فصلنامه علمی پژوهشی دانش سرمایه‌گذاری
دوره ۱٤ / شماره ۳ (پیاپی ۵۵) / پائیز ۱٤۰٤
صفحه ۹۳ تا ۱۱۷

بررسی عملکرد روش شبکه عصبی بازگشتی (RNN) حافظه طولانی کوتاه مدت (LSTM) در پیش بینی قیمت سهام

حسین عامری شهرابی
کارشناسی ارشد رشته مهندسی مالی دانشگاه خوارزمی
hosseinameri1995@gmail.com

احسان طیبی ثانی
استادیار دانشکده علوم مالی دانشگاه خوارزمی(نویسنده مسئول)
Ehsan_taieby_sani@khu.ac.ir

تاریخ دریافت: ۱۴۰۰/۰۵/۱۶ تاریخ پذیرش: ۱۴۰۱/۰۴/۰۶

چکیده

پیش‌بینی یکی از مهم‌ترین و چالش‌برانگیزترین مسائل در امور اقتصادی و مالی است که همواره مورد توجه افراد و محققان زیادی بوده است و از اهمیت بالایی برخوردار است. همواره برای پیش‌بینی در بازارهای مالی دو رویکرد سنتی و هوشمند وجود داشته که در روش سنتی این پیش‌بینی بر اساس مدل‌های آماری و در روش هوشمند بر اساس مدل‌های هوش مصنوعی است. با وجود مدل‌های فراوانی که برای پیش‌بینی ارائه‌شده است کماکان تلاش برای ساخت مدلی که بتواند متغیرهای مؤثر بیشتری را برای پیش‌بینی مورد استفاده قرار دهد و بتواند فاکتورهایی مانند زمان، ریسک و بازده را هم در نظر بگیرد، ادامه دارد. در پژوهش حاضر به پیش‌بینی بازدهی قیمت سهام در بورس اوراق بهادار تهران با استفاده از شبکه عصبی بازگشتی حافظه طولانی کوتاه‌مدت (LSTM) پرداخته شده است. از میان شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، ۲۰ شرکت برتری که حداقل در دوره زمانی ۱۳۹۴/۰۲/۱۱ تا ۱۳۹۹/۰۱/۲۲ حضور داشته و در آن بازه زمانی گپ قابل توجه قیمت نداشته‌اند (زیرا این موضوع منجر به خطای پیش‌بینی می‌گردید) به‌عنوان نمونه انتخاب شدند و داده‌ی تعدیل‌شده آن‌ها (به‌منظور حذف گپ ناشی از افزایش سرمایه و تقسیم سود نقدی) از سایت TSECLIENT دریافت گردید. میانگین و انحراف معیار خطای پیش‌بینی مدل شبکه عصبی بازگشتی (LSTM) حاکی از معناداری آماری مدل مذکور بوده است.

واژه‌های کلیدی: شبکه عصبی بازگشتی، حافظه طولانی کوتاه‌مدت، پیش‌بینی قیمت سهام.

۱- مقدمه

سرمایه‌گذاری با خرید و فروش سهام عرضه شده در بورس اوراق بهادار یکی از گزینه‌های پرسود در بازار سرمایه است اما بازار سرمایه دارای دستگاهی غیرخطی و نوسان پذیر است (شیلر[1]،۱۹۸۷). به همین خاطر تاکنون مدل‌های بسیاری که از ابزار و شیوه های مختلف بهره می‌بردند، برای پیش‌بینی روند بازار معرفی‌شده‌اند که هرکدام بازده و عملکردی متفاوت داشته‌اند. در حقیقت مدل‌های اقتصادی را می‌توان از طریق بررسی میزان دقت پیش‌بینی مورد آزمون قرار داد، بدین‌صورت که اگر مدلی درست عمل کند باید بتواند پیش‌بینی صحیحی از آینده متغیر نیز ارائه نماید.

از سوی دیگر مسائل اقتصادی و تجاری به شدت تحت تأثیر عوامل سیاسی، اجتماعی و فرهنگی هستند که بسیاری از این پارامترها ناشناخته بوده و به‌سختی قابل‌اندازه‌گیری‌اند. (آریوو همکاران[2]،۲۰۱۴) امروزه پیش‌بینی متغیرهای اقتصادی از اهمیت ویژه‌ای برای مدیران استراتژیک در بخش‌های خصوصی و دولتی جهت تنظیم امور و مناسبات اقتصادی برخوردار است به طوری که نیاز به ابزار و شیوه‌های پیش‌بینی متغیرها با کمترین مقدار خطا محسوس است. دستیابی به پیش‌بینی‌های با میانگین مطلق درصد خطای حدود ۱۰ درصد شاید آسان باشد ولی هزینه‌های خطا بسیار زیاد خواهد بود و تحقیقی که بتواند در کاهش چنددرصدی خطا کمک کند بسیار سودمند و مورد استفاده خواهد بود؛ زیرا در مقیاس کلان خطای ۱ درصد می‌تواند به تفاوت میلیونی یا میلیاردی سرمایه منجر شود؛ بنابراین انتخاب روشی که بتواند جواب را با حداقل خطای ممکن پیش‌بینی کند الزامی است.

به دلیل اهمیت و جایگاه ویژه بازارهای مالی و تأثیری که بر بازارهای موازی می‌گذارد که نشان‌دهنده نقش پررنگ آن در اقتصاد هر کشور است، پیش‌بینی در این حوزه از اهمیت خاصی برخوردار است و به بخش جدایی‌ناپذیر و پراهمیتی در این حوزه تبدیل‌شده است(بیگی، ۱۳۹۸). هر چه میزان دقت این پیش‌بینی بیشتر باشد و خطای کمتری داشته باشد باعث اعتماد بیشتر سرمایه‌گذاران می‌شود زیرا از این طریق می‌توانند ریسک خود را به حداقل برسانند. به منظور بالا بردن دقت پیش‌بینی می‌توان عوامل مهم تأثیرگذار بر بازارهای مالی را شناسایی و به کمک آن‌ها خطاهای پیش‌بینی را کمتر کرد (بیک و کیم[3]،۲۰۱۸). در این پژوهش با کمک تعدادی از متغیرهای کلان اقتصادی این کار انجام می‌شود.

بنابراین امروزه در راستای افزایش اطمینان از سرمایه گذاری و کاهش ریسک، با در اختیار داشتن اطلاعات مربوط به قیمت سهام درگذشته و استفاده از روش‌های علمی متفاوت، می‌توان به پیش‌بینی قیمت آن در آینده پرداخت. برخی از پژوهشگران نیز به مقایسه عملکرد این روش‌ها در تحقیقات خود می‌پردازند تا مدلی که خطای کمتری دارد را شناسایی کنند.

اهمیت پیشبینی نوسانات قیمت در بازارهای سرمایه از جنبه مدیریت ریسک برای بسیاری از سهامداران بازار سرمایه باهدف بهینه‌سازی پرتفوی دارایی‌های مالی مهم است. از سوی دیگر یکی از راههای شناسایی و اندازه‌گیری ریسک در بازارهای مالی برای مقامات ناظر و قانونگذار، رصد و پیش‌بینی نوسانات قیمت سهام و شاخص‌های قیمت

[1] Shiller,1987
[2] Ariyo et al.,2014
[3] Baek, & Kim,2018

بازاری است (مارکویتز[1]،۱۹۵۲). در این راستا، یکی از ویژگی‌هایی که ممکن است در رفتار بازارهای مالی مشاهده شود، وجود حافظه بلندمدت است، به این معنی‌که در صورت نامانا بودن سری زمانی، اثر شوک‌های خارجی بر نوسانات آن سری دائمی خواهد بود. در این رابطه علیرغم بهبودهای حاصل شده در توسعه مدل‌های اقتصادسنجی برای شناسایی این مسئله (نظیر ارایه مدلهای میانگین شرطی با حافظه بلندمدت(ARFIMA)) در پیش‌بینی صحیح داده‌های مالی، گروهی از محققین پیش‌تر به سراغ مدلهای هوش مصنوعی (به ویژه شبکه عصبی مصنوعی) و استفاده از قابلیت‌های آنها برای پیشبینی داده‌های سری زمانی رفته‌اند. اخیرا گروه دیگری از محققین نیز به دنبال ترکیب مدلهای اقتصادسنجی و هوش مصنوعی جهت استفاده از قابلیت‌های هر یک از آنها برای پیش‌بینی بهتر داده‌های سری زمانی مالی بوده‌اند. در این راستا محققین با ترکیب مدل‌های پیش‌بینی کننده فردی و طراحی مدلهای ترکیبی، اقدام به ارزیابی قدرت پیش‌بینی آن با مدلهای فردی کرده‌اند که در اغلب یافته‌ها، مدلهای ترکیبی خطای پیشبینی کمتری نسبت به مدلهای فردی داشته‌اند (ذوالفقاری و سحابی[2]،۲۰۱۷).

ازاینرو در مطالعه حاضر، برای نخستین بار در مطالعات داخلی، اقدام به ارایه مدل شبکه عصبی بازگشتی (RNN) حافظه طولانی کوتاه مدت (LSTM)برای پیش بینی قیمت سهام نموده است و سپس بررسی قدرت روش مزبور هستیم. بدین منظور در ابتدا به پیش‌بینی قیمت کوتاه‌مدت سهام ۲۰ شرکت برتر در بورس اوراق بهادار تهران پرداخته شده و در ادامه به بررسی عملکرد روش شبکه عصبی بازگشتی حافظه طولانی کوتاه‌مدت (LSTM) در پیش‌بینی قیمت سهام پرداخته می‌شود. دلیل انتخاب روش شبکه عصبی بازگشتی طولانی کوتاه‌مدت این است که در نسخه‌های اولیه از شبکه عصبی بازگشتی حافظه بلندمدت وجود ندارد و در صورتی‌که در ابتدای دنباله ویژگی مهمی به دست آید به گام‌های بعدی قابلیت انتقال ندارد.

در غالب پژوهش‌های انجام شده همانطور که در قسمت پیشینه تحقیق ارایه شده است به کرات از مدلهای شبکه عصبی مصنوعی معمولی مانند ANN در مدل‌های ترکیبی استفاده شده است. ولیکن تابحال پژوهشی که حاصل ترکیب مدل شبکه‌های عصبی بازگشتی(RNN) حافظه طولانی کوتاه مدت (LSTM) برای پیش بینی قیمت سهام باشد، استفاده نشده است و به نوعی این پژوهش اولین تحقیق در این حوزه می‌باشد که بازگوکننده نوآوری این مقاله می‌باشد. در مطالعات خارجی صورت گرفته در خصوص مقایسه این دو پیش‌بینی کنندها، کلیه یافته‌ها نشاندهنده برتری غالب مدل‌های شبکه‌های عصبی بازگشتی بر شبکه عصبی معمولی مانند (ANN) می‌باشد.

