

دریافت مقاله: ۱۴۰۴/۰۸/۱۶

پذیرش مقاله: ۱۴۰۴/۰۹/۱۶

نوع مقاله: پژوهشی

کارایی سازوکار توجه در شبکه‌های عصبی عمیق برای پیش‌بینی قیمت سهام: شواهد تجربی از بورس تهران

رامین خوجانی*

گروه پژوهشی علوم داده زاگرس، دانشگاه آیت الله بروجردی، بروجرد، ایران.
پست الکترونیکی: Khochiany@abru.ac.ir

یونس نادمی

گروه پژوهشی علوم داده زاگرس، دانشگاه آیت الله بروجردی، بروجرد، ایران.
پست الکترونیکی: younesnademi@abru.ac.ir

مجید ابتیاع

گروه پژوهشی علوم داده زاگرس، دانشگاه آیت الله بروجردی، بروجرد، ایران.
پست الکترونیکی: majid.ebtia@gmail.com

چکیده

میانگین عملکرد برای پنج سهم منتخب، مدل پیشنهادی در معیار RMSE به مقدار ۰/۰۳۵ و در معیار MAPE به مقدار ۰/۰۴۶ دست یافت، در حالی که بهترین مدل مرجع (CNN-LSTM) به ترتیب مقادیر ۰/۰۵۲ برای RMSE و ۰/۰۶۶ برای MAPE را ثبت کرد. این بهبود دقت که در معیار ضریب تعیین (R^2) نیز آشکار است (رسیدن به میانگین ۰/۸۶۵ در برابر ۰/۸۰۳ در مدل مرجع)، اثربخشی هم‌کردار هم‌آمیختگی، BiLSTM و سازوکار توجه را تأیید می‌کند. اگر چه پیچیدگی محاسباتی این ساختار نسبتاً بیشتر است، اما دقت بسیار بالاتر در پیش‌بینی قیمت‌های سهام، آن را برای کاربردهای سرمایه‌گذاری در بورس تهران مفید می‌سازد. این پژوهش نخستین مطالعه‌ای است که سازوکار توجه را در پیش‌بینی قیمت سهام بورس تهران به کار گرفته و یافته‌ها نشان می‌دهد مدل پیشنهادی می‌تواند به ابزاری کارآمد برای تحلیلگران و معامله‌گران تبدیل شود.

کلید واژه‌ها: پیش‌بینی قیمت سهام، یادگیری عمیق،

مدل ترکیبی، سازوکار توجه، بورس اوراق بهادار تهران

هدف از این پژوهش بررسی کارایی سازوکار توجه در ساختار ترکیبی شبکه‌های CNN و BiLSTM برای پیش‌بینی قیمت سهام بورس تهران است. در مدل پیشنهادی (CNN-BiLSTM-Attention) ابتدا با لایه‌های هم‌آمیختگی و ویژگی‌های غیرخطی محلی استخراج می‌شوند و سپس با BiLSTM به صورت دوطرفه پردازش می‌شوند. سازوکار توجه نیز به صورت خودکار به بخش‌های مهم توالی قیمتی وزن‌دهی می‌کند. داده‌های تجربی شامل قیمت‌ها و شاخص‌های تکنیکی پنج سهم منتخب (فولاد، فملی، وبملت، فارس، خودرو) در بورس تهران در بازه ۱۴۰۰/۱/۱ تا ۱۴۰۴/۵/۱۵ (مجموعاً ۹۲۹ رکورد برای آموزش و آزمون) استفاده شد. داده‌ها به صورت ۸۰ درصد برای آموزش و ۲۰ درصد برای آزمون تقسیم شدند. نتایج کلیدی نشان داد مدل CNN-BiLSTM-Attention در مقایسه با سایر مدل‌های مرجع به طور معنی‌داری دقت پیش‌بینی بالاتری ارائه می‌کند. در

