۵۰ شرکت برتر در نمودار ۹ پیش‌بینی سری شاخص با استفاده از مدل ARIMA آورده شده است. شیوه‌ی کار بدین صورت است که برای هر محاسبه از میان مدل‌های برازش شده، مدلی که بهترین AIC را دارد، انتخاب می‌شود و بر اساس آن پیش‌بینی انجام می‌گردد. در حقیقت تعداد ضرایب موجود در مدل با استفاده از این معیار انتخاب شده و برآورد می‌شوند.

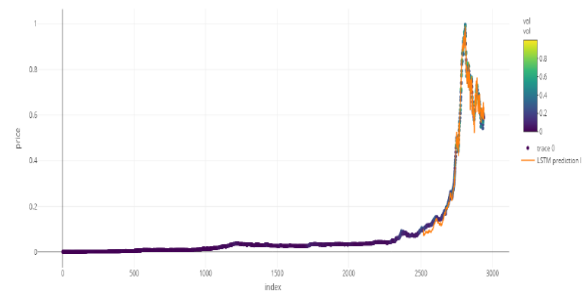


(شکل-۹): پیش‌بینی سری شاخص با استفاده از مدل ARIMA

در ادامه مقادیر دو معیار MAPE (میانگین درصد قدر مطلق خطا) و RMSE (جذر میانگین مربعات خطا) محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۵۳۷ و RMSE برابر با ۰/۰۵۱۷ شده است. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۵ درصد می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل اتکایی را برای شاخص ارائه نمی‌دهد.

شبکه از یک پیچش با تعداد زیادی فیلتر، تابع فعال ساز Relu، بهینه ساز و لایه‌هایی با اتصالات کامل استفاده خواهد شد. گام‌های زمانی در یادگیری، با توجه به پارامترها برابر ۱۵ قرار داده می‌شوند و ۴ روز آینده را به عنوان خروجی مورد نظر، شبکه پیش‌بینی خواهد کرد. نحوه‌ی کار به این صورت است که داده‌های پنجره زمانی اول را شبکه به عنوان ورودی دریافت می‌کند و ۴ داده بعدی را تخمین زده و سپس همین عملیات را از داده دوم شروع کرده و ادامه می‌دهد.

برای طراحی این شبکه ابتدا یک شبکه عصبی پیچشی با تابع فعال ساز Relu استفاده می‌شود و سپس خروجی‌ها به عنوان ورودی به یک حافظه کوتاه مدت طولانی داده می‌شوند. در این مدل نیز از لایه‌هایی با اتصالات کامل استفاده شده است. گام‌های زمانی در یادگیری با توجه به پارامترها برابر ۱۵ قرار داده می‌شوند و ۴ روز آینده را به عنوان خروجی مورد نظر، شبکه پیش‌بینی خواهد کرد. نحوه‌ی کار به این صورت است که داده‌های پنجره زمانی اول را شبکه به عنوان ورودی دریافت می‌کند و ۴ داده بعدی را تخمین زده و سپس همین را از داده دوم شروع کرده و ادامه می‌دهد.



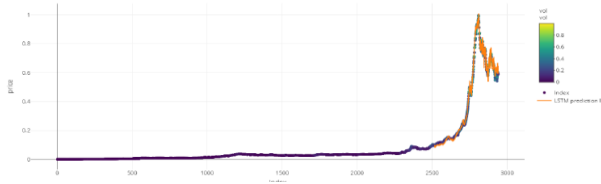
(شکل-۸): نمودار پیش‌بینی توسط مدل عصبی پیچشی با حافظه کوتاه مدت طولانی.

در ادامه مجدداً مقادیر دو معیار MAPE و RMSE محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۷۰۲ و RMSE برابر با ۰/۰۵۸۶ می‌باشد. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۷ درصد می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل اتکایی را برای شاخص ارائه نمی‌دهد.