برخلاف شبکه عصبی مصنوعی سنتی، روش LSTM به کمک ماژول‌ها (واحدهای) مختلفی که دارد هر ویژگی دنباله را در خود ذخیره می‌کند و با توجه به ورود ویژگی‌های جدید ویژگی قبلی را حذف نمی‌نماید. در شبکه عصبی LSTM ما با مفاهیم جدیدی مواجه می‌شویم که در شبکه عصبی بازگشتی سنتی وجود ندارند. در این شبکه اصطلاحاً سه دروازه وجود دارد که از طریق آن شبکه نسبت به کنترل جریان داده درون خود اقدام می‌کند.

بنابراین LSTM در پیش‌بینی مسائل سری زمانی بسیار قدرتمند عمل می‌کند، زیرا قادر به ذخیره‌سازی اطلاعات گذشته است. در مصالح پیش‌بینی قیمت سهام، این ویژگی بسیار مهم است زیرا قیمت پیشین سهام برای

[1] Markowitz
[2] Zolfaghari and Sahabi

پیش‌بینی قیمت آینده آن عاملی حیاتی محسوب می‌شود؛ بنابراین هدف از پژوهش حاضر بررسی عملکرد روش شبکه عصبی بازگشتی (RNN) حافظه طولانی کوتاه مدت (LSTM) در پیش بینی قیمت سهام می باشد.

پس از پیاده‌سازی روش پیش‌بینی بر روی هر یک از ۲۰ نماد معرفی‌شده در فصول گذشته، به محاسبه ۴ نوع خطای پیش‌بینی پرداخته شد. پس از محاسبه انواع خطا ، لازم بود تا از لحاظ آماری به بررسی میانگین و واریانس خطاهای متناظر با روش در پیش‌بینی قیمت ۲۰ سهام پرداخته شود. پس‌ازانجام محاسبات نتایج حاصل حاکی از آن بود که در پیش‌بینی قیمت آینده ۲۰ نماد، از لحاظ آماری فرض صفر مبنی بر وجود خطای پیش بینی در روش LSTM را نمی‌توان رد نمود فلذا می‌توان دریافت که قدرت پیش‌بینی روش LSTM مناسب است.

مبانی نظری پژوهش

شبکه عصبی مصنوعی

یک شبکه عصبی مصنوعی ایده‌ای برای پردازش اطلاعات است که از سیستم عصبی زیستی الهام گرفته و مانند مغز به پردازش اطلاعات می‌پردازد. عنصر کلیدی این ایده، ساختار جدید سیستم پردازش اطلاعات است. این سیستم از شمار زیادی عناصر پردازشی فوق‌العاده به‌هم‌پیوسته به نام نورون‌ها تشکیل‌شده که برای حل یک مسئله باهم هماهنگ عمل می‌کنند. شبکه‌های عصبی مصنوعی نیز مانند انسان‌ها با مثال یاد می‌گیرند و یک شبکه عصبی برای انجام وظیفه‌های مشخص مانند شناسایی الگوها و دسته‌بندی اطلاعات، در طول یک پروسه یادگیری تنظیم می‌شود. در سیستم‌های زیستی، یادگیری با تنظیماتی در اتصالات سیناپسی که بین اعصاب قرار دارد همراه است. از این روش در شبکه‌های عصبی نیز استفاده می‌شود. شبکه‌های عصبی مصنوعی با پردازش داده‌های تجربی، دانش یا قانون نهفته در ورای داده‌ها را به ساختار شبکه منتقل می‌کند که به این عمل یادگیری می‌گویند. اصولاً توانایی یادگیری مهم‌ترین ویژگی یک سیستم هوشمند است. سیستمی که قابلیت یادگیری داشته باشد، منعطف‌تر است و ساده‌تر برنامه‌ریزی می‌شود، بنابراین بهتر می‌تواند در مورد مسائل و معادلات جدید پاسخگو باشد (جورابیان، ۱۳۸۶).

یک شبکه عصبی قادر است تا با متغیرهای ورودی به‌صورت موازی کار کند که نتیجه آن، توانایی کنترل داده‌های بی‌شماری با سرعت‌بالا است. دلیل اصلی که شبکه عصبی را بسیار نیرومند و متمایز می‌کند توانایی یافتن الگو است (چن و همکاران[1]،۲۰۰۷).

ساختار غیرخطی شبکه عصبی برای یافتن ارتباط بین اتفاقات و رخدادهای واقعی جهان استفاده می‌شود. شبکه‌ی عصبی جزء روش‌های چندبعدی به‌منظور پیش‌بینی می‌باشند که تنها چهارچوب غیرمالی رویدادها را در نظر نمی‌گیرد بلکه فرآیندهای خطی را نیز در نظر می‌گیرد. برای مثال شبکه‌های عصبی این توانایی رادارند تا برای مدل کردن سری زمانی خطی مورداستفاده قرار بگیرند که درواقع مطلب فوق را تأیید می‌کند (Moghar & Hamiche,2020).

[1] Chen et al,2007

یکی از کاربردهای اصلی شبکه عصبی درزمینه پیش‌بینی است که اخیراً فعالیت‌های فراوانی درزمینه پیش‌بینی با استفاده از شبکه عصبی انجام‌گرفته است. توانایی پیش‌بینی دقیق آینده یکی از اهداف اساسی فرآیند تصمیم‌گیری، برنامه‌ریزی، زمان‌بندی، خرید، تعیین استراتژی، سیاست‌گذاری و فرآیند زنجیره تأمین است و هنوز هم پیش‌بینی یک جزء بااهمیت در فعالیت‌های بشری است (راعی و همکاران،۱۳۹۱).

انواع شبکه عصبی

شبکه‌های عصبی تنوع بسیار زیادی دارند، درواقع ما ۲۷ نوع شبکه عصبی داریم دودسته‌ی پرکاربرد در شبکه‌های عصبی، شبکه‌های عصبی پیشرو[1] (پیش‌خور) و شبکه‌های عصبی بازگشتی[2] می‌باشند که در این پژوهش به تشریح دسته دوم می پردازیم.

شبکه عصبی بازگشتی[3]

شبکه عصبی بازگشتی که به آن شبکه عصبی مکرر نیز گفته می‌شود، نوعی از شبکه عصبی مصنوعی است که در تشخیص گفتار، پردازش زبان طبیعی[4] و همچنین در پردازش داده‌های ترتیبی استفاده می‌شود. بسیاری از شبکه‌های عمیق مانند شبکه عصبی متحرک[5] شبکه‌های پیش‌خور هستند؛ یعنی سیگنال در این شبکه‌ها فقط در یک‌جهت از لایه ورودی، به لایه‌های مخفی و سپس به لایه خروجی حرکت می‌کند و داده‌های قبلی به حافظه سپرده نمی‌شوند؛ اما شبکه‌های عصبی بازگشتی یک‌لایه بازخورد دارند که در آن خروجی شبکه به همراه ورودی بعدی، به شبکه بازگرداننده می‌شود. می‌تواند به علت داشتن حافظه داخلی، ورودی قبلی خود را به خاطر بسپارد و از این حافظه برای پردازش دنبال‌هایی از ورودی‌ها استفاده کند. به بیان ساده، شبکه‌های عصبی بازگشتی شامل یک حلقه بازگشتی هستند که موجب می‌شود اطلاعاتی را که از لحظات قبلی به دست آورده‌ایم از بین نروند و در شبکه باقی بمانند (سلوین و همکاران[6]،۲۰۱۷).

مفهوم شبکه‌های عصبی بازگشتی در سال ۱۹۸۰ ایجاد گردید. این‌گونه از شبکه‌های عصبی به‌طور خاص برای پردازش داده‌های سری یا دنباله‌دار مفید هستند و در آن‌ها هر نورون یا واحد پردازشی قادر به حفظ حالت داخلی یا همان حافظه به‌منظور حفظ اطلاعات مرتبط با ورودی قبلی است. این ویژگی به‌طور خاص در کاربردهای مختلف مرتبط با داده‌های سری اهمیت اساسی پیدا می‌کند.

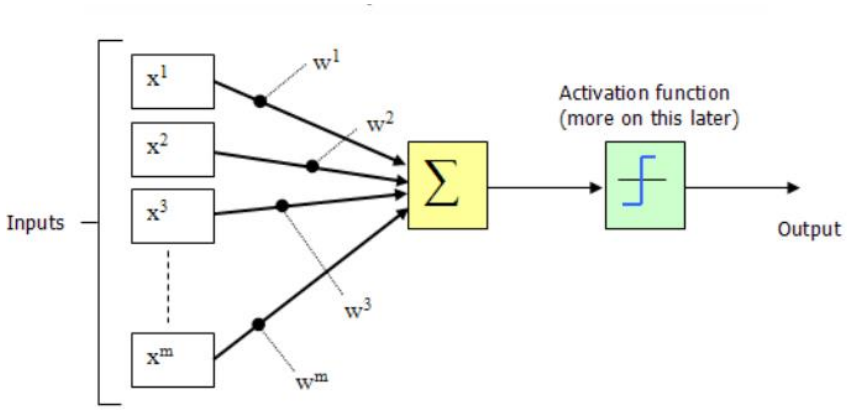
ایده اصلی پشت این نوع از معماری بهره‌برداری از این ساختار سری داده است. نام این شبکه عصبی از این واقعیت به دست می‌آید که این نوع از شبکه‌ها به‌صورت بازگشتی عمل می‌کنند؛ یعنی یک عملیات برای تک‌تک المان‌های یک دنباله انجام می‌گیرید و خروجی آن وابسته به ورودی فعلی و عملیات قبلی است. این مهم از طریق

[1]-Feed Forward Neural Networks
[2]- Recurrent Neural Networks
[3]- Recurrent Neural Network (RNN)
[4]- natural language processing
[5] -convolutional neural network
[6] Selvin et al,2017

تکرار یک خروجی از شبکه در زمان t با ورودی شبکه در زمان ۱ +t انجام می‌شود. (خروجی از مرحله قبل با ورودی تازه در مرحله جدید ترکیب می‌شوند.) این چرخه‌ها اجازه وجود اطلاعات از یک گام زمانی به گام زمانی بعدی را موجب می‌شوند. به عبارت بهتر این نوع شبکه‌ها دارای حلقه‌ای در درون خود هستند که به‌وسیله آن می‌توانند اطلاعات را در حین خواندن ورودی از نورون‌ها عبور دهند. (زرمبا و همکاران[1]،۲۰۱۴)

ساختار شبکه عصبی

شبکه‌ی عصبی یک سامانه پردازش داده‌ها است که از مغز انسان ایده گرفته و پردازش داده‌ها را به عهده پردازنده‌های کوچک و بسیار زیادی می‌سپارد که به‌صورت شبکه‌ای به‌هم‌پیوسته و موازی با یکدیگر برای حل یک مسئله رفتار می‌کنند. در این شبکه‌ها به کمک دانش برنامه‌نویسی، ساختار داده‌ای طراحی می‌شود که می‌تواند همانند نورون عمل کند که به آن ساختار داده گره گفته می‌شود. در این ساختار با ایجاد شبکه‌ای بین این گره‌ها و اعمال یک الگوریتم آموزشی به آن، شبکه را آموزش می‌دهند. در این حافظه یا شبکه عصبی گره‌ها دارای دو حالت فعال و غیرفعال‌اند و هر یال دارای یک وزن است. یال‌های با وزن مثبت، موجب تحریک یا فعال کردن گره غیرفعال بعدی می‌شوند و یال‌های با وزن منفی، گره متصل بعدی را غیرفعال می‌کنند (رمضانیان و همکاران،۱۳۹۸).



