فصلنامه مهندسی مدیریت نوین
سال دهم، شماره سوم، پاییز ۱۴۰۳

پیش‌بینی‌پذیری بورس اوراق بهادار تهران با استفاده از مدل‌های یادگیری عمیق (مدل هیبریدی CNN-LSTM)

مهدی حیدرزاده[۱]، مژگان صفا[۲]، میرفیض فلاح شمس[۳]، حسین جهانگیرنیا[۴]

چکیده

یادگیری عمیق، زیرمجموعه‌ای از کلاس گسترده‌تر از روش‌های یادگیری ماشین مبتنی شبکه‌های عصبی است که اخیراً در حوزه‌های مختلفی ازجمله پیش‌بینی سری‌های زمانی در بازارهای مالی، توجهات زیادی را به خود جلب کرده است. در این تحقیق، ابتدا بر اساس مدل-های یادگیری عمیق مبتنی بر شبکه‌های LSTM و CNN حرکت شاخص بورس اوراق بهادار تهران پیش‌بینی می‌شود. در ادامه با ترکیب دو مدل مذکور، مدل هیبریدی یادگیری عمیق CNN-LSTM به‌منظور پیش‌بینی شاخص بورس اوراق بهادار تهران مورد استفاده قرار گرفت. در مرحلهٔ بعد، به‌منظور ارزیابی عملکرد مدل‌های پیش‌بینی مذکور، سه معیار سنجش کارایی میانگین درصد قدر مطلق خطای متقارن (SMAPE)، میانگین مطلق درصد خطا (MAPE) و ریشه میانگین مربع خطا (RMSE) مورد استفاده قرار گرفت. در این تحقیق از داده‌های روزانه شاخص بورس اوراق بهادار تهران در دورهٔ زمانی ۱۳۹۵/۴/۲۳ – ۱۴۰۰/۱/۲۶ استفاده شده است. نتایج برآورد مدل‌ها در پیش‌بینی شاخص بورس اوراق بهادار تهران با گام یک‌روزه و مقایسه معیارهای سنجش کارایی، حاکی از برتری عملکرد مدل پیشنهادی CNN-LSTM در مقایسه با دو مدل دیگر است. مدل LSTM در رتبه بعدی دقت و کارایی پیش‌بینی قرار می‌گیرد. با توجه به نتایج ارائه‌شده در این تحقیق، به فعالین بازارهای مالی در ایران پیشنهاد می‌گردد مدل‌های

---

[۱] دانشجوی دکتری گروه مدیریت، واحد قم، دانشگاه آزاد اسلامی، قم، ایران.

[۲] استادیار گروه حسابداری، واحد قم، دانشگاه آزاد اسلامی، قم، ایران، (نویسنده مسئول).

Mozhgan_safa@yahoo.com

[۳] دانشیار گروه مدیریت، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران.

[۴] استادیار گروه حسابداری، واحد قم دانشگاه آزاد اسلامی، قم، ایران.

تاریخ وصول ۱۴۰۲/۱۱/۲۵ **تاریخ پذیرش ۱۴۰۳/۳/۹**

تلفیقی یادگیری عمیق را به‌منظور افزایش کارایی و دقت پیش‌بینی‌های خود مورد توجه قرار دهند.

کلیدواژگان: پیش‌بینی‌پذیری، مدل‌های یادگیری عمیق (Deep Learning)، بورس اوراق بهادار تهران.

طبقه‌بندی JEL : G17, D10, L10

۱- مقدمه

تأثیر بازار بورس در تأمین مالی و توسعهٔ کشور بر کسی پوشیده نیست؛ بنابراین یافتن روشی مناسب برای پیش‌بینی بازار سهام اهمیت بالایی یافته است. شوک‌های ناگهانی بازار، ریزش و سقوط قیمت‌ها باعث می‌شود تعداد زیادی از سرمایه‌گذاران بازار را ترک کنند. افزایش ابزارهای کارای مرتبط با متغیرهای مالی، دامنه فرصت‌های سرمایه‌گذاری جهانی را برای فعالین بازارهای مالی گسترش می‌دهد. دلایل اصلی برای پیشرفت این ابزارها را می‌توان چنین بیان نمود؛ اولاً، این ابزارها، روش‌های مؤثری برای سرمایه‌گذاران به‌منظور محافظت از ریسک‌های بالقوه بازار فراهم می‌آورند. دوماً، فرصت‌های کسب سود را نیز برای کسانی که از موقعیت‌های بازار استفاده می‌کنند، فراهم می‌کنند؛ بنابراین ارائهٔ مدل‌های کارا برای پیش‌بینی شاخص‌ها و متغیرهای بورس از اهمیت و کاربرد خاصی برخوردار است. متغیرهای مالی و اقتصادی (مخصوصاً در بازار سهام) عمدتاً با روابط غیرخطی سروکار داشته؛ بنابراین پیش‌بینی وضعیت آتی بازار سهام با استفاده از مدل‌های مرسوم خطی قابل اتکا نیست.