(جدول-۶): نتایج میانگین خطای روش

شاخص	MAPE	RMSE
مقدار	۰/۰۷۰۲	۰/۰۵۸۶

۵-۱۱- مقایسه چهار روش برای پیش‌بینی شاخص کل در سه بخش قبل مشاهده کردیم که هیچ‌یک از روشهای مورد استفاده در پژوهش دارای دقت خوبی برای پیش‌بینی یک روز آینده شاخص کل را ندارند. اما در این قسمت سعی داریم یک مقایسه با استفاده از دو معیار RMSE و MAPE انجام بدهیم. این دو شاخص برای شبکه



(شکل-۱۱): نمودار پیش‌بینی توسط مدل LSTM و GARCH

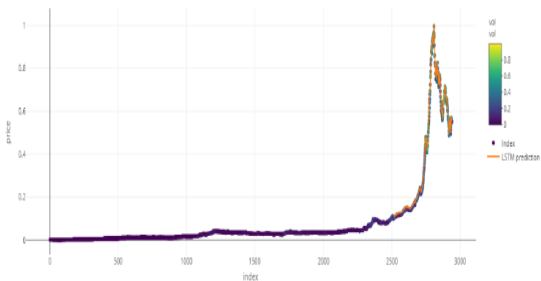
در ادامه مجدداً مقادیر دو معیار MAPE و RMSE محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۳۱۶ و RMSE برابر ۰/۰۲۰۷ می‌باشد. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۳ درصد می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل اتکایی را برای شاخص ارایه می‌دهد.

(جدول-۱۰): نتایج میانگین خطای روش

شاخص	MAPE	RSME
مقدار	۰/۰۳۱۶	۰/۰۲۰۷

آموزش حافظه کوتاه مدت طولانی پیچشی

در آخرین مدل سازی انجام شده، برای طراحی شبکه از یک حافظه کوتاه مدت طولانی حاوی عمل پیچش درون سلول LSTM استفاده می‌شود. با این کار ضرب ماتریس با عمل جمع شدن در هر دروازه داخل سلول حافظه جایگزین می‌شود. مجدداً از تابع فعال ساز Relu و لایه‌هایی با اتصالات کامل استفاده می‌شود. گام‌های زمانی در یادگیری، با توجه به پارامترها برابر ۱۵ قرار داده می‌شوند و ۴ روز آینده را به عنوان خروجی مورد نظر، شبکه پیش‌بینی خواهد کرد. نحوه‌ی کار به این صورت است که داده‌هایی پنجره زمانی اول را شبکه به عنوان ورودی دریافت می‌کند و ۴ داده بعدی را تخمین زده و سپس همین عملیات را از داده دوم شروع کرده و ادامه می‌دهد.



(شکل-۱۲): نمودار حافظه کوتاه مدت طولانی پیچشی

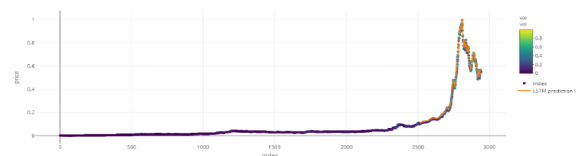
در ادامه مجدداً مقادیر دو معیار MAPE و RMSE محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۶۳۱ و RMSE برابر با ۰/۰۵۳۱ می‌باشد. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۷ درصد

(جدول-۸): نتایج میانگین خطای روش

شاخص	MAPE	RMSE
مقدار	۰/۰۵۳۷	۰/۰۵۱۷

شبکه عصبی LSTM

در این روش ابتدا وقفه‌های اول تا هشتم شاخص ۵۰ شرکت برتر و ارزش معاملات را محاسبه کرده و بعنوان متغیر مستقل به شبکه عصبی معرفی می‌کنیم و به آموزش مدل می‌پردازیم. روش کار بدین صورت است که هر روز یک داده به مدل اضافه می‌شود و شبکه عصبی دوباره آموزش داده می‌شود و یک پیش‌بینی برای روز بعد ارایه می‌دهد و اینکار آنقدر ادامه پیدا می‌کند تا اینکه همه داده‌ها پوشش داده شوند.