شکل ۱: ساختار شبکه عصبی

اطلاعات از طریق ورودی یا همان دندریت وارد نورون می‌شوند که با مقادیر ($x_1,\ldots\ldots,x_m$) در تصویر قابل‌مشاهده هستند. در مدل شبکه عصبی مصنوعی به هر ورودی یک وزن ($w_1,\ldots\ldots,w_m$) اختصاص می‌دهیم. این وزن‌ها درواقع اهمیت ورودی‌ها برای ما هستند، یعنی هر چه وزن بیشتر باشد، ورودی برای آموزش شبکه مهم‌تر است.

[1] Zaremba et al., 2014

سپس تمامی ورودی‌ها باهم جمع (Σ) شده و به‌صورت یک‌لایه به آکسون وارد می‌شوند. در مرحله بعد تابع فعالیت[1] را بر روی داده‌ها اعمال می‌کنیم. پس‌ازآنجام محاسبات در این مرحله اطلاعات ما از طریق سیناپس‌های خروجی وارد نورون دیگر می‌شوند و این مرحله تا جایی ادامه پیدا می‌کند که به لایه خروجی برسیم (رمضانیان و همکاران،۱۳۹۸).

یادگیری در شبکه‌های عصبی

شبکه‌های عصبی این توانایی رادارند که سیگنال‌های ورودی را طبق طراحی که انجام می‌دهند پردازش کرده و به سیگنال‌های خروجی موردنظر تبدیل کنند. معمولاً پس از طراحی و پیاده‌سازی شبکه عصبی، تنظیم سیگنال-های ورودی باید به‌گونه‌ای باشد که سیگنال‌های خروجی بتوانند شبکه خروجی موردنظر را به وجود بیاورند. مجموعه این اتفاقات آموزش شبکه عصبی می‌گویند. می‌توان گفت آموزش فرآیندی است که نتیجه آن یافتن بهترین وزن‌ها برای ورودی هر سلول شبکه عصبی است. درواقع آموزش به دنبال این موضوع است که وزن‌هایی را برای سلول‌های ورودی شبکه عصبی انتخاب نماید که خروجی به‌دست‌آمده کمترین خطا را نسبت به حالت مطلوب داشته باشد.

در مراحل ابتدایی آموزش، مقادیر را به‌صورت تصادفی انتخاب می‌کنیم. در طول مدت‌زمان انجام فرآیند آموزش توسط شبکه عصبی، پارامترها به مقدار واقعی‌شان نزدیک‌تر می‌شوند. درمجموع می‌توان از دو روش با ناظر و بدون ناظر برای آموزش دادن به شبکه‌های عصبی استفاده کرد (اندرسون و همکاران[2]،۱۹۹۲).

مکانیسم تست

مکانیسم تست اطلاعات را از ساختار شبکه عصبی بازیابی می‌کند. ساختار شبکه عصبی، نقش حافظه‌ای را دارد که توسط مکانیسم آموزش پر می‌شود و سپس توسط مکانیسم تست از اطلاعات ذخیره‌شده بهره‌برداری صورت می‌گیرد. مکانیسم تست اطلاعاتی را که مکانیسم آموزش با الگوریتم استخراج کرده، استفاده می‌نماید و جوابی متناسب با هر سؤال ارائه می‌دهد. کارایی یک شبکه عصبی با فراخوانی آن در فاز تست تعیین می‌شود، درواقع شبکه عصبی سؤالات کاربران را در فاز تست دریافت می‌کند. به این نمونه‌ها، نمونه تست الگوی فراخوانی می‌گویند. نمونه تست یا الگویی فراخوانی، فاقد خروجی هستند و شبکه عصبی باید خروجی این نمونه‌ها را تخمین بزند. مکانیسم تست یا مکانیسم فراخوانی شبکه عصبی، خروجی نمونه‌های تست را با استفاده از نگاشت به‌دست‌آمده در فاز آموزش تخمین می‌زند. هر چه پاسخ شبکه عصبی برای نمونه‌های تست به پاسخ واقعی نزدیک‌تر باشد نشان‌دهنده این است که نگاشت تشکیل‌شده به‌وسیله شبکه عصبی به نگاشت سیستم مسئله نزدیک‌تر است (موقار[3]،۲۰۲۰).

[1]-Activation Function
[2] Anderson et al,1992
[3] Moghar,2020

شبکه عصبی بازگشتی حافظه طولانی کوتاه‌مدت[1]

نسخه‌های اولیه شبکه عصبی بازگشتی در ذخیره‌سازی اطلاعات مرتبط با ورودی‌های گذشته به مدت طولانی ناتوان است. این موضوع شبکه را در مدل‌سازی ساختارهای بلندمدت تضعیف می‌کند و این «فراموشی» باعث می‌شود تا این نوع از شبکه‌ها در زمان تولید دنباله در معرض ناپایداری قرار گیرند. مشکلی که وجود دارد این است که اگر پیش‌بینی‌های شبکه تنها وابسته به چند ورودی اخیر باشد و این ورودی‌ها خود نیز توسط شبکه تولیدشده باشند، شانس بسیار کمی برای تصحیح و جبران اشتباهات گذشته توسط شبکه وجود دارد. داشتن یک حافظه بلندمدت‌تر دارای اثر تثبیت‌کننده است چراکه حتی اگر شبکه نتواند از تاریخچه اخیر خود درک صحیحی پیدا کند، باز بااین‌وجود قادر است با نگاه درگذشته پیش‌بینی خود را کامل کند. حافظه طولانی کوتاه‌مدت یا به‌اختصار LSTM یک معماری شبکه عصبی بازگشتی است که برای ذخیره‌سازی و دسترسی بهتر به اطلاعات نسبت به نسخه اولیه آن طراحی‌شده است. برخلاف شبکه عصبی بازگشتی سنتی که در آن محتوا در هر گام زمانی از نو بازنویسی می‌شود در یک شبکه عصبی بازگشتیLSTM ، شبکه قادر است نسبت به حفظ حافظه فعلی تصمیم‌گیری کند. LSTM در پیش‌بینی برای مسائل دارای توالی بسیار قدرتمند عمل می‌کند، زیرا قادر به ذخیره‌سازی اطلاعات گذشته است. در مصالح پیش‌بینی قیمت سهام، این ویژگی بسیار مهم محسوب می‌شود زیرا قیمت پیشین سهام برای پیش‌بینی قیمت آینده آن عاملی حیاتی محسوب می‌شود(سلوین و همکاران،۲۰۱۷).

مروری بر پیشینه پژوهش

ذوالفقاری و همکارانش در پژوهش خود در سال ۱۳۹۹ به معرفی یک مدل ترکیبی از شبکه یادگیری عمیق (Learning Deep) و مدل‌های منتخب خانواده GARCH جهت پیش‌بینی کوتاه‌مدت بازدهی روزانه شاخص کل بورس اوراق بهادار تهران پرداخته اند. در پژوهش مذکور از میان مدل‌های شبکه یادگیری عمیق، شبکه عصبی بازگشتی مبتنی بر حافظه کوتاه‌مدت و بلندمدت(LSTM-RNN) انتخاب و از مدل‌های دارای حافظه‌کوتاه مدت GARCHو EGARCH در ساختار آن استفاده می‌شود. همچنین دو متغیر مستقل قیمت نفت و نرخ دلار در ساختار مدل ترکیبی، کمک فراوانی به آن در پیش‌بینی دقیقتر داده‌های مالی می‌کند. نتایج تحقیق نشان میدهد که مدل‌های ترکیبی دقت پیشبینی باالتری نسبت به مدلهای تکی دارند. همچنین براساس معیارهای ارزیابی خطای پیشبینی RMSE و MAPE ،مدل EGARCH-LSTM-RNN برپایه توزیع GED دارای خطای پیشبینی کمتری نسبت به ۲۳ مدل دیگر دارد. در این راستا، معیار بررسی صحت پیشبینی دیبولد-ماریانو (DM) نیز یافته‌های فوق را تایید می‌کند.

جعفر باباجانی و همکارانش در ۱۳۹۸، در این پژوهش با رویکرد ترکیبی، با به‌کارگیری شبکه عصبی بازگشتی مبتنی بر الگوریتم کلونی زنبورعسل مصنوعی(ABC-RNN) ، درصدد ارائه مدلی بهینه برای پیش‌بینی قیمت سهام در بورس تهران بودند. . نتایج نشان‌دهنده آن است که استفاده از شبکه عصبی بهینه‌شده با الگوریتم کلونی زنبورعسل مصنوعی، دقت قابل‌ملاحظه‌ای در مقایسه با سایر روش‌های پیش‌بینی دارد (باباجانی و دیگران، ۱۳۹۸).