* نویسنده مسئول

۱- مقدمه

بازار سهام به عنوان یکی از ارکان اصلی بازارهای مالی، نقشی اساسی در تجهیز منابع، تأمین مالی فعالیت‌های تولیدی و انتقال اطلاعات درباره چشم‌انداز اقتصاد کلان دارد و نوسانات شدید آن می‌تواند به کاهش اعتماد سرمایه‌گذاران و تشدید بی‌ثباتی مالی منجر شود. از این رو، پیش‌بینی دقیق قیمت و بازده سهام برای مدیریت ریسک، جلوگیری از بحران‌های مالی و بهبود تصمیم‌گیری سرمایه‌گذاران، مدیران سبد و نهادهای ناظر اهمیت زیادی دارد. با این حال، سری‌های زمانی مالی به دلیل ویژگی‌هایی نظیر نالیستایی، ناهمسان‌واری، جهش‌های ناگهانی و روابط غیرخطی پیچیده، پیش‌بینی‌پذیری دشواری دارند و مدل‌های کلاسیک اقتصادسنجی مانند خانواده ARMA-GARCH علی‌رغم کارایی نسبی در برآورد نوسان، در دوره‌های نوسان بالا و ساختارهای پیچیده بازار با محدودیت دقت مواجه‌اند. [۱]

در دهه اخیر، پیشرفت‌های چشم‌گیر در حوزه هوش مصنوعی و یادگیری عمیق، زمینه‌ساز کاربرد گسترده شبکه‌های عصبی عمیق برای مدل‌سازی پویایی پیچیده بازار سهام شده است. مطالعات متعددی نشان داده‌اند که مدل‌های یادگیری عمیق در مقایسه با روش‌های متعارف یادگیری ماشین و اقتصادسنجی در شناسایی الگوهای غیرخطی و وابستگی‌های بلندمدت سری‌های مالی و در نتیجه بهبود دقت پیش‌بینی، عملکرد برتری دارند. در این میان، شبکه‌های عصبی هم‌آختی (CNN) به واسطه توانایی در استخراج ویژگی‌های محلی و کوتاه‌مدت از توالی داده‌ها، و شبکه‌های عصبی حافظه طولانی-کوتاه‌مدت (LSTM) به دلیل ساختار دروازه‌ای و قابلیت حفظ وابستگی‌های زمانی بلندمدت، به طور گسترده در پیش‌بینی قیمت سهام به کار گرفته شده‌اند. [۲]

ترکیب این دو ساختار در قالب مدل‌های ترکیبی CNN-LSTM و CNN-BiLSTM، امکان بهره‌گیری هم‌زمان از استخراج ویژگی‌های مکانی و مدل‌سازی پویایی‌های

زمانی سری‌های مالی را فراهم کرده و نتایج تجربی در ادبیات نشان می‌دهد که این مدل‌های ترکیبی، در مقایسه با مدل‌های منفرد، دقت پیش‌بینی بالاتری ارائه می‌کنند. افزون بر این، نسخه دوطرفه (BiLSTM) LSTM با پردازش توالی داده در دو جهت گذشته و آینده، توانایی بیشتری در بازنمایی وابستگی‌های پیچیده زمانی داشته و در حوزه پیش‌بینی قیمت و نوسان سهام نتایج امیدوارکننده‌ای ارائه کرده است. در کنار این تحولات، سازوکار توجه به عنوان ابزاری برای تمرکز پویا بر بخش‌های اثرگذارتر داده، به معماری‌های مبتنی بر LSTM و BiLSTM افزوده شده و شواهد تجربی حاکی از آن است که استفاده از لایه‌های توجه می‌تواند بهبود معناداری در شاخص‌های عملکردی همچون RMSE، MAE و R2 ایجاد کند. [۳]

در بازارهای نوظهوری نظیر بورس اوراق بهادار تهران نیز، کاربرد مدل‌های یادگیری عمیق برای پیش‌بینی قیمت و بازده سهام مورد توجه قرار گرفته و برخی پژوهش‌ها نشان داده‌اند که ترکیب CNN و LSTM می‌تواند بر دقت پیش‌بینی در مقایسه با مدل‌های خطی و نیز برخی الگوریتم‌های متعارف یادگیری ماشین بیفزاید. با وجود این، هنوز مطالعات اندکی به ارزیابی نظام‌مند مدل‌های ترکیبی مبتنی بر CNN، BiLSTM و سازوکار توجه در بستر بورس تهران و در افق‌های زمانی مختلف پرداخته‌اند و جای خالی بررسی دقیق اثر افزودن لایه توجه بر عملکرد این مدل‌های ترکیبی احساس می‌شود [۲].

نوآوری‌ها و مشارکت‌های اصلی این پژوهش را می‌توان به‌طور خلاصه در چهار محور زیر برشمرد:

۱. طراحی و ارائه یک معماری ترکیبی جدید مبتنی بر ترکیب شبکه هم‌آختی (CNN)، شبکه دوطرفه LSTM و BiLSTM و سازوکار توجه برای پیش‌بینی قیمت سهام که به‌طور هم‌زمان از قابلیت استخراج ویژگی‌های مکانی و مدل‌سازی وابستگی‌های زمانی دوطرفه بهره می‌گیرد.