(شکل-۱۰): پیش‌بینی سری شاخص ۵۰ شرکت برتر توسط مدل

LSTM

در ادامه مجدداً مقادیر دو معیار MAPE و RMSE محاسبه شده‌اند. همانگونه که در جدول زیر مشاهده می‌شود مقدار MAPE برابر شده است با ۰/۰۶۰۷ و RMSE برابر با ۰/۰۵۰۱ می‌باشد. این مسئله نشان می‌دهد که میانگین خطای این روش بطور تقریبی برابر ۶ درصد می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل اتکایی را برای شاخص ارایه نمی‌دهد.

(جدول-۹): نتایج میانگین خطای روش

شاخص	MAPE	RMSE
مقدار	۰/۰۶۰۷	۰/۰۵۰۱

شبکه عصبی LSTM-GARCH

در این روش ابتدا وقفه‌های اول تا هشتم شاخص کل و ارزش معاملات را محاسبه کرده و بعنوان متغیر مستقل به شبکه عصبی معرفی می‌کنیم و به آموزش مدل می‌پردازیم. یعنی سری نوساناتی که از مدل GARCH گرفته می‌شود نیز به عنوان متغیر کمکی به مدل اضافه می‌کنیم. روش کار بدین صورت است که هر روز یک داده به مدل اضافه می‌شود و شبکه عصبی دوباره آموزش داده می‌شود و سپس با تلفیق دو مدل LSTM و GARCH پیش‌بینی برای روز بعد ارایه می‌دهد و اینکار آنقدر ادامه پیدا می‌کند تا اینکه همه داده‌ها پوشش داده شوند.

روش‌های فراوانی به جهت پیش‌بینی قیمتی ارائه شده است. مدل‌های اقتصادسنجی و استفاده از هوش مصنوعی و شبکه عصبی از این دست تلاش‌ها هستند. در سال‌های اخیر با استفاده فراگیر از کامپیوترهای با توان محاسباتی بالا در معاملات و تحلیل بازار و همچنین امکان دسترسی به حجم بالایی از داده‌های مرتبط با کشف قیمت، نیاز به مدلی که بتواند از میان این حجم انبوه از اطلاعات، به استخراج اطلاعاتی که بیشترین ارتباط را با قیمت داشته باشند، احساس می‌شود. مدل یادگیری عمیق به صورت پاسخی به این نیاز فراگیر ارائه شده است. ساختار و توابع استفاده شده در این مدل به نحوی هستند که در ابتدا به استخراج ویژگی‌های با تاثیر بالا بر روی قیمت می‌پردازد، سپس با استفاده از این ویژگی‌های منتخب به مدل‌سازی رفتار قیمت به منظور پیش‌بینی آن می‌پردازد. اگرچه این مدل در حوزه‌های مهندسی کاربرد فراوانی دارد، استفاده از آن در بازارهای مالی با سرعتی روزافزون رو به افزایش است.

در فصل گذشته مدل‌های سری زمانی ARIMA, LSTM و LSTM + Garch برآورد شد. در این بخش با استفاده از ۲۵۰۰ داده‌ی ابتدایی، مدل‌ها را برآزش داده و توسط سه مدل، روز بعد را پیش‌بینی کرده‌ایم. متغیرهای مورد استفاده در مدل‌ها شاخص‌های ۵۰ شرکت و شاخص کل بعنوان متغیرهای هدف و همچنین متغیر ارزش معاملات نیز بعنوان یک متغیر کمکی به مدل‌ها اضافه شده‌اند. لازم بذکر است وقفه‌های این متغیرها را با استفاده از نمودار تابع خود همبستگی جزئی نیز به مدل‌ها وارد کردیم. در ادامه با استفاده از یک پنجره داده از مبدا تا زمان هر روز یک مدل جدید با اضافه کردن داده‌ای جدید برآزش داده و یک پیش‌بینی برای روز بعد ارائه دادیم و این کار ادامه پیدا کرد تا اینکه همه داده‌ها پوشش داده شدند.