پیش‌بینی‌پذیری بازارهای مالی از سابقهٔ طولانی مطالعاتی و تجربی برخوردار است (Ang & Bekaert, 2007). علی‌رغم اختلاف‌نظر در مورد کارایی بازارها، مطالعات نشان‌دهندهٔ این است که بازارهای مالی تا حدودی دارای ویژگی پیش‌بینی‌پذیری است (Bollerslev et al, 2014). در بین رویکردهای پیش‌بینی روند آتی سهام، روش‌های اقتصادسنجی یا آماری مبتنی بر تجزیه‌وتحلیل متغیرهای بازار بر اساس داده‌های تاریخی، به‌طور گسترده مورد استفاده فعالان و متخصصین قرار گرفته است. در این رویکردها، روش‌های مختلف خطی و غیرخطی برای پیش‌بینی وضعیت آتی بازار سهام استفاده می‌شود. در این روش‌ها، فرض نرمال بودن تابع چگالی و همچنین مستقل و یکسان

بودن تابع توزیع حیاتی است؛ بااین‌حال شواهدی از ویژگی‌هایی مانند دنبالهٔ پهن و قلهٔ بلند، بازگشت به میانگین واریانس ناهمسان، خوشه‌ای بودن تلاطم‌ها، اثر اهرمی و حافظهٔ بلندمدت نشان می‌دهد که این فرض‌ها در مورد سری‌های زمانی بازده سهام و متغیرهای مالی برقرار نیستند (Vadiei & Hosseini, 2012).

از اواسط دههٔ ۱۹۷۰ تلاش‌های گسترده‌ای در زمینهٔ قابلیت پیش‌بینی قیمت‌های سهام با استفاده از روش‌های ریاضی جدید، سری زمانی طولانی و ابزار پیشرفته‌تری مثل هوش مصنوعی آغاز شد و آزمون‌های زیادی بر روی اطلاعات قیمت و شاخص سهام در کشورهایی مثل انگلستان، آمریکا، کانادا، آلمان و ژاپن صورت گرفت تا وجود یا فقدان ساختاری معین در اطلاعات قیمت سهام نشان داده شود و از این راه فرضیهٔ گام‌های تصادفی را نقض کنند (Kara and Baykan, 2011).

یادگیری ماشینی (ML) زیرمجموعه‌ای از هوش مصنوعی است که با استفاده از الگوریتم‌های مختلف، عملکرد خود بر روی موضوعی خاص را به‌تدریج بهبود می‌دهد. یادگیری ماشینی برای کشف الگوها و جستجوی تغییرات کوچک، بر اساس بررسی و مقایسهٔ داده‌هایی از مقادیر کوچک تا حجم‌های عظیم داده‌ها استوار است. مطالعات فراوانی انجام شده است که نشان می‌دهد پیش‌بینی‌ها براساس مدل‌های یادگیری عمیق ML عملکرد تقریباً بهتری نسبت به تکنیک‌های پیش‌بینی سری زمانی کلاسیک دارند. در همین حال، استفادهٔ گسترده از سیستم‌های تجارت الکترونیکی خودکار همراه با افزایش تقاضا برای بازده بالاتر، محققان و دست‌اندرکاران را مجبور به ادامهٔ کار به‌منظور دستیابی به مدل‌های بهتر کرد. در چند دههٔ اخیر، یکی از زیرشاخه‌های یادگیری ماشینی، به نام یادگیری عمیق (DL)، در پیش‌بینی‌های مالی مورد استقبال قرار گرفته است. هرچند یادگیری عمیق در ابتدای مسیر توسعه است، اما تحقیقات و سرمایه‌گذاری‌های شرکت‌های بزرگ در این حوزه، گسترش روزافزون کاربردهای یادگیری عمیق را نشان می‌دهد. هم‌اکنون یادگیری عمیق در زمینه‌ها و علوم مختلف داده‌کاوی، پردازش تصویر و صدا، رباتیک و پزشکی کاربردهای زیادی داشته است. بر اساس پیش‌بینی‌های مراکز علمی، در سال‌های آتی، یادگیری عمیق در بسیاری از زمینه‌ها، به‌صورت مستقیم یا غیرمستقیم استفاده خواهند کرد.

با توجه به اینکه روش یادگیری عمیق، ظرفیت‌ها و توانایی قابل‌توجهی در استخراج اطلاعات معتبر از مجموعه داده‌ها و الگوها دارد، بسیاری از تحقیقات اخیر بر کاربرد الگوریتم‌های DP در حوزهٔ مالی تمرکز کرده‌اند و ادغام الگوریتم‌های یادگیری عمیق با پیش‌بینی‌های بازار به‌عنوان یکی از مباحث جذاب در ادبیات مالی است (Cavalcante et al, 2016).

در این تحقیق تلاش می‌شود مدل هیبریدی یادگیری عمیق CNN-LSTM به‌منظور پیش‌بینی شاخص بورس اوراق بهادار تهران معرفی شده و کارایی این مدل با ساختارهای سنتی در این حوزه ازجمله مدل LSTM و مدل CNN مقایسه شود. در این راستا از معیارهای سنجش کارایی میانگین مطلق درصد خطا (MAPE) و ریشه میانگین مربع خطا (RMSE[1]) استفاده می‌شود.

۲- ادبیات تحقیق

همان‌طور که پیشتر بیان گردید یادگیری عمیق (DP) زیرشاخه‌ای از روش‌های یادگیری ماشین بر اساس شبکه‌های عصبی است. شبکه‌های عصبی، الگوریتم‌های محاسباتی با توانایی یادگیری و تعمیم‌پذیری هستند. شبکه‌های عصبی مدل‌های الکترونیکی شبیه‌سازی شده بر اساس ساختار عصبی مغز انسان هستند. واژه «عمیق» در «یادگیری عمیق» به تعداد لایه‌هایی اشاره می‌کند که در ساختار مدل‌ها به‌منظور استخراج ویژگی‌های خاص داده‌های ورودی به کار گرفته می‌شوند.