[1]- Long-Short Term Memory (LSTM)

احمدخان بیگی و عبدالوند در سال ۱۳۹۶، با توجه به ارتباطات غیرخطی موجود میان متغیرهای مؤثر بر قیمت سهام، شبکه‌های عصبی مصنوعی را یکی از مناسب‌ترین رویکردهای موجود جهت پیش‌بینی قیمت سهام دانستند. در این مقاله سعی شده تا از طریق ترکیب نگاشت‌های آشوبی و الگوریتم رقابت استعماری، زاویه حرکتی مستعمرات به سمت استعمارگر اصلاح‌شده و به‌این‌ترتیب احتمال قرارگیری در دام نقطه بهینه محلی تا حد ممکن کاهش یابد. هدف این مقاله معرفی و مقایسه عملکرد رویکرد پیشنهادی با سایر الگوریتم‌های بهینه‌سازی جستجوی پیشین است. ازاین‌رو با استفاده از اطلاعات قیمتی روزانه سهام شرکت ایران‌خودرو بین سال‌های ۱۳۸۹ تا ۱۳۹۵ به آموزش شبکه عصبی با الگوریتم‌های بهینه‌سازی مختلف بپردازند. جهت ارزیابی میزان عملکرد رویکردها، از سه دیدگاه: میزان دقت پیش‌بینی (آماره‌های اندازه‌گیری خطا R2,RMSE)، میزان حافظه مصرفی و زمان اجرایی استفاده شد، نتایج حاکی از آن است که رویکرد پیشنهادی از عملکرد بهتری نسبت به سایر رویکردهای پیشین برخوردار است. (احمدخان بیگی و عبدالوند، ۱۳۹۶)

بیات و باقری در سال ۱۳۹۶، در این پژوهش به پیش‌بینی قیمت سهام ۱۰ شرکت از شرکت‌های پذیرفته‌شده در بورس و تعدادی از شرکت‌های حاضر در فرابورس با استفاده از الگوریتم کرم شب‌تاب پرداخته‌اند. در این پژوهش از الگوریتم آموزش‌دیده برای پیش‌بینی قیمت سهام استفاده می‌گردد. طرز کار الگوریتم نیز به این صورت است که داده‌های روزانه, ماهانه یا سالانه (N) به الگوریتم داده می‌شود. الگوریتم به پیش‌بینی قیمت سهام برای روز، ماه یا سال N+۱ ام می‌پردازد. برای محاسبه‌ی خطای پیش‌بینی از محاسبه خطای نسبی استفاده‌شده است. محاسبات انجام خطای کمتر از ۶٪ را برای پیش‌بینی نشان می‌دهد؛ بنابراین آن‌ها در این پژوهش نتیجه گرفتند که الگوریتم کرم شب‌تاب قابلیت پیش‌بینی قیمت سهام رادار است. (بیات و باقری، ۱۳۹۶)

خوات و لی به کاربرد گسترده منطق فازی در پیش‌بینی تغییرات آینده، قیمت سهام و سایر مسائل مالی سری زمانی و همچنین مقبولیت فراگیر روش‌های داده‌کاوی در توانایی یادگیری و شناسایی روابط بین متغیرها، در پژوهش خود به مقایسه عملکرد منطق فازی[1] و ادراک چندلایه‌ای[2] که نوعی شبکه عصبی مصنوعی است، پرداخته‌اند تا اثربخشی هریک از آن‌ها را بررسی نمایند. منطق فازی روشی برای نتیجه‌گیری قطعی از اطلاعات مبهم و نادقیق است. شبکه عصبی مصنوعی عملکرد بهتری نسبت به مدل‌های رگرسیون آماری دارد. این روش به تجزیه‌وتحلیل عمیق مجموعه داده‌های بزرگ، به‌ویژه مواردی که تمایل به نوسان در مدت‌زمانی کوتاه دارند، می‌پردازد. در این پژوهش روش‌های مذکور بر روی‌داده‌های قیمت سهام چند شرکت پیاده‌سازی و در نهایت اثربخش هر یک از آن‌ها بررسی گردید. (خوانت و لی[3]،۲۰۱۷)

ژانگ و همکاران[4] (۲۰۱۹)، در مطالعه‌ی خود با عنوان پیش‌بینی شاخص بازار بر اساس شبکه‌ی معماری جدیدی از شبکه‌ی مولد متقابل[5] پرداختند. آن‌ها در پژوهش خود حافظه طولانی کوتاه‌مدت[6] را به‌عنوان

[1]- Fuzzy Logic (FL)
[2]- Multilayer Perception (MLP)
[3] Khuat and Le, 2017
[4] Zhang & et al
[5] generative adversarial network
[6] Long short-term memory

تعیین‌کننده‌ی پیش‌بینی قیمت بسته شدن سهام در نظر گرفتند. همچنین آن‌ها شبکه‌ی مولد متقابل را با چندلایه پرسپترون[1] به‌عنوان تشخیص‌دهنده پیشنهاد دادند. نتایج این پژوهش نشان داد که می‌تواند در مقایسه با سایر مدل‌ها در یادگیری ماشین و یادگیری عمیق عملکرد بهتری داشته باشد.

سیدرا و همکاران[2] (۲۰۲۰) در پژوهش خود اقدام به ارائه مدلی دو رگه برای برمبنای یادگیری ماشین و یادگیری عمیق برای پیش بینی شاخص نیفتی[3] ۵۰ بورس هند پرداخته‌اند. داده های مربوط به شاخص مذکور را از دسامبر ۲۰۱۴ تا جولای ۲۰۲۰ مورد بررسی قرار داده‌اند. برای افزایش قدرت پیش‌بینی مدل رگرسیون پیشنهادی خود از شبکه حافظه کوتاه مدت-بلندمدت[4] با روش اعتبارسنجی گام پیش‌رو[5] و جستجوی شبکه‌ای[6] استفاده نمودند که در نهایت نشان دادند مدل‌های بر مبنای حافظه کوتاه مدت-بلندمدت بر مبنای داده‌های هفته آخر بیشترین قدرت پیش‌بینی کنندگی را دارا است.

کومار[7] و همکاران(۲۰۲۱) در پژوهش اخیر خود نتیجه می گیرند که مدل‌های مختلفی مانند میانگین متحرک یکپارچه بازگشتی خودکار (ARIMA)، ماشین برداری (SVM)، شبکه عصبی مصنوعی (ANN)، XGBoost و شبکه عصبی بازگشتی (RNN) از دقت بالاتری برخوردارند. در این مطالعه، یک مدل حافظه طولانی کوتاه مدت (LSTM) برای پیش‌بینی دقت بازار سهام پیشنهاد شده و مدل پیشنهادی با میانگین متحرک (MA) و مدل‌های XGBoost مقایسه می‌شود. آزمایش‌ها بر روی مجموعه داده‌های تاریخی در بورس اوراق بهادار بمبئی، هند (BSE30) انجام شده است. این مدل همچنین از طریق معیارهای عملکرد ریشه میانگین مربعات خطا (RMSE) و میانگین درصد مطلق خطا (MAPE) ارزیابی می‌شود و نشان می‌دهد که مدل LSTM نسبت به مدل میانگین متحرک عملکرد بهتری دارد.

پاندی کومار و همکاران[8] (۲۰۲۲) از الگوریتم یادگیری عمیق شبکه عصبی حافظه طولانی کوتاه مدت (LSTM) همراه با روش حرکت قیمت سهام[9] برای پیش‌بینی حرکت قیمت در طی روز جهت پیش‌بینی کوتاه‌مدت استفاده می‌کند. مجموعه داده از تاریخ، قیمت بازگشایی، بالاترین، پایین ترین، قیمت پایانی و حجم معاملات استفاده می کند. پیش‌بینی‌های دقت حرکت قیمت بر روی سهام بانک دولتی هند (SBI) صورت گرفت و یک سال از مجموعه داده‌های معاملاتی مورد استفاده برای آموزش الگوریتم آزمایش شدند. یافته ها حاکی از آن بود که الگوریتم پیشنهادی پیش بینی دقت حرکت قیمت را تا ۹۸.۹٪ درست پیش‌بینی نموده است.

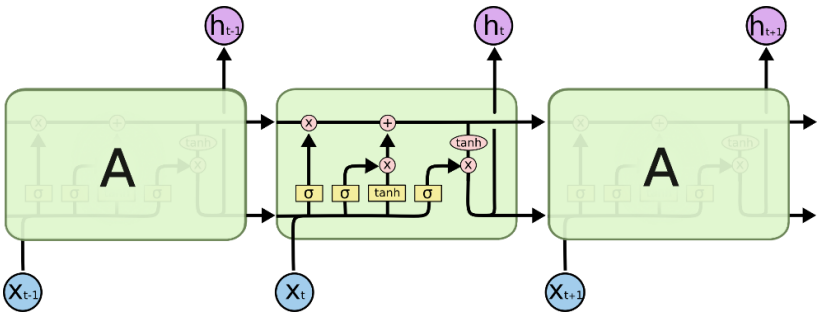
[1] multilayer perceptron
[2] Sidra et al
[3] The NIFTY 50 is a benchmark Indian stock market index that represents the weighted average of 50 of the largest Indian companies listed on the National Stock Exchange.
[4] Long-short term memory(LSTM)
[5] Walk-forward Validation
[6] Grid Searching
[7] kumar
[8] Pandikumar
[9] Price action

روش تحقیق

پژوهش حاضر از نظر هدف کاربردی می باشد. جامعه آماری این پژوهش، به شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران به دلیل شفافیت بیشتر قیمت‌ها محدودشده است. از میان شرکت‌های پذیرفته شده در آن نیز، ۲۰ شرکت برتری که حداقل در دوره زمانی ۱۳۹۴/۰۲/۱۱ تا ۱۳۹۹/۰۱/۲۲ حضورداشته و در آن بازه زمانی گپ قابل توجه قیمت نداشته‌اند (زیرا این موضوع منجر به خطای پیش‌بینی می‌گردید) به‌عنوان نمونه انتخاب شدند و داده‌ی تعدیل‌شده آن‌ها (به‌منظور حذف گپ ناشی از افزایش سرمایه و تقسیم سود نقدی) از سایت TSECLIENT دریافت گردید.

روش شبکه عصبی بازگشتی (RNN) حافظه طولانی کوتاه‌مدت (LSTM)

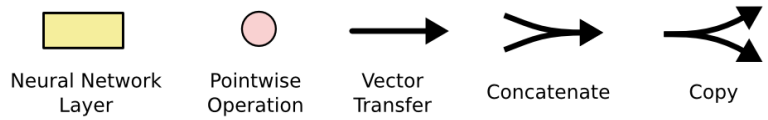
شبکه‌های LSTM ساختار دنباله یا زنجیره‌ای مانند شکل زیر دارند و ماژولِ تکرارشونده آن‌ها ساختار متفاوتی نسبت به شبکه‌های عصبی بازگشتی اولیه دارد. آن‌ها به‌جای داشتن تنها یک‌لایه شبکه عصبی، ۴ لایه دارند که طبق ساختار ویژه‌ای با یکدیگر در تعامل و ارتباط هستند (هاچریتر و شیمدبر[1]، ۱۹۹۷).