هدف این پژوهش بررسی این موضوع است که آیا می‌توان با روش شبکه عصبی LSTM شاخص کل و ۵۰ شرکت برتر را پیش‌بینی کرد یا خیر. سپس به بررسی این موضوع می‌پردازیم که آیا این پیش‌بینی نسبت به پیش‌بینی توسط روش ARIMA عملکرد بهتری داشته است؟ در فصل گذشته به این نتیجه رسیدیم که برای شاخص ۵۰ شرکت برتر مدل LSTM به تنهایی نمی‌تواند بهتر از مدل‌های خطی سری زمانی، آینده این شاخص را تبیین کند، اما با اضافه کردن مدل از نوع گارچ برای نوسانات این شاخص در این مدل به نتیجه به مراتب بهتری رسیدیم که نشان می‌دهد دقت این مدل در این حالت حدوداً دو برابر شده است. به عبارتی MAPE از ۰/۰۷ به ۰/۰۴ کاهش پیدا کرد که این خود پیشرفت چشمگیری است. همچنین مدل سری زمانی ARIMA برای این شاخص نیز پیش‌بینی قابل‌قبولی در مقابل مدل مذکور ارائه نداده است. به عبارتی معیار MAPE برای این مدل بسیار بیش از مدل LSTM+Garch شده است. برای شاخص کل با اینکه این مدل بهتر از نوع معمولی خود برآورد انجام داده ولی باز معیار MAPE بزرگ بوده و تفاوت چندانی با مدل سری‌های زمانی ندارد. در واقع این روش برای شاخص کل دقت قابل‌قبولی ارائه نداده است و

می‌باشد که با توجه به دامنه تغییرات ۵ درصدی شاخص، این روش پیش‌بینی قابل‌اتکایی را برای شاخص ارائه نمی‌دهد.

(جدول-۱۱): نتایج میانگین خطای روش

شاخص	MAPE	RMSE
مقدار	۰/۰۶۳۱	۰/۰۵۳۱

مقایسه چهار روش برای پیش‌بینی شاخص ۵۰ شرکت

در سه بخش قبل مشاهده کردیم که روش‌های LSTM و ARIMA مورد استفاده در پژوهش دارای دقت خوبی برای پیش‌بینی یک روز آینده شاخص ۵۰ شرکت برتر نیستند اما روش LSTM-GARCH در پیش‌بینی شاخص ۵۰ شرکت برتر عملکرد بهتری نسبت به پیش‌بینی شاخص کل دارد؛ چرا که دارای خطای پیش‌بینی کمتری نسبت به حالت قبل است. یک مقایسه با استفاده از دو معیار RMSE و MAPE برای پیش‌بینی شاخص ۵۰ شرکت برتر در زیر انجام شده است. این دو شاخص برای شبکه عصبی LSTM-GARCH تلفیقی از همه کمتر شده است که نشان می‌دهد این روش از دقت بیشتری برای پیش‌بینی شاخص ۵۰ شرکت بزرگ برخوردار است. بعد از آن نیز مدل ARIMA انتخاب شده با استفاده از معیار AIC دقت بیشتری داشته است. در نهایت هم مدل LSTM, CNN با بیشترین خطا نسبت به دو روش دیگر شاخص ۵۰ شرکت برتر را پیش‌بینی می‌کند.

(جدول-۱۲): نتایج میانگین خطای روش‌ها

شاخص	MAPE	RMSE
ARIMA	۰/۰۵۲۷	۰/۰۵۱۷
LSTM	۰/۰۶۰	۰/۰۵۰۱
LSTM-GARCH	۰/۰۳۱۶	۰/۰۲۰۷
CNN	۰/۰۶۳۱	۰/۰۵۳۱

۵- نتیجه گیری

بازار کارا، بازاری است که در آن اطلاعات با سرعت بالایی بر قیمت سهام تأثیر می‌گذارد و قیمت‌ها خود را با توجه به این اطلاعات تعدیل می‌کنند. در واقع بازار کارا به سرمایه‌گذاران این اطمینان را می‌دهد که همه آنها از اطلاعات مشابهی آگاهی دارند، پس بازار کارا، بازاری است که بازتاب‌دهنده آن اطلاعاتی است که در بازار موجود بوده و راهنمای سرمایه‌گذاران است. در بازار کارا نمی‌توان به بازده غیرعادی دست یافت. از طرفی انتقاداتی به بازار کارا وارد است و این اعتقاد که با پیش‌بینی قیمت می‌توان به بازده غیرعادی دست یافت در محافل علمی و عملی دارای پیروان فراوانی است. به این منظور تلاش‌های زیادی در جهت پیش‌بینی قیمت صورت پذیرفته است. تاکنون