بر اساس بررسی‌های انجام‌شده در ایران تحقیقی که با استفاده از روش یادگیری عمیق به پیش‌بینی متغیرهای مالی و اقتصادی بپردازد یافت نگردید بااین‌حال در سال‌های اخیر، مطالعات زیادی با استفاده از مدل‌های یادگیری عمیق به آزمون پیش‌بینی‌پذیری متغیرهای بازارهای مالی در کشورهای دیگر پرداخته‌اند ازجمله، لیو[2] (۲۰۱۹) در تحقیقی با استفاده از الگوریتم‌های یادگیری عمیق به پیش‌بینی نوسانات بازدهی شاخص اس‌اند پی ۵۰۰ (S & P 500) پرداخته است. در این مطالعه کارایی مدل‌های یادگیری عمیق و مدل‌های واریانس ناهمسانی شرطی خودرگرسیو تعمیم‌یافته (GARCH) مورد ارزیابی

[1] Root Mean Square Error (RMSE)
[2] Liu

قرار گرفت. نتایج این تحقیق نشان‌دهندهٔ کارایی بالاتر مدل‌های یادگیری عمیق در پیش-بینی نوسانات بازدهی شاخص اس‌اندپی ۵۰۰ است.

کانکالوز و همکاران[1] (۲۰۱۹) در تحقیقی با عنوان کاربرد روش یادگیری عمیق در بازار ارز، به پیش‌بینی نوسانات این بازار بر پایه مدل‌های گوناگون یادگیری عمیق پرداختند.

لانگ و همکاران[2] (۲۰۱۸) با استفاده از دو مدل یادگیری عمیق، تلاش کردند شاخص قیمت سهام در کشور چین را مدل‌سازی و پیش‌بینی نمایند. نتایج نشان داد که عملکرد مدل ترکیبی در پیش‌بینی شاخص سهام از مدل‌های رقیب بالاتر است.

تیپریسیتی[3] (۲۰۱۸) تلاش نمود با استفاده از الگوریتم‌های یادگیری عمیق به پیش‌بینی قیمت سهام بازارهای جهانی بپردازد. در این تحقیق با استفاده از ترکیب یادگیری عمیق و تحلیل اخبار از منابع اینترنتی، مدلی هیبریدی به‌منظور پیش‌بینی قیمت سهام پیشنهاد کرده است.

کراوس و همکاران[4] (۲۰۱۷) با استفاده از مدل‌های یادگیری ماشین، شبکه‌های عصبی عمیق و الگوریتم جنگل تصادفی به پیش‌بینی بازده سهام با گام یک‌روزه برای شاخص S & P500 پرداختند. نتایج این تحقیق حاکی از کارایی الگوریتم جنگل تصادفی در پیش-بینی شاخص اس اند پی ۵۰۰ است.

۳- روش تحقیق

این پژوهش حاضر ازلحاظ هدف کاربردی و ازلحاظ ماهیت و روش علّی است. ازنظر ویژگی و جهت داده‌ها پس رویدادی است. در این تحقیق، حرکت شاخص بورس اوراق بهادار تهران با استفاده از مدل‌های یادگیری عمیق مبتنی بر شبکه‌های LSTM، CNN و همچنین مدل هیبریدی یادگیری عمیق CNN-LSTM پیش‌بینی می‌شود و عملکرد این مدل‌ها، با استفاده از معیارهای سنجش مورد ارزیابی قرار می‌گیرد. در این راستا از داده‌های روزانه شاخص بورس اوراق بهادار تهران در دورهٔ زمانی ۱۳۹۵/۴/۲۳ –

[1] Gonçalves et al.
[2] Long et al.
[3] Tipirisetty.
[4] Krauss et al.

۱۴۰۰/۱/۲۶ استفاده می‌شود. در ادامه به تشریح مدل‌های یادگیری عمیق مورد استفاده در این مطالعه پرداخته می‌شود.

۳-۱ شبکه‌های بازگشتی (RNN)

شبکه‌های عصبی بازگشتی (RNN) به‌منظور شبیه‌سازی رفتار انسان در استفاده از اطلاعات قبلی، توسط متخصصان حوزهٔ یادگیری ماشین پیشنهاد شد. شبکه‌های عصبی بازگشتی متشکل از حلقهٔ بازگشتی هستند و تفاوت اصلی RNN با دیگر ساختارها این است که در این ساختار داده‌های ورودی دورهٔ جاری و دوره‌های گذشته به‌صورت هم‌زمان به‌کار گرفته می‌شوند. در واقع خروجی مدل در دورهٔ جاری به داده‌های دورهٔ قبل نیز وابسته است. ازنظر تئوریک یک شبکهٔ عصبی بازگشتی استاندارد (اگر به‌اندازهٔ کافی بزرگ باشد) باید بتواند دنباله‌هایی با هر پیچیدگی‌ای را تولید کند اما در عمل این شبکه در ذخیره‌سازی اطلاعات مرتبط با داده‌های گذشته به مدت طولانی ناتوان است. این ویژگی، علاوه بر تضعیف قدرت این شبکه در مدل‌سازی ساختارهای بلندمدت، باعث می‌شود تا این نوع از شبکه‌ها در زمان تولید دنباله در معرض ناپایداری قرار بگیرند. مشکلی که به وجود خواهد آمد این است که اگر پیش‌بینی‌های شبکه تنها وابسته به چند ورودی اخیر باشد برای اصلاح و جبران اشتباهات گذشته توسط شبکه، شانس بسیار کمی وجود خواهد داشت. یک راه‌حل برای این مشکل، تزریق نویز به پیش‌بینی-های انجام‌شده توسط شبکه قبل از تغذیه آنها به گام زمانی بعدی است. این راه‌حل به تقویت شبکه در قبال ورودی‌های غیرمنتظره منجر می‌شود. حافظه طولانی کوتاه‌مدت (LSTM[1]) یک ساختار شبکهٔ عصبی بازگشتی است که به‌منظور ذخیره‌سازی و دسترسی بهتر به اطلاعات نسبت به نسخهٔ سنتی آن طراحی گردیده است.