۲. به‌کارگیری مدل پیشنهادی بر داده‌های واقعی بورس اوراق بهادار تهران و ارزیابی عملکرد آن با تکیه بر

شاخص‌های مختلف خطا و برازش.

۳. انجام مقایسه‌ای نظام‌مند بین مدل پیشنهادی و چند مدل مرجع شامل LSTM ساده، CNN-LSTM و BiLSTM بدون سازوکار توجه از منظر دقت، پایداری پیش‌بینی و حساسیت نسبت به نوسانات شدید بازار.

۴. تحلیل میزان بهبود حاصل از افزودن لایه توجه به ساختار CNN-BiLSTM و نشان دادن این‌که مدل پیشنهادی، به‌طور متوسط، منجر به کاهش معنادار خطاهای RMSE و MAE و افزایش ضریب تعیین R^2 نسبت به مدل‌های مرجع می‌شود.

به‌طور مشخص، نتایج تجربی پژوهش نشان می‌دهد که مدل CNN-BiLSTM-Attention در مقایسه با سایر مدل‌های مرجع به‌طور معنی‌داری دقت پیش‌بینی بالاتری ارائه می‌کند. در میانگین عملکرد برای پنج سهم منتخب، مدل پیشنهادی در معیار RMSE به مقدار ۰/۰۳۵ و در معیار MAPE به مقدار ۰/۰۴۶ دست یافت، در حالی که بهترین مدل مرجع (CNN-LSTM) به ترتیب مقادیر ۰/۰۵۲ برای RMSE و ۰/۰۶۶ برای MAPE را ثبت کرد. این بهبود دقت که در معیار ضریب تعیین (R^2) نیز آشکار است (رسیدن به میانگین ۰/۸۶۵ در برابر ۰/۸۰۳ در مدل مرجع)، اثربخشی هم‌کردار هم‌آمیختی، BiLSTM و سازوکار توجه را تأیید می‌کند.

ساختار مقاله نیز بدین‌صورت است که در بخش ۲، مبانی نظری و پیشینه پژوهش مرور می‌شود، سپس در بخش ۳، روش‌شناسی مدل پیشنهادی تشریح می‌شود، در بخش ۴، نتایج پژوهش گزارش شده و در نهایت در بخش ۵، نتیجه‌گیری و پیشنهادها می‌باشد که یافته‌ها را جمع‌بندی کرده و چشم‌انداز تحقیقات آینده را مطرح می‌نماید.

۲- مبانی نظری و مروری بر پیشینه پژوهش

پیش‌بینی قیمت سهام یکی از چالش‌برانگیزترین حوزه‌ها در علوم مالی و مهندسی داده است، چرا که ماهیت بازارهای مالی پیچیده، غیرخطی و غیرایستا بوده و

تحت تأثیر عوامل متعددی از جمله متغیرهای اقتصاد کلان، سیاست‌های دولتی، رفتار سرمایه‌گذاران و حتی اخبار و رویدادهای اجتماعی قرار دارد. روش‌های سنتی در تحلیل سری‌های زمانی مالی نظیر مدل‌های خودتوضیح یکپارچه میانگین متحرک و مدل‌های نوسان شرطی تعمیم‌یافته توانایی محدودی در مدل‌سازی الگوهای غیرخطی و وابستگی‌های بلندمدت دارند. به همین دلیل، در دو دهه اخیر تمرکز پژوهش‌ها به سمت بهره‌گیری از الگوریتم‌های یادگیری ماشین و به‌ویژه یادگیری عمیق معطوف شده است [۵].

یادگیری عمیق به‌عنوان شاخه‌ای پیشرفته از هوش مصنوعی، امکان کشف الگوهای پیچیده و غیرخطی در داده‌های بزرگ و پرنوفه مالی را فراهم ساخته است. مزیت اصلی این رویکردها در توانایی خودکار آنها در استخراج ویژگی‌های پنهان و نمایه‌سازی چندلایه از داده‌ها نهفته است. برخلاف مدل‌های سنتی که نیازمند تعریف دستی شاخص‌ها و ویژگی‌ها بودند، شبکه‌های عصبی عمیق قادرند به‌طور مستقیم از داده‌های خام، ساختارهای مفید و معنادار