International Conference on Computer Modelling and Simulation (pp. 106-112). IEEE.

[6]. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational research*, 270(2), 654-669

[7]. Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques—Part II: Soft computing methods. *Expert systems with applications*, 36(3), 5932-5941.

[8]. Avellaneda, M., & Lee, J. H. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, 10(7), 761-782.

[9]. Baker, M., Bradley, B., & Wurgler, J. (2011). Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*, 67(1), 40-54.

[10]. Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99(2), 427-446.

[11]. Bogomolov, T. (2013). Pairs trading based on statistical variability of the spread process. *Quantitative Finance*, 13(9), 1411-1430.

[12]. Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected idiosyncratic skewness. *The Review of Financial Studies*, 23(1), 169-202.

[13]. Britz, D. (2015). Recurrent neural networks tutorial. URL: <http://www.wildml.com/2015/09/recurrentneural-networks-tutorialpart-1-introduction-to-rnns/> [accessed June 28, 2019].

[14]. Abarbanell, J. S. (1997). Fundamental analysis, future earnings, and stock prices. *Journal of accounting research*, 35(1), 1-2.

[15]. Adebisi, A. A. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*.

[16]. Aggarwal, C. C. (2018). *Neural networks and deep learning*. Cham: Springer International Publishing.

[17]. Artificial neural network. [wikipedia.org](https://en.wikipedia.org): <https://en.wikipedia.org>

[18]. Bollerslev, T. M. (2014). Stock return predictability and variance risk premia: statistical inference and international evidence. *Journal of Financial and Quantitative Analysis*, 49(3), 633-661.

[19]. Chen, W.-C. (2008). Nonlinear dynamics and chaos in a fractional-order financial system. *Chaos, Solitons & Fractals*, 36(5), 1305-1314

[20]. Chen, W.-C. (2008). Nonlinear dynamics and chaos in a fractional-order financial system. *Chaos*,

با توجه به محاسبات زمان‌بر و پیچیده‌ای که دارد، نتوانسته است بهتر از مدل‌های سری زمانی معمولی عمل کند. لذا در این پژوهش به این نتیجه رسیدیم که استفاده از این ابزار بخصوص، همیشه کارایی نخواهد داشت. بطور مثال برای شاخص کل مشاهده شد که این روش تاثیر قابل توجهی در بهبود پیش‌بینی آینده شاخص کل ندارد، در حالی که برای شاخص ۵۰ شرکت، این مدل تا حد زیادی کیفیت و دقت پیش‌بینی را بالاتر برده است.

همچنین در این روش فقط یک روز بعد شاخص پیش‌بینی شده است و با توجه به اینکه معیارهای اندازه‌گیری خطا در کل بالا برآورد شده‌اند، می‌توان گفت این روش برای پیش‌بینی‌های بلند مدت کارایی چندانی ندارد.

محدودیت‌های تحقیق

از آنجایی که در بازار بورس تهران دامنه نوسان سهام محدود می‌باشد، این محدودیت نوسان خود به خود روی شاخص بورس هم تاثیر می‌گذارد. وجود دامنه نوسان می‌تواند روی شکل نمودارها و داده‌های روزانه تاثیر داشته باشد.

پیشنهادها

به محققان و پژوهشگران پیشنهاد می‌شود که از سایر مدل‌های شبکه عصبی نیز، نظیر شبکه عصبی بازگشتی، شبکه عصبی پرسپترون و شبکه‌های هیبریدی برای پیش‌بینی شاخص بورس اوراق بهادار تهران استفاده کنند و معیارهای فوق را در دوره‌های زمانی متفاوت‌تری نظیر ماهانه یا سالانه و همچنین اثر صعودی یا نزولی بودن بازار را بر نتیجه تحقیق به آزمایش بگذارند.