برخلاف شبکهٔ عصبی بازگشتی استاندارد (RNN)، در یک شبکهٔ عصبی بازگشتی (LSTM)، شبکه قادر است نسبت به حفظ حافظه فعلی از طریق گیت‌های معرفی‌شده تصمیم بگیرد. به‌طور مفهومی، اگر یک واحدLSTM ویژگی مهمی در دنباله ورودی در مراحل ابتدایی را تشخیص دهد قادر است این اطلاعات را طی مسیر طولانی انتقال داده و اینگونه وابستگی‌های بلندمدت احتمالی را دریافت و حفظ نماید.

[1]Long Short-Term Memory

شبکه‌های LSTM از واحدهای LSTM تشکیل شده‌اند. واحدهای LSTM با هم ادغام شده و لایه‌های LSTM تشکیل می‌شوند. یک واحد LSTM از سلول‌هایی تشکیل می‌شود که دارای گیت ورودی، گیت خروجی و گیت فراموشی است. معادلات (۱) تا (۵) فرم عمومی شبکهٔ LSTM را ارائه می‌کنند.

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (۱)$$

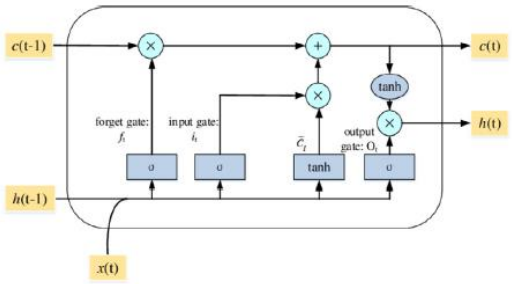
$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (۲)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (۳)$$

$$c_t = f_t * c_{t-1} + i_t * \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (۴)$$

$$h_t = o_t * \sigma_h(c_t) \quad (۵)$$

در اینجا xt، بردار ورودی واحد LSTM، f_t بردار فعال‌سازی گیت فراموشی، i_t بردار فعال‌سازی گیت ورودی، o_t: بردار فعال‌سازی گیت خروجی، h_t بردار خروجی واحد LSTM، c_t بردار حالت سلول، σ_g تابع سیگموئید، σ_cتابع تانژانت هایپربولیک، W، U: ماتریس‌های وزن که باید آموزش داده‌شده و برازش شوند و b پارامترهای بردار بایاس هستند. ساختار یک بلوک از LSTM در شکل (۲) ارائه‌شده است.



شکل (۱) ساختار مدل LSTM

۳-۲ مدل‌های کانولوشنی (CNN[1])

شبکه‌های عصبی کانولوشنی یا پیچشی (CNN) شاخه‌ای از شبکه‌های عصبی عمیق هستند که در پیش‌بینی‌های سری زمانی در یادگیری ماشین استفاده می‌شوند. در شبکه-

[1] Convolutional Neural Network

های عصبی کانولوشنی به‌منظور حداقل‌سازی پیش‌پردازش‌ها از نوعی از پرسپترون‌های چندلایه استفاده می‌شود. به‌جای شبکهٔ عصبی کانولوشنی گاهی از این شبکه‌ها با نام شبکه‌های عصبی تغییرناپذیر با انتقال[1] یا تغییرناپذیر با فضا[2] هم یاد می‌شود. این نام‌گذاری بر مبنای ساختار این شبکه است. ساختار شبکه‌های کانولوشنی از فرایندهای زیستی الگوسازی شده‌اند که در آنها نورون‌ها تنها در یک ناحیهٔ محدود به تحریک، (که به آن ناحیهٔ پذیرش گفته می‌شود) واکنش نشان می‌دهند. نواحی پذیرش نورون‌های مختلف به‌صورت جزئی با هم همپوشانی داشته به‌گونه‌ای که پوشش کل میدان را نتیجه می‌دهند.

ساختار یک شبکهٔ عصبی کانولوشنی (CNN) از لایه‌های مختلفی تشکیل شده است. لایه‌های پنهان کانولوشنی یا ادغامی یا کاملاً متصل هستند. لایه‌های کانولوشنی ورودی وظیفهٔ فیلتر کردن را به عهده دارند، سپس نتیجه را به لایهٔ بعدی منتقل می‌نمایند. ممکن است شبکه‌های عصبی کانولوشنی از لایه‌های ادغام[3] محلی یا سراسری تشکیل شده باشند که خروجی‌های خوشه‌های نورونی در یک لایه را در یک تک نورون در لایهٔ بعدی ادغام می‌کنند. روش حداکثر تجمیع[4] یک مثالی است که حداکثر مقدار بین خوشه‌های نورونی در لایهٔ پیشین را مورد استفاده قرار می‌دهد. میانگین تجمیع[5] که مقدار میانگین خوشه‌های نورونی در لایهٔ پیشین را در نظر می‌گیرد، می‌توان مثالی دیگر در نظر گرفت. شبکه‌های عصبی کانولوشنی با اشتراک وزن‌ها در لایه‌های پیچشی، باعث ایجاد حداقل حافظه و بیشترین کارایی می‌شوند. عملیات کانولوشن (فیلتر) استاندارد در معادلهٔ (۶) نمایش داده شده است. t نشانگر زمان، x نشان‌دهندهٔ ورودی و a نشان‌دهندهٔ متغیر است.