شکل ۲: ماژول‌های تکرارشونده در LSTM

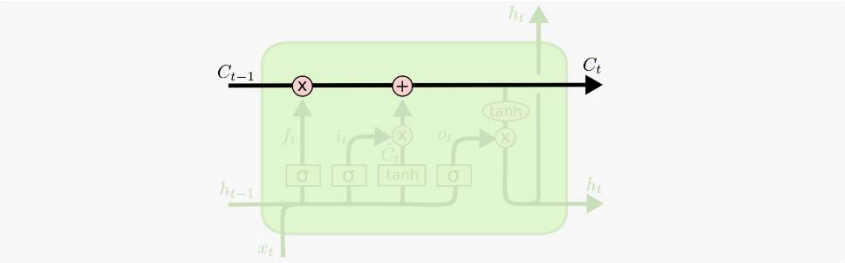
برای درک بهتر علائم استفاده‌شده در شکل بالا، در شکل زیر راهنمای علائم آمده است که به ترتیب از راست به چپ کپی کردن، وصل کردن، بردار انتقال، عملیات نقطه‌به‌نقطه و یک‌لایه‌ی شبکه عصبی هستند. به‌عبارتی‌دیگر هر خط یک بردار را به‌صورت کامل از خروجی یک گره به ورودی گره دیگر انتقال می‌دهد. دایره‌های صورتی نمایش‌دهنده عملیات نقطه به نقطه مانند «جمع‌کردن دو بردار» هستند. مستطیل‌های زرد، لایه‌های شبکه‌های عصبی هستند که شبکه پارامترهای آن‌ها را یاد می‌گیرد. خط‌هایی که باهم ادغام می‌شوند نشان‌دهنده الحاق و خط‌هایی که چند شاخه می‌شوند نشان‌دهنده‌ای این موضوع است که محتوای آن‌ها کپی و به بخش‌های مختلف ارسال می‌شود.

[1] Hochreiter and Schmidhuber, 1997



شکل ۳: علائم مورداستفاده در شکل LSTM

ایده اصلی LSTM سلول حالت[1] است که در حقیقت یک خط افقی است که در بالای شکل زیر قرار دارد. سلول حالت را می‌توان به‌صورت یک تسمه‌نقاله تصور کرد که از اول تا آخر دنباله یا همان زنجیره با تعاملات خطیِ جزئی درحرکت است (یعنی ساختار آن بسیار ساده است و تغییرات کمی در آن اتفاق می‌افتد).

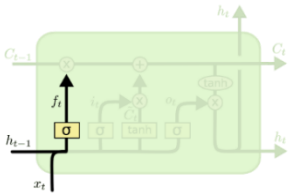


شکل ۴: سلول حالت LSTM

LSTM این توانائی را دارد که اطلاعات جدیدی را به سلول حالت اضافه یا اطلاعات آن را حذف کنید. این کار توسط ساختارهای دقیقی به نام دروازه‌ها[2] انجام می‌شود. دروازه‌ها راهی هستند برای ورود اختیاری اطلاعات. آن‌ها از یک‌لایه شبکه عصبی سیگموید[3] به همراه یک عملگر ضرب نقطه‌به‌نقطه تشکیل‌شده‌اند. خروجی لایه سیگموید عددی بین صفر و یک است که نشان می‌دهد چه مقدار از وروی باید به خروجی ارسال شود. مقدار صفر یعنی هیچ اطلاعاتی نباید به خروجی ارسال شود درحالی‌که مقدار یک یعنی تمام ورودی به خروجی ارسال شود. LSTM دارای ۳ دروازه مشابه برای کنترل مقدار سلول حالت است.

قدم اول در LSTM تصمیم در مورد اطلاعاتی است که می‌خواهیم آن‌ها را از سلول حالت پاک‌کنیم. این تصمیم توسط یک‌لایه سیگموید به نام «دروازه فراموشی»[4] انجام می‌شود. این دروازه با توجه به مقادیر h_{t-1} و x_t، برای هر عدد، مقدار صفر یا یک را در سلول حالتِ C_{t-1} به خروجی می‌برد. مقدار یک یعنی به‌صورت کامل مقدار حال حاضرِ سلول حالت (C_{t-1}) را به C_t ببر و مقدار صفر یعنی به‌صورت کامل اطلاعات سلول حالت کنونی (C_{t-1}) را پاک‌کن و هیچ مقداری از آن را به C_t نبر. شکل زیر تکمیل‌کننده توضیحات این دروازه است:

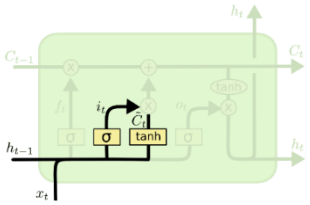
[1] - Cell State
[2] - Gates
[3] - Sigmoid
[4] - Forget Gate



$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

شکل ۵: دروازه فراموشی LSTM

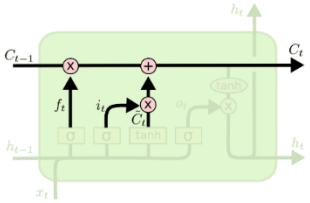
قدم بعدی این است که تصمیم بگیریم چه اطلاعات جدیدی را می‌خواهیم در سلول حالت ذخیره کنیم. این تصمیم دوبخشی است. ابتدا یک‌لایه سیگموید به نام دروازه ورودی[1] داریم که تصمیم می‌گیرد چه مقادیری به‌روز خواهند شد. مرحله بعدی یک‌لایه تانژانت هایپربولیک[2] است که برداری از مقادیر به نام C_t~ می‌سازد که می‌توان آن‌ها را به سلول حالت اضافه کرد. در مرحله بعد، ما این دو مرحله را باهم ترکیب می‌کنیم تا مقدار سلول حالت را به‌روز کنیم. شکل زیر تکمیل‌کننده توضیحات ارائه‌شده است:



$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

شکل ۶: دروازه ورودی و تانژانت هایپربولیک در LSTM

حال زمان آن فرا رسیده است که سلول حالت قدیمی یعنی C_{t-1} را سلول حالت جدید یعنی C_t به‌روز کنیم. در مراحل قبلی تصمیم گرفته شد که چه کنیم و در حال حاضر تنها لازم است تصمیماتی را که گرفته شد عملی کنیم. لذا مقدار قبلی سلول حالت را در f_t ضرب می‌کنیم که یعنی فراموش کردن اطلاعاتی که پیش‌تر تصمیم گرفتیم آن‌ها را فراموش کنیم. سپس $i_t * C_t$~ را به آن اضافه می‌کنیم. در حال حاضر مقادیر جدید سلول حالت با توجه به تصمیماتی که پیش‌تر گرفته‌شده بود به‌دست‌آمده‌اند. شکل زیر تکمیل‌کننده‌ی توضیحات ارائه‌شده است:

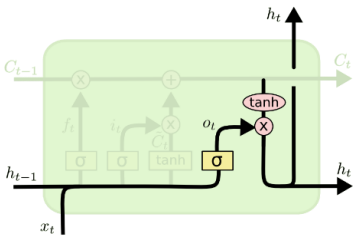


$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

شکل ۷: مرحله اعمال تغییرات در LSTM

[1] - Input Gate
[2] - Hyperbolic Tangent

در نهایت باید تصمیم گرفت قرار است چه اطلاعاتی را به خروجی ببریم. این خروجی با در نظر گرفتن مقدار سلول حالت خواهد بود، ولی از فیلتر مشخصی عبور خواهد کرد. در ابتدا، یک‌لایه سیگموید داریم که تصمیم می‌گیرد چه بخشی از سلول حالت قرار است به خروجی برده شود. سپس مقدار سلول حالت (پس از به‌روز شدن در مراحل قبلی) را به یک‌لایه تانژانت هایپر بولیک (تا مقادیر بین ۱- و ۱+ باشند) می‌دهیم و مقدار آن را در لایه خروجی[1] سیگموید قبلی ضرب می‌کنیم تا تنها بخش‌هایی که مدنظرمان است به خروجی برود. شکل زیر تکمیل‌کننده توضیحات ارائه‌شده است:



$$o_t = \sigma\left(W_o\ [h_{t-1}, x_t]\ +\ b_o\right)$$
$$h_t = o_t * \tanh\left(C_t\right)$$

شکل ۸: دروازه خروجی LSTM

خطای پیش‌بینی

در علم آمار، دقت پیش‌بینی به درجه نزدیکی مقدار پیش‌بینی‌شده توسط مدل به مقدار واقعی تعریف می‌شود. هرچقدر خطای پیش‌بینی مدل پایین‌تر باشد، دقت آن بیشتر است. لذا به‌منظور محاسبه دقت پیش‌بینی، می‌توان از معیارهای محاسبه خطا استفاده نمود. در این پژوهش از چهار معیار میانگین مربعات خطا[2]، جذر میانگین مربعات خطا[3]، میانگین درصد خطای مطلق[4] و میانگین قدر مطلق خطا[5] که کاربرد گسترده‌ای در محاسبه خطای پیش‌بینی مدل‌ها دارند، استفاده‌شده است. در ادامه به توضیح و ارائه نحوه محاسبه هر یک از این دو معیار پرداخته‌شده است.

میانگین مربعات خطا

معیاری برای محاسبه مقدار خطای پیش‌بینی یک مدل و درنتیجه دقت آن مدل در پیش‌بینی است. این معیار همواره مقداری مثبت دارد و هرچه مقدار آن به صفر نزدیک‌تر باشد، مدل دقت بیشتری دارد. (والچ و گوفینت[6]،۱۹۸۹) فرمول آن به‌صورت زیر است:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(f_i - y_i)^2 \qquad (۱)$$

[1] - Output Gate
[2] - Mean Squared Error (MSE)
[3] - Root Mean Squared Error (RMSE)
[4] - Mean Absolute Percentage Error (MAPE)
[5] - Mean Absolute Deviation
[6] Wallach and Goffinet, 1989

N= تعداد نقاط پیش‌بینی

f_i= مقدار پیش‌بینی‌شده توسط مدل

y_i= مقدار واقعی

جذر میانگین مربعات خطا

این معیار جذر معیار قبلی که همان میانگین مربعات خطا بود، است. این معیار همواره مقداری مثبت دارد و هرچه مقدار آن به صفر نزدیک‌تر باشد، مدل دقت بیشتری دارد. فرمول آن به‌صورت زیر است:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - y_i)^2} \quad (۲)$$

N= تعداد نقاط پیش‌بینی

f_i= مقدار پیش‌بینی‌شده توسط مدل

y_i= مقدار واقعی

میانگین درصد خطای مطلق

معیاری برای سنجش دقت پیش‌بینی مدل (به‌خصوص مدل‌های پیش‌بینی روند) در آمار است. (De Myttenaere et al., 2016) فرمول آن به‌صورت زیر است:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - f_i}{y_i}\right| \quad (۳)$$

N= تعداد نقاط پیش‌بینی

f_i= مقدار پیش‌بینی‌شده توسط مدل

y_i= مقدار واقعی

میانگین قدر مطلق خطا

$$MAD = \frac{1}{N}\sum_{i=1}^{N}|y_i - f_i| \quad (۴)$$

این معیار نیز معیاری برای سنجش دقت پیش‌بینی مدل در آمار است و فرمول آن به‌صورت زیر است:

N = تعداد نقاط پیش‌بینی

f_i= مقدار پیش‌بینی‌شده توسط مدل

y_i= مقدار واقعی

یافته‌ها

به‌منظور توصیف داده‌های خام قیمت تعدیل‌شده هر یک از نمادهای مورد بررسی در پژوهش، جدول (۱) که شامل آمار توصیفی است، آماده گردید.

جدول۱: آمار توصیفی

نماد	تعداد	حداکثر	حداقل	میانگین	انحراف معیار	دامنه	چارک اول	چارک دوم	چارک سوم	چولگی	کشیدگی
فملی	۲۹۵۷	۳۷۷۰۰.۰	۵۷.۷	۱۲۷۴.۱	۳۰۰۲.۱	۳۷۶۴۲.۳	۲۳۵.۳	۵۷۴.۳	۸۹۴.۷	۷.۱	۶۲.۵
همراه	۱۶۱۷	۵۲۱۵۰.۰	۱۳۶۸.۴	۶۴۲۰.۶	۶۸۷۲.۶	۵۰۷۸۱.۶	۲۵۵۶.۰	۴۰۲۵.۹	۷۴۸۶.۵	۳.۴	۱۴.۰
حکشتی	۲۵۲۴	۵۸۰۰۰.۰	۳۰۰.۳	۲۸۳۹.۷	۴۵۷۵.۹	۵۷۶۹۹.۷	۶۵۳.۷	۲۳۸۴.۳	۳۴۷۶.۷	۶.۳	۵۱.۴
کچاد	۳۵۳۸	۲۵۳۹۰.۰	۵۵۰.۵	۱۹۶۶.۵	۲۰۷۳.۴	۲۴۸۳۹.۶	۱۰۶۶.۰	۱۵۰۳.۳	۲۱۵۸.۴	۶.۲	۴۹.۸
فارس	۱۶۴۱	۳۲۱۸۰.۰	۹۲۵.۳	۲۵۵۵.۸	۳۲۷۱.۰	۳۱۲۵۴.۸	۱۲۹۲.۴	۱۵۳۶.۰	۲۲۷۸.۴	۴.۸	۲۷.۲
اخابر	۲۵۲۲	۱۷۶۳۰.۰	۲۷۹.۰	۱۵۳۴.۱	۱۸۳۲.۸	۱۷۳۵۱.۰	۸۲۸.۷	۱۰۷۳.۳	۱۳۶۴.۳	۵.۰	۳۰.۵
فخوز	۲۶۲۱	۴۰۷۸۰.۰	۱۱۴.۳	۲۹۳۱.۵	۴۸۷۷.۳	۴۰۶۶۵.۷	۳۶۷.۹	۱۳۷۲.۴	۲۵۹۶.۳	۴.۰	۲۰.۹
کگل	۳۲۱۲	۲۵۳۷۰.۰	۴۶.۲	۱۱۳۲.۹	۲۲۶۱.۴	۲۵۳۲۳.۸	۱۱۱.۲	۴۶۴.۱	۹۴۷.۸	۵.۰	۳۳.۲
فولاد	۲۹۲۲	۲۳۰۲۰.۰	۴۵.۰	۸۴۵.۸	۱۸۶۶.۶	۲۲۹۷۵.۰	۱۲۲.۴	۳۶۴.۳	۵۵۶.۹	۶.۱	۴۷.۸
مبین	۱۱۸۹	۳۵۴۴۰.۰	۱۴۳۸.۱	۵۶۳۱.۹	۵۵۸۷.۹	۳۴۰۰۱.۹	۲۵۲۳.۷	۳۳۵۸.۴	۷۲۳۰.۹	۲.۷	۸.۵
وغدیر	۳۹۸۷	۱۸۵۴۰.۰	۳۵۷.۵	۱۲۸۰.۳	۱۴۸۳.۷	۱۸۱۸۲.۵	۷۵۶.۹	۹۶۴.۲	۱۳۰۸.۶	۷.۰	۶۰.۵
وامید	۲۶۶۶	۴۹۳۹۰.۰	۹۰.۸	۲۰۵۲.۳	۴۴۶۹.۱	۴۹۲۹۹.۲	۳۱۲.۳	۹۵۲.۱	۱۴۴۹.۸	۶.۱	۴۶.۱
وبصادر	۲۰۹۸	۴۴۶۰.۰	۵۸.۱	۲۸۸.۵	۳۷۴.۲	۴۴۰۱.۹	۱۰۶.۲	۲۴۲.۱	۳۲۰.۳	۵.۹	۴۴.۹
تاپیکو	۱۵۳۸	۲۳۰۰۰.۰	۹۵۴.۴	۲۱۲۱.۶	۲۲۷۹.۹	۲۲۰۴۵.۶	۱۱۵۸.۰	۱۳۴۳.۷	۲۲۷۴.۲	۵.۲	۳۲.۸
وبپاسار	۲۹۳۲	۵۱۶۰.۰	۲۹.۲	۲۰۴.۷	۴۰۴.۵	۵۱۳۰.۸	۵۴.۹	۱۴۲.۴	۱۹۲.۳	۷.۷	۷۰.۴
شبندر	۱۵۰۸	۵۳۰۵۰.۰	۲۵۲.۳	۴۱۸۸.۳	۵۱۴۳.۱	۵۲۷۹۷.۷	۱۶۳۳.۶	۲۷۷۸.۰	۴۸۰۸.۳	۴.۵	۲۸.۵
ومعادن	۳۸۵۵	۷۷۵۰.۵	۸.۴	۳۱۸.۱	۶۸۷.۰	۷۷۴۲.۲	۳۱.۰	۹۶.۱	۲۹۷.۸	۵.۵	۳۹.۵
شپنا	۲۱۲۷	۳۸۹۷۰.۰	۳۹۱.۹	۲۱۵۲.۲	۳۲۶۶.۶	۳۸۵۷۸.۱	۶۹۳.۹	۸۸۱.۳	۳۰۲۷.۶	۶.۰	۴۹.۷
پارسان	۱۸۶۳	۴۰۸۰۰.۰	۱۰۷۷.۳	۳۴۶۹.۴	۳۶۸۸.۸	۳۹۷۲۲.۷	۱۴۰۸.۶	۲۹۷۰.۸	۳۸۷۶.۹	۵.۳	۳۷.۲
رمپنا	۲۵۰۶	۵۲۹۸۰.۰	۱۰۸.۲	۳۰۲۷.۳	۵۴۳۴.۵	۵۲۸۷۱.۸	۵۴۰.۳	۲۱۰۶.۴	۲۷۱۶.۵	۴.۷	۲۸.۲

مرحله ۱: پیش‌پردازش داده

تمامی مراحل پیاده‌سازی شبکه عصبی بازگشتی (RNN) حافظه طولانی کوتاه‌مدت (LSTM)، با استفاده از زبان برنامه‌نویسی پایتون و کتابخانه‌های زیر انجام‌شده است:

- Numpy
- Pandas
- Matplotlib
- Sklearn
- Keras

در ادامه هر یک از مراحل پیش‌پردازش داده به تفکیک توضیح داده‌شده است.

بارگذاری مجموعه داده

در گام اول، لازم است تا داده دریافت شده از نرم‌افزار TSECLIENT را با استفاده از زبان برنامه‌نویسی پایتون بارگذاری نمود. به این منظور از دستور read_csv از کتابخانه Pandas استفاده‌شده است.

جداسازی مجموعه آموزشی و آزمایشی

در بسیاری از الگوریتم‌های پیش‌بینی، لازم است تا داده به دو مجموعه آموزشی و آزمایشی تقسیم گردد. داده آموزشی به‌منظور یادگیری مدل، کشف روابط موجود و یافتن مقدار مناسب پارامترهای مدل استفاده می‌شود. داده آزمایشی نیز به‌منظور امتحان کردن توانمندی مدل ساخته‌شده، استفاده می‌شود. به‌منظور تعیین افق پیش‌بینی به ادبیات حوزه مراجعه گردید و مقالات مشابه مبنای تصمیم‌گیری از قرار ونگ و همکارانش[1]، بابو و همکارانش[2] و همچنین اسکانبرگ و همکارانش[3] ازجمله محققینی بودند که پژوهش آنان مبنای کار قرار گرفت و با توجه به آن‌ها افق زمانی پیش‌بینی تعیین گردید. لذا در این پژوهش، بازه زمانی مورد بررسی از هر سهم از تاریخ ۱۳۹۴/۰۲/۱۱ تا ۱۳۹۹/۰۱/۲۲ است که داده ۱۰ روز آخر، به‌عنوان داده آزمایشی و داده ماقبل آن، داده آموزشی در نظر گرفته‌شده است.