۶- منابع:

[1] Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for the stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications*, 113, 457-480.

[2] Abu-Mostafa, Y. S., & Atiya, A. F. (1996). Introduction to financial forecasting. *Applied intelligence*, 6(3), 205-213.

[3]. Adhikari, R. (2015). A mutual association-based nonlinear ensemble mechanism for time series forecasting. *Applied Intelligence*, 43(2), 233-250.

[4]. Akita, R., Yoshihara, A., Matsubara, T., & Uehara, K. (2016, June). Deep learning for stock prediction using numerical and textual information. In 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS) (pp. 1-6). IEEE.

[5]. Ariyo, A. A., Adewumi, A. O., & Ayo, C. K. (2014, March). Stock price prediction using the ARIMA model. In the 2014 UKSim-AMSS 16th

تجاری، سیستم های پیشنهاد دهنده، مدیریت دانش مشتری، داده کاوی، متن کاوی و فناوری اطلاعات در پزشکی.
Khosravi.280@gmail.com



محمدرضا نوروزی مدرک کارشناسی را از مرکز آموزش عالی محلات در رشته مهندسی کامپیوتر گرایش نرم افزار دریافت کرده است و زمینه پژوهش و علائق، یادگیری عمیق و سیستم های پیشنهاد دهنده می باشد

آدرس رایانه : norouzi_reza78@yahoo.com.

Solitons & Fractals. 36(5),1305-1314

[21] Ahmad Khan Beigi S. Abdolvand N, "Stock price prediction with Artificial Neural Network hybrid approach and colonial competition algorithm based on chaos theory." *Journal of Financial Management Strategy* 18 (2016): 73-27 (In Persian).

[22] Bayat, Ali; Asadi, Lida, "Optimization of Stock Portfolio: Utility of Birds Algorithm and Markowitz Model." *Financial Engineering and Managing Securities* 8.32 (2017): 63-85 (In Persian).

[23] Alizadeh Noodehi, et al, "Comparison of Efficiency and Risk in Technical Method with Balanced and Reinvestment and Purchase and Maintenance Strategy in Tehran Stock Exchange." *Journal of Investment Knowledge* 16 (2015): (in Persian).

[24] Habibzadeh M, Izadpour M, "Predicting profitability with neural network approach and comparing it with support vector machine (SVM) and C5 decision tree" *Financial knowledge of securities analysis* 13.46 (2020):39-56 (In Persian).

[25] Zarei Q, et al, "Comparison of fuzzy neural network methods with fuzzy wavelet neural network in predicting stock prices of banks admitted to Tehran Stock Exchange" *Financial Management Strategy* 6.3 (2017):109-138 (In Persian).

[26] Afsar, A., Halil, F, "Portfolio Optimization with Hybrid Approach of Technical Analysis and Data Mining Methods." *Journal on New Research in Decision Making* 2.2 (2017): 1-22 (In Persian).

[27] Heydari Zare B, Kordloi H, "Stock price prediction using artificial neural network" *Journal of Industrial Strategic Management* 7.17 (2010):49-56 (In Persian).



آرش خسروی مدرک کارشناسی خود

را در رشته مهندسی نرم افزار در سال ۱۳۸۲ از دانشگاه صنعتی اصفهان، مدرک ارشناسی ارشد خود را در رشته مهندسی فناوری اطلاعات در سال ۹۲ و مدرک دکتری خود

را در رشته مهندسی فناوری اطلاعات، گرایش سیستمهای اطلاعاتی در سال ۹۶ از دانشگاه صنعتی مالزی اخذ کرده است. ایشان در حال حاضر به عنوان هیات علمی مرکز آموزش عالی محلات مشغول به کار هستند. زمینه های پژوهشی مورد علاقه ایشان عبارتند از: هوش