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t-a) \qquad (۶)$$

معادلهٔ (۷) جزئیات مربوط به معماری شبکه را ارائه می‌دهد که در آن W نشانگر وزن، x نشان‌دهندهٔ ورودی، b نشان‌دهندهٔ بایاس و z نشان‌دهندهٔ خروجی نورون‌ها است.

[1] Shift Invariant
[2] Space Invariant
[3] Pooling Layer
[4] Max Pooling
[5] Average Pooling

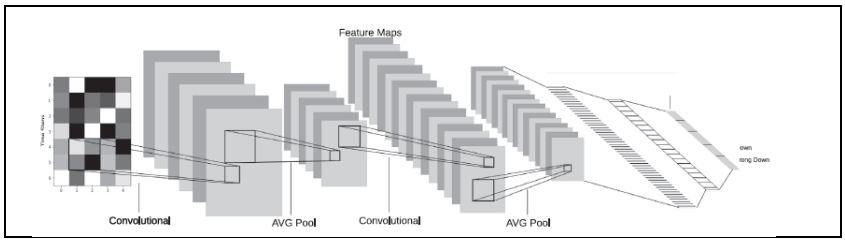
در انتهای شبکه، از تابع softmax برای گرفتن خروجی استفاده می‌شود. معادلهٔ (۸) و (۹) تابع softmax را نشان می‌دهد که در آن y نشان‌دهندهٔ خروجی است.

$$z_i = \sum_j W_{i,j}\, x_j + b_i. \tag{۷}$$

$$y = \mathrm{softmax}(z) \tag{۸}$$

$$\mathrm{softmax}(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_j)} \tag{۹}$$

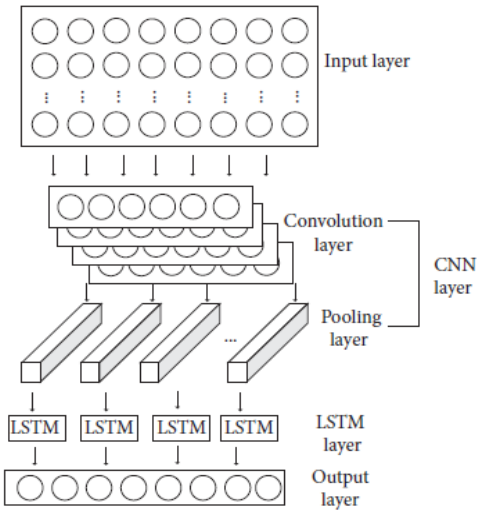
شکل (۱) نمایی کلی از معماری شبکه‌های کانولوشنی در یادگیری عمیق را نشان می‌دهد.



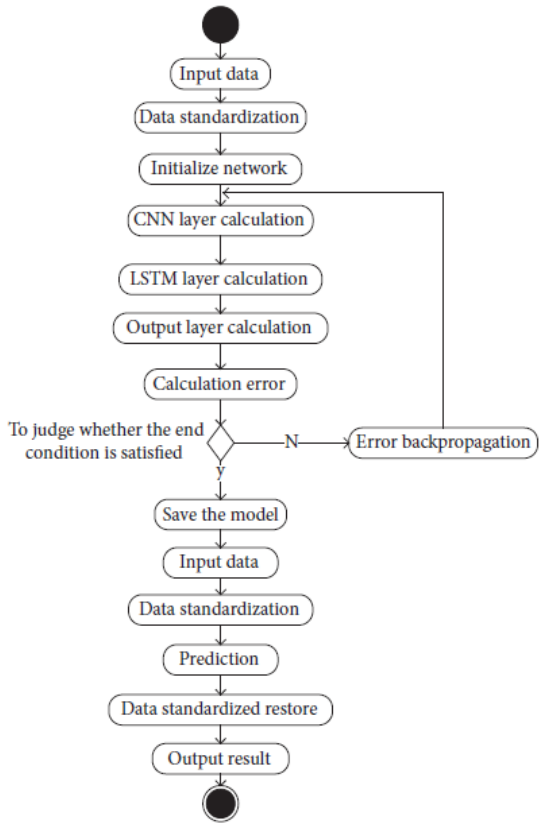
شکل (۲) ساختار شبکه‌های کانولوشنی در یادگیری عمیق

۳-۳ مدل هیبریدی CNN-LSTM

همان‌طور که پیش‌تر بیان شد مدل‌های یادگیری عمیق مبتنی بر شبکه‌های عصبی کانولوشنی یا پیچشی (CNN) در پیش‌بینی‌های سری زمانی در یادگیری ماشین استفاده می‌شوند بااین‌حال بر اساس مطالعات تجربی، این مدل‌ها بر استخراج ویژگی‌ها متمرکز می‌شوند و از طرفی مدل‌های یادگیری عمیق مبتنی بر شبکه‌های LSTM دارای ویژگی تعمیم‌دهی بر اساس توالی زمانی هستند. بر اساس ویژگی‌های دو مدل مذکور، مدل هیبریدی CNN-LSTM به‌منظور پیش‌بینی داده‌های سری زمانی بازارهای مالی به وجود آمده است. شکل (۳)، نمایی کلی از ساختار این مدل را ارائه می‌کند. همان‌طور که در این نگاره قابل مشاهده است این مدل شامل لایهٔ ورودی، بخش لایهٔ کانولوشنی CNN تک‌بعدی، لایهٔ ادغام و بخش لایهٔ پنهان LSTM است. در ادامه و در شکل (۴) مکانیسم آموزش و پیش‌بینی به‌وسیله مدل CNN-LSTM ترسیم شده است.