تغییر مقیاس ویژگی

در این پژوهش لازم است تا قیمت سهام در طول زمان به بازه‌ی ۰ تا ۱ تغییر مقیاس داده شود. به این منظور در این پژوهش، از تابع MinMaxScaler از کتابخانه Sklearn و ماژول Sklearn.preprocessing در زبان برنامه‌نویسی پایتون استفاده‌شده است که فرمول محاسبه آن به‌صورت زیر است:

X= مقدار اولیه

$$X_{sc} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (۴)$$

X_{sc}= مقدار تغییر مقیاس یافته

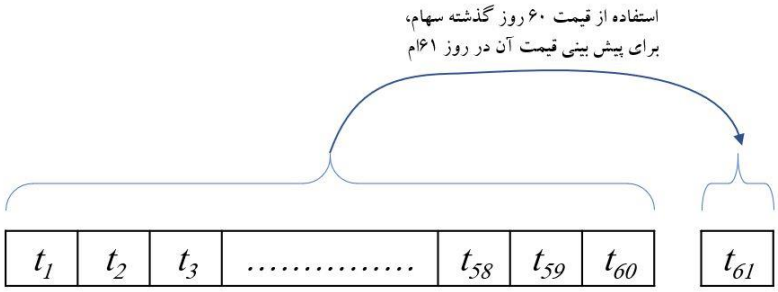
X_{min}= کمترین مقدار

X_{max}= بیشتر مقدار

تهیه ساختار مورد نیاز برای شبکه

در ابتدا لازم است تا داده را به ساختاری مشابه شکل (۹) تبدیل نمود:

[1] Weng et al.,2018
[2] Babu and Reddy,2015
[3] Schöneburg,1990



شکل ۹:گام زمانی

داده ورودی به تابع LSTM از کتابخانه Keras، می‌بایست در قالب مشخصی (یک آرایه سه‌بُعدی) باشد. لذا در گام دوم داده با استفاده از تابع reshape ازکتابخانه NumPyبه یک آرایه سه‌بعدی، با نمونه‌های داده آموزشی، تعداد ۶۰ گام زمانی و یک ویژگی در هر گام تبدیل شدند.

مرحله ۲: پیاده‌سازی الگوریتم‌ها

واردکردن توابع مورد نیاز از کتابخانه Keras

- Sequential= این مدل، دسته‌ای خطی از لایه‌ها است و برای مقداردهی اولیه شبکه عصبی استفاده می‌شود.
- Dense= برای افزودن یک‌لایه به‌شدت متصل شبکه عصبی[1] که برای دریافت خروجی استفاده می‌شود.
- LSTM= برای افزودن لایه شبکه حافظه طولانی کوتاه-مدت
- Dropout= لایه‌ای که برای جلوگیری از بیش برازش[2] مدل نسبت به داده آموزشی استفاده می‌شود. به این صورت که هر بار به‌صورت تصادفی درصدی از رئوس[3] شبکه را به همراه یال‌های متصل به آن‌ها حذف می‌کند تا مدل بیش‌ازحد به داده آموزشی برازش نشود.

ساخت شبکه LSTM

پس از تعریف لایه‌های مذکور، یک مدل پیش‌بینی، با استفاده از بهینه‌ساز محبوب Adam و تابع هزینه میانگین مربعات خطا ساخته می‌شود. سپس، مدل برای اجرا روی ۱۰۰ دوره (تعداد دفعاتی که مدل باید اجرا شود تا پارامترهای شبکه تنظیم شوند) و اندازه دسته ۳۲ برازش می‌شود و در هر دوره مدام میانگین مربعات خطا کاهش می‌یابد.

[1] - Densely-Connected Neural Network
[2] - Overfitting
[3] - Units

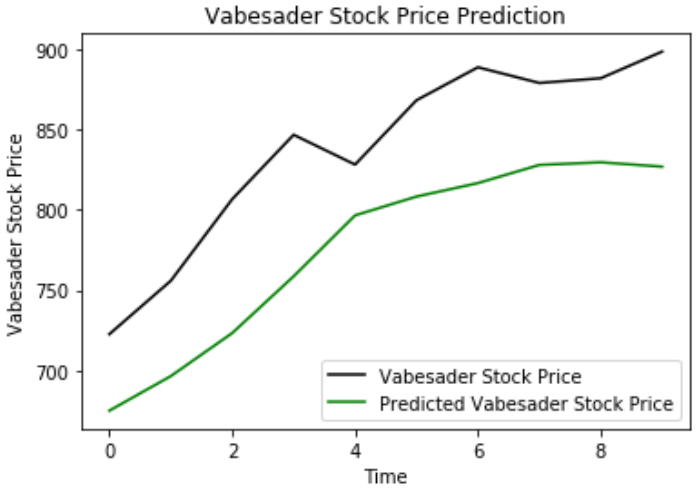
مرحله ۳: پیش‌بینی

به‌منظور پیش‌بینی قیمت‌های آینده سهام، می‌بایست مجموعه داده آزمایشی را بارگذاری و سپس چند مرحله زیر طی شود:

۱) اضافه کردن ۶۰ داده آخر از مجموعه داده آموزشی به ابتدای مجموعه داده آزمایشی
۲) تغییرمقیاس ویژگی مجموعه داده آزمایشی
۳) تغییر ساختار داده آزمایشی به قالب ورودی موردنیاز شبکه
۴) پیاده‌سازی مدل پیش‌بینی بر روی مجموعه داده آزمایشی
۵) تغییر مقیاس مجدد مقادیر پیش‌بینی‌شده توسط مدل، به مقیاس واقعی

مرحله ۴: مصورسازی

با استفاده از تابع pyplot از کتابخانه matplotlib در زبان برنامه‌نویسی پایتون، به‌رسم روند واقعی قیمت سهام و روند پیش‌بینی‌شده توسط مدل پرداخته‌شده است. نمودار زیر خروجی حاصل از پیاده‌سازی این الگوریتم بر روی یکی از نمادها است و نمودار سایر نمادها نیز در پیوست آمده است. در این نمودار محور افقی زمان است که از ۱ الی ۱۰ روز شماره‌گذاری شده است. محور عمودی نیز قیمت است. نمودار مشکی رنگ مقدار واقعی و نمودار سبزرنگ قیمت پیش‌بینی‌شده نماد مورد بررسی است.



شکل۱۰: خروجی LSTM برای نماد وبصادر

جدول ۲:خطاهای پیش‌بینی

	LSTM			
نماد	MAD	MSE	RMSE	MAPE
مبین	688.5	612720.6	782.8	4.7
وبصادر	61.7	4071.5	63.8	7.4
وپاسار	118.9	16966.5	130.3	14.4
وامید	418.2	240014.6	489.9	3.4
تاپیکو	451.2	273447.7	522.9	10.1
شبندر	265.6	93325.8	305.5	2.8
فارس	265.3	91193.6	302.0	3.4
همراه	2673.4	7934902.1	2816.9	12.4
پارسان	317.7	129189.5	359.4	4.0
شپنا	212.2	59913.6	244.8	3.9
اخابر	627.1	420803.8	648.7	8.1
حکشتی	965.6	1284929.0	1133.5	4.7
رمپنا	2787.7	8362633.8	2891.8	15.9
فخوز	386.6	198325.1	445.3	2.9
فولاد	157.6	45185.2	212.6	3.3
کگل	660.5	480266.3	693.0	9.8
کچاد	204.4	63980.5	252.9	3.0
ومعادن	318.6	126015.1	355.0	12.0
وغدیر	194.5	54925.8	234.4	4.2
فملی	265.2	92595.1	304.3	3.8

پس از محاسبه انواع خطا برای روش پیش‌بینی، لازم بود تا از لحاظ آماری به بررسی میانگین و واریانس انواع خطاها در پیش‌بینی قیمت ۲۰ سهام پرداخته شود. پس از انجام محاسبات مربوط به میانیگین و انحراف معیار انواع خطاهای پیش‌بینی نمی‌توان فرض صفر مبنی بر صفر بودن میانیگین خطای پیش بینی هر یک از معیارهای زیر رد نمود. در نتیجه می‌توان گفت که قدرت پیش‌بینی روش LSTM مناسب است. در جدول زیر میانگین و انحراف معیار خطاهای پیش‌بینی محاسبه‌شده است و نتایج مذکور قابل‌مشاهده است.

جدول ۳:میانگین و انحراف معیار خطاهای پیش‌بینی

	LSTM			
	MAD	MSE	RMSE	MAPE
میانگین	۶۰۲	۱۰۲۹۲۷۰	۶۵۹.۵	۶.۷
انحراف معیار	۷۴۲	۲۳۹۱۵۰۰	۷۷۰.۹	۴.۱

نتیجه‌گیری

در این پژوهش به توضیح نحوه آماده‌سازی داده‌ها، ساخت مدل و پیش‌بینی قیمت آینده سهام ۲۰ شرکت برتر بازار بورس، توسط روش شبکه عصبی بازگشتی (RNN) حافظه طولانی کوتاه‌مدت (LSTM) پرداخته شد. پس از پیاده‌سازی روش پیش‌بینی بر روی هر یک از ۲۰ نماد معرفی‌شده در فصول گذشته، به محاسبه ۴ نوع خطای پیش‌بینی پرداخته شد. پس از محاسبه انواع خطا ، لازم بود تا از لحاظ آماری به بررسی میانگین و واریانس خطاهای متناظر با روش در پیش‌بینی قیمت ۲۰ سهام پرداخته شود. پس‌ازانجام محاسبات نتایج حاصل حاکی از آن بود که در پیش‌بینی قیمت آینده ۲۰ نماد، از لحاظ آماری فرض صفر بودن مبنی بر وجود خطای پیش بینی در روش LSTM را نمی‌توان رد نمود. و لذا می‌توان دریافت که قدرت پیش‌بینی روش LSTM مناسب است.

به عنوان یک نتیجه نهایی می‌توان بیان کرد که همانند سایر مطالعات پیشین (نظیر ذوالفقاری و همکاران، (۱۳۹۹)؛ سیدرا و همکاران[1] (۲۰۲۰)؛ کومار[2] و همکاران(۲۰۲۱)؛ پاندی کومار و همکاران[3] (۲۰۲۲)) مدل‌های ترکیبی مبتنی بر وجود حافظه طولانی کوتاه‌مدت نسبت به سایر مدلهای سری زمانی توانایی بهتری جهت پیش‌بینی داده‌های مالی را دارد. شایان ذکر است که به دلیل عدم نرمال بودن بازدهی سهام، توزیع مدلها بر اساس توزیع t نتایج بهتری نسبت به توزیع نرمال ارائه دادند. درنهایت با توجه به اینکه پیش‌بینی تغییرپذیری یا نوسان در مدیریت ریسک، ارزشگذاری سبدسرمایه، قیمت‌گذاری مشتقات و ... کاربردهای فراوانی دارد؛ و ازدیدگاه معامله‌گران بازار مشتقات، درک نوسان پذیری، پیش‌بینی دقیق آن و حفاظت از سبد دارایی‌های در مقابل هزینه‌هایی که این متغیر به ارزش کل تحمیل می‌کند، از اهمیت دوچندانی برخوردار است، استفاده از چنین ساختاری که تقریبا تمامی عوامل اثرگذار (اخبار سیاسی، شوک‌های داخلی و خارجی و ...) در این شاخص‌ها را در نظر می‌گیرد، می‌تواند مفید و کاربردی باشد .ارتباط تحقیق حاضر با تحقیقات پیشین از این جهت است که در تحقیقات گذشته غالب مدل‌های اقتصادسنجی مورد استفاده قرارگرفته است و تفاوت تحقیق حاضر با تحقیقات قبل در این می‌باشد که در این تحقیق مدلسازی بصورت ترکیبی انجام شده است که البته نتایج نیز نسبت به تحقیقات قبل بهبود را نشان می‌دهد. براساس نتایج و مطالب این مطالعه و همچنین بررسی تحقیقات گذشته، پیشنهادات زیر در خصوص پژوهش و تحقیقات آتی ارایه می‌شود: استفاده از ترکیب مدلهای خانواده اقتصادسنجی و الگوریتم‌های فراابتکاری و همچنین مدل‌های توسعه یافته خانواده GARCH نظیر FIGARCH و FIEGARCH و استفاده از مدل ترکیبی معرفی شده در پژوهش حاضر جهت پیش‌بینی سایر داده‌های اقتصادی و مالی نظیر نرخ ارز، شاخص صنعت و قیمت سهام و نرخ سکه.