شکل (٣). ساختار مدل هیبریدی CNN-LSTM



شکل (٤). مکانیسم آموزش و پیش‌بینی با استفاده از مدل یادگیری عمیق هیبریدی CNN-LSTM

۴-۳ معیارهای ارزیابی

برای ارزیابی کارایی روش‌های پیش‌بینی سری زمانی شاخص بورس اوراق بهادار تهران، از سه معیار اصلی استفاده می‌شود:

۱- میانگین درصد قدر مطلق خطای متقارن (SMAPE[1])

۲- میانگین مطلق درصد خطا[2] (MAPE)

۳- ریشه میانگین مربع خطا (RMSE[3])

این معیارها به‌صورت روابط زیر تعریف می‌شوند:

$$SMAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|\hat{y}_t - y_t|}{\frac{|\hat{y}_t|+|y_t|}{2}} \times 100 \qquad (۱۰)$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|\hat{y}_t - y_t|}{|y_t|} \times 100 \qquad (۱۱)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \qquad (۱۲)$$

جاییکه $\hat{y}_t$ و y_t به ترتیب مقادیر واقعی و پیش‌بینی‌شده در زمان t هستند.

۴- نتایج برآورد مدل‌های یادگیری عمیق

در این تحقیق داده‌های شاخص بورس اوراق بهادار تهران به‌عنوان متغیر اصلی تحقیق به‌منظور ورود به پروسه مدل‌سازی و پیش‌بینی با استفاده از مدل‌های یادگیری عمیق (DL) در ابتدا با استفاده از رابطهٔ زیر، نرمال‌سازی می‌شوند:

$$NINDEX = \frac{(INDEX - \mu)}{\sqrt{\sigma^2}} \qquad (۱۳)$$

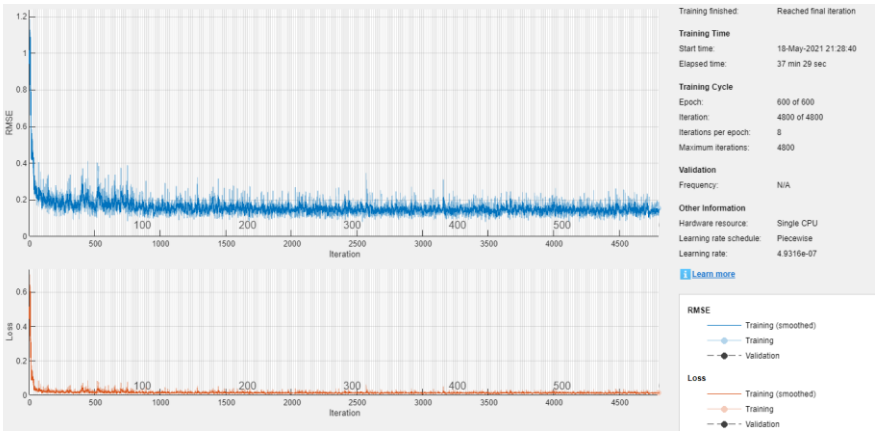
در این رابطه μ و $\sqrt{\sigma^2}$ به ترتیب میانگین و انحراف معیار داده‌های شاخص بورس در دورهٔ مورد بررسی هستند. پس از مرحلهٔ پیش‌پردازش و نرمال‌سازی، داده‌های نرمال-سازی شده به دو بخش داده‌های آموزش و داده‌های آزمون تقسیم می‌شوند. در این تحقیق ۸۰٪ داده‌ها به بخش آموزش تخصیص داده می‌شود و ۲۰٪ باقیمانده به‌منظور

[1] Symmetric Mean Absolute Percentage Error (SMAPE)
[2] Mean Absolute Percentage Error (MAPE)
[3] Root Mean Square Error (RMSE)

آزمون در مدل یادگیری عمیق مورد استفاده قرار می‌گیرد. لازم به ذکر است که در مرحلهٔ آموزش اوزان در مدل‌های یادگیری عمیق بر اساس شبکه‌های عصبی به‌صورتی تعدیل می‌شوند که بتوان الگوهای موجود در سری زمانی را شناسایی کرد. در مرحلهٔ تست، بر اساس داده‌های واقعی و اندازه‌گیری معیارهای سنجش کارایی، قدرت پیش‌بینی و تعمیم-دهی مدل به داده‌های واقعی سنجیده می‌شود.

٤-١ فرآیند آموزش

شکل (۵) فرایند آموزش مدل ترکیبی یادگیری عمیق مبتنی بر شبکه CNN-LSTM را نشان می‌دهد. همان‌طور که در این شکل مشخص است بهینه‌سازی فرایند یادگیری در این مدل بر اساس حداقل‌سازی RMSE و تابع زیان انجام شده است. در جدول (۱) تنظیمات هایپرپارمترهای مدل که بر اساس برآورد چندین مدل و انتخاب مدل با بالاترین کارایی انجام شده است را نشان می‌دهد.



شکل (۵). فرآیند آموزش مدل یادگیری عمیق هیبریدی مبتنی بر CNN-LSTM

منبع: محاسبات تحقیق

جدول (۱) تنظیمات هایپرپارامترای مدل CNN-LSTM

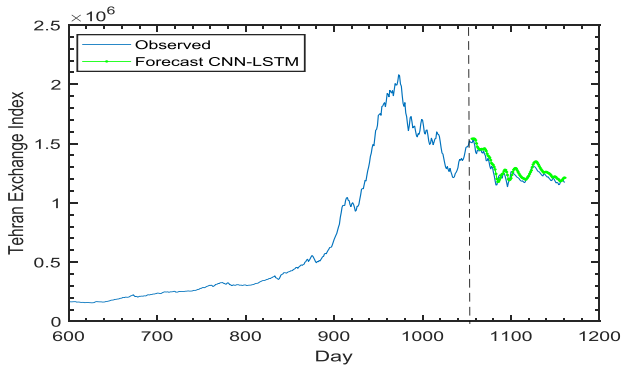
Convolution layer filters	۳۲
Convolution layer activation function	Tanh
Convolution layer padding	Same
Pooling layer padding	Same
Pooling layer activation function	Relu
Number of hidden units in LSTM layer	٦٤
LSTM layer activation function	Tanh

Learning rate	۰.۰۰۱
Optimizer	Adam
Loss function	mean_absolute_error
Epochs	۶۰۰
Iteration per Epoch	8
Iteration	4800

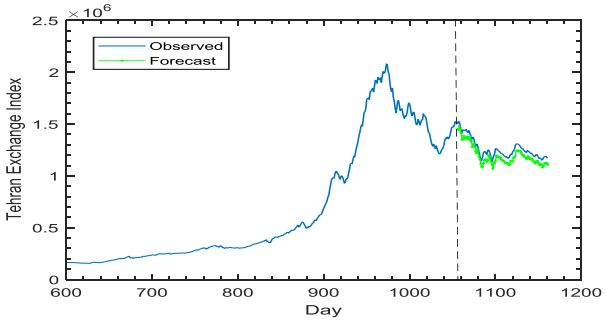
در جدول (۱) تمامی تنظیمات هایپر پارامترهای مدل مورد بررسی را نشان می‌دهد.

٤-۲ پیش‌بینی براساس مدل‌های یادگیری عمیق

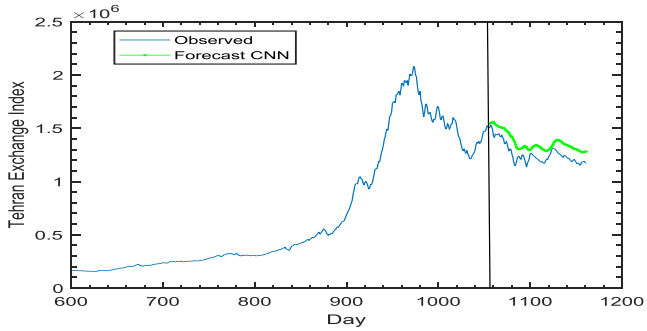
در ادامه و در شکل (۶) تا (۸) نتایج پیش‌بینی شاخص بورس اوراق بهادار در دورهٔ آزمون داده‌ها به ترتیب با استفاده از سه مدل LSTM، CNN و مدل هیبریدی CNN-LSTM ارائه‌شده است. نتایج ارائه‌شده نشان‌دهندهٔ عملکرد و دقت بهتر مدل هیبریدی CNN-LSTM در پیش‌بینی با گام یک روز به جلو نسبت به دو مدل دیگر است. بااین‌حال به‌منظور بررسی دقیق‌تر این موضوع باید کارایی این مدل‌ها در پیش‌بینی وضعیت شاخص بورس بر اساس سه معیار سنجش کارایی بررسی شود.



شکل (٦). پیش‌بینی شاخص بورس اوراق بهادار تهران بر اساس مدل CNN-LSTM



شکل (۷). پیش‌بینی شاخص بورس اوراق بهادار تهران بر اساس مدل LSTM



شکل (۸). پیش‌بینی شاخص بورس اوراق بهادار تهران بر اساس مدل CNN

در این بخش از سه معیار سنجش کارایی به‌منظور ارزیابی عملکرد مدل‌های پیش‌بینی شاخص بورس اوراق بهادار بر اساس روش یادگیری عمیق مبتنی بر شبکه‌های عصبی استفاده می‌شود. نتایج محاسبات معیارهای میانگین درصد قدر مطلق خطای متقارن (SMAPE)، میانگین مطلق درصد خطا (MAPE) و ریشه میانگین مربع خطا (RMSE) در جدول (۲) ارائه شده است. بر اساس نتایج ارائه‌شده در جدول (۲)، مدل CNN-LSTM بهترین عملکرد را داشته است و مدل LSTM در رتبهٔ بعدی دقت و کارایی پیش‌بینی ارائه شده است.