در این پژوهش نیز مانند تمامی پژوهش‌ها، محدودیت‌هایی وجود داشت که در ادامه به برخی از آن‌ها اشاره‌شده است. در برخی از شرکت‌ها، به دلیل افزایش سرمایه‌های سنگین، قیمت سهم‌ها بسیار کاهش پیدا می‌کند. لذا در هنگام باز شدن سهم، انتظارات تورمی، رقیق شدن قیمت سهم و شرایط بازار باعث می‌شوند که تقاضا افزایش پیدا کند. از طرفی با توجه به آنکه پس از افزایش سرمایه و تقسیم سود نقدی، سهم‌ها بدون دامنه نوسان بازگشایی

[1] Sidra et al
[2] kumar
[3] Pandikumar

می‌شوند لذا گپ نسبتاً بزرگ قیمتی ایجاد می‌شود. با توجه به آنکه شبکه عصبی بازگشتی LSTM دارای حافظه‌ی بلندمدت است، لذا وجود این گپ‌ها باعث کاهش دقت پیش‌بینی می‌شوند.

فهرست منابع

* احمدخان بیگی, سهیل, عبدالوند, ندا. (۱۳۹۶). پیش‌بینی قیمت سهام با رویکرد ترکیبی شبکه عصبی مصنوعی و الگوریتم رقابت استعماری مبتنی بر تئوری آشوب .راهبرد مدیریت مالی۷۳-۲۷، (۳) ۵.
* بیگی، شاهرخ (۱۳۹۸). بررسی تعاملات نرخ ارز و بازار سهام، رویکرد سیستم معادلات هم‌زمان، پایان‌نامه کارشناسی‌ارشد، دانشگاه خوارزمی تهران.
* راعی،رضا ؛ علیدوست اقدم، محمد ؛ فیضی، اصغر (۱۳۹۱)، شبکه‌های عصبی در امور مالی، تهران: انتشارات نگاه دانش.
* جورابیان، محمود (۱۳۸۶)،شبکه‌های عصبی مصنوعی، انتشارات دانشگاه شهیدچمران.
* رمضانیان کشتلی، مریم و مومنی کشتلی، رضا،۱۳۹۸،کاربرد الگوریتم های یادگیری ماشین در پیش بینی تصمیمات سیستم های قضایی،پنجمین کنفرانس ملی علوم انسانی و مطالعات مدیریت،تهران،
* باباجانی، جعفر و تقوا، محمدرضا و بولو، قاسم و عبدالهی، محسن (۱۳۹۸). پیش‌بینی قیمت سهام در بورس تهران با استفاده از شبکه عصبی بازگشتی بهینه شده با الگوریتم کلونی زنبور عسل مصنوعی. راهبرد مدیریت مالی تابستان ۱۳۹۸ - شماره ۲۵ علمی-پژوهشی (وزارت علوم) ISC (۳۵ صفحه - از ۱۹۵ تا ۲۲۹)
* بیات, علی, باقری, زینب. (۱۳۹۶). پیش‌بینی قیمت سهام با استفاده از الگوریتم کرم شب‌تاب (FA). دانش مالی تحلیل اوراق بهادار، ۱۴۵-۱۳۵ ، (۳۵) ۱۰.
* حیدری زارع، بهزاد و کردلویی، حمیدرضا (۱۳۸۹). پیش‌بینی قیمت سهام با استفاده از شبکه عصبی مصنوعی. پژوهشگر (مدیریت) (Journal of Industrial Strategic Management) صفحه ۴۹-۵۶.
* منجمی، سید امیرحسین و ابزری، مهدی و رعیتی شوازی، علیرضا (۱۳۸۸). پیش‌بینی قیمت سهام در بازار بورس اوراق بهادار با استفاده از شبکه عصبی فازی و الگوریتم‌های ژنتیک و مقایسه آن با شبکه عصبی مصنوعی. اقتصاد مقداری (بررسی های اقتصادی)
* نصیر زاده، فرزانه (۱۳۹۱). ارزیابی توانایی مدل‌های داده‌کاوی در پیش‌بینی قیمت سهام. یازدهمین همایش ملی حسابداری ایران
* ARIYO, A. A., ADEWUMI, A. O. & AYO, C. K. Stock price prediction using the ARIMA model. 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, 2014. IEEE, 106-112.
* Anderson, D., & McNeill, G. (1992). Artificial neural networks technology. Kaman Sciences Corporation, 258(6), 1-83.
* Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. Expert Systems with Applications, 113, 457-480.

* Das, S. R., Mishra, D., & Rout, M. (2019). Stock market prediction using Firefly algorithm with evolutionary framework optimized feature reduction for OSELM method. Expert Systems with Applications: X, 4, 100016.
* HAJIZADEH, E., SEIFI, A., ZARANDI, M. F. & TURKSEN, I. 2012. A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, 39**,** 431-436.
* KHUAT, T. T. & LE, M. H. 2017. An application of artificial neural networks and fuzzy logic on the stock price prediction problem. *JOIV: International Journal on Informatics Visualization,* 1**,** 40-49.
* SELVIN, S., VINAYAKUMAR, R., GOPALAKRISHNAN, E., MENON, V. K. & SOMAN, K. Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017 international conference on advances in computing, communications and informatics (icacci), 2017. IEEE, 1643-1647.
* SHILLER, R. J. 1987. The volatility of stock market prices. Science, 235, 33-37.
* ZAREMBA, W., SUTSKEVER, I. & VINYALS, O. 2014. Recurrent neural network regularization. arXiv preprint arXiv:1409.2329.
* Moghar, A., & Hamiche, M. (2020). Stock Market Prediction Using LSTM Recurrent Neural Network. Procedia computer science, 170, 1168-1173.
* Zhang, X., Chen, Y., & Yang, J. Y. (2007). Stock Index Forecasting Using PSO Based Selective Neural Network Ensemble. Paper presented at the IC-AI.
* Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K. and Soman, K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. Proc. Int. Conf. Advances in Computing, Communications and Informatics (ICACCI), Udupi, India.
* S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
* Arnaud de Myttenaere, Boris Golden, Bénédicte Le Grand, Fabrice Rossi, Mean Absolute Percentage Error for regression models, Neurocomputing, Volume 192, 2016, Pages 38-48, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2015.12.114.
* Babu AS, Reddy SK (2015) Exchange Rate Forecasting using ARIMA, Neural Network and Fuzzy Neuron. J Stock Forex Trad 4:155. doi:10.4172/2168-9458.1000155
* Markowitz, Harry,(1959). Portfolio selection: efficient diversification of investments, Wiley, New York, NY ,x,344pp.
* H. Zhang, T.-W. Weng, P.-Y. Chen, C.-J. Hsieh, and L. Daniel. Efficient neural network robustness certification with general activation functions. Advances in Neural Information Processing Systems, 2018.
* E. Schöneburg, Stock price prediction using neural networks: A project report, Neurocomputing, Volume 2, Issue 1, 1990, Pages 17-27, ISSN 0925-2312, https://doi.org/10.1016/0925-2312(90)90013-H.
* Zolfaghari, M., & Sahabi, B. (2017). Impact of foreign exchange rate on oil companies risk in stock market: A Markov-switching approach. Journal of Computational and Applied Mathematics, 317, 274-289.
* R. Kumar, P. Kumar and Y. Kumar, "Analysis of Financial Time Series Forecasting using Deep Learning Model," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021, pp. 877-881, doi: 10.1109/Confluence51648.2021.9377158.
* Sidra, M., , Jaydip, S., & Abhishek, D.(2020). Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. International Journal of Business Forecasting and Marketing Intelligence
* Zhang, K., Zhong, G., Dong, J., Wang, S., & Wang, Y. (2019). Stock market prediction based on generative adversarial network. Procedia computer science, 147, 400-406

* Pandikumar, S., Sethupandian, S. B., Saravanan, M. S., Prasad, S. N., & Arun, M. (2022). Deep Learning based Long Short-Term Memory Recurrent Neural Network for Stock Price Movement Prediction. *Indian Journal of Science and Technology*, *15*(11), 474-480.

Financial Economics
Vol. (14) Issue (55) Autumn 2024

Evaluation of the performance of recursive neural network (RNN) Long-Short Term Memory (LSTM) method in stock price forecasting

Hossein Ameri Shahrabi
M.A. Graduated of kharazmi University, Department of financial Engineering, Tehran, Iran
hosseinameri1995@gmail.com

Ehsan Taieby sani
Associate professor of kharazmi University, Department of financial management, Tehran, Iran
(Correspondent author)
Ehsan_taieby_sani@khu.ac.ir

Abstract
Forecasting is one of the most important and challenging issues in economic and financial affairs that has always been considered by many people and researchers and is of great importance. There have always been two traditional and intelligent approaches to forecasting in financial markets. In the traditional method, this forecasting is based on statistical models and in the intelligent method, it is based on artificial intelligence models. In this research, due to the importance of investing and especially investing in the stock market, we predicted the stock price return on the stock exchange through Recursive Neural Network (RNN) of long-short term memory (LSTM). The research variable is the stock price of the top 20 (in market cap) companies on the stock exchange for the period of the 11th Feb 2015 to 22th Jan 2022. We considered the data of the last 10 days as experimental data and the previous data as educational data. Initially, we calculated the mean and standard deviation of the prediction error of both models; The mean and standard deviation of the predictive error of the recursive neural network (LSTM) model indicate the statistical significance of the model.
Key words: abnormalities, price gaps, patterns, Heteroscedasticity.