جدول (۲). معیارهای سنجش دقت و کارایی پیش‌بینی مدل‌های یادگیری عمیق

	CNN-LSTM	LSTM	CNN
SMAPE	۰/۰۰۷۰	۰/۰۱۲۵۵۹	۰/۰۲۰۴۴
MAPE	۰/۰۲۸۹	۰/۰۴۸۹	۰/۰۸۵۸
RMSE	۴۳۳۴۱	۶۳۹۵۷	۱۱۲۶۵۰

منبع: یافته‌های پژوهش

۵- بحث و نتیجه‌گیری

اخیراً در حوزه یادگیری ماشینی، گرایشی با نام یادگیری عمیق (DL)، در پیش‌بینی‌های مالی، توجه بسیاری را جلب کرده است. بااینکه یادگیری عمیق در سال‌های ابتدایی توسعه خود قرار دارد، اما روند تحقیقات، مقالات و سرمایه‌گذاری‌های شرکت‌های بزرگ در این حوزه، نشان‌دهندهٔ گسترش روزافزون کاربردهای یادگیری است. با توجه به ظرفیت‌ها و توانایی قابل توجه روش یادگیری عمیق در استخراج اطلاعات معتبر از

مجموعه داده‌ها و الگوهای شناسایی قدرتمند، بسیاری از مقالات اخیر بر کاربرد تکنیک-های DP در حوزهٔ مالی متمرکز شده‌اند و ادغام الگوریتم‌های یادگیری عمیق با پیش‌بینی-های بازار به‌عنوان یکی از جذاب‌ترین مباحث در ادبیات مالی در نظر گرفته می‌شود (Cavalcante et al, 2016).

در این تحقیق تلاش شد مدل هیبریدی یادگیری عمیق مبتنی بر ساختار شبکه‌های عصبی CNN-LSTM به‌منظور پیش‌بینی شاخص بورس اوراق بهادار تهران معرفی شده و کارایی این مدل با ساختارهای سنتی در این حوزه ازجمله مدل LSTM و مدل CNN مقایسه شود. در این راستا از معیارهای سنجش کارایی میانگین درصد قدر مطلق خطای متقارن (SMAPE)، میانگین مطلق درصد خطا (MAPE) و ریشهٔ میانگین مربع خطا (RMSE) استفاده شد. نتایج پیش‌بینی بر اساس سه مدل مذکور، مدل هیبریدی یادگیری عمیق مبتنی بر ساختار شبکه‌های عصبی CNN-LSTM برترین عملکرد را داشته است و مدل یادگیری عمیق مبتنی بر شبکه‌های عصبی LSTM در رتبهٔ بعدی قرار دارد. نتایج این تحقیق نشان می‌دهد که مدل‌های یادگیری عمیق می‌تواند به‌عنوان ابزاری با کارایی بالا در پیش‌بینی روندهای حرکتی متغیرهای مالی مورد استفاده قرار گیرد از اینرو به فعالان بازارهای مالی، استفاده و توجه ویژه به این ابزار پیشنهاد می‌شود.

References:

- Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of arima and artifi- cial neural networks models for stock price prediction. Journal of Applied Math- ematics, 2014.#
- Andrew Ang & Geert Bekaert (2007). "Stock Return Predictability: Is it There?" Review of Financial Studies, Society for Financial Studies, vol. 20(3), pages 651-707.#
- Bollerslev, T., Marrone, J., Xu, L., & Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. Journal of Financial and Quantitative Analysis, 49 (03), 633–661.#
- Krauss, C., Do, X. A., Huck, N. (2017). Deep neural networks, gradient boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research 259(2), 689–702.#
- Christopher Krauss; Xuan Anh Do and Nicolas Huck, (2017), Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500, European Journal of Operational Research, 259, (2), 689-702.#

- Gonçalves, R., Miguel,R. V, Pereira. F., Rocha. A., (2019). Deep learning in exchange markets. Information Economics and Policy.#
- Kara, Y. and O. Baykan (2011). "Predicting Direction of Stock Price Index Changes Using Artificial Neural Networks and Support Vector Machines: The Sample of the Istanbul Stock Exchange", Expert Systems with Applications, Vol. 36, No. 2, pp. 3355-3366.#
- Krauss, C., Do, X. A., Huck, N. (2017). Deep neural networks, gradient- R.C. Cavalcante, R.C. Brasileiro, V.L.F. Souza, J.P. Nobrega, A.L.I. Oliveira, (2016). Computational Intelligence and Financial Markets: A Survey and Future Directions, Expert Systems with Applications, 55,194-211.#
- Tipirisetty, Abhinav, "Stock Price Prediction using Deep Learning" (2018). Master's Projects. 636. DOI: https://doi.org/10.31979/etd.bzmm-36m7.#
- Tipirisetty, Abhinav, (2018). Stock Price Prediction using Deep Learning, Master's Projects. 636.#
- Vadiei, Mohammad Hossein, & Hosseini, Seyyed Mohammad. (2012). The relationship between performance evaluation criteria and abnormal stock return. journal of empirical research in accounting, 1(4), 73-87. #
- W. Long, Z. Lu and L. Cui, (2018). Deep learning-based feature engineering for stock price movement prediction, Knowledge-Based Systems .#
- Yan, L.; Zhang, H. T.; Goncalves, J.; Xiao, Y.; Wang, M. et al. (2020): A machine learning-based model for survival prediction in patients with severe COVID-19 infection. MedRxiv: 2020.2002.2027.20028027#
- Yang Liu, (2019). Novel Volatility Forecasting Using Deep Learning – Long Short Term Memory Recurrent Neural Networks, Expert Systems with Applications.#

COPYRIGHTS
© 2023 by the authors. Licensee Advances in Modern Management Engineering Journal. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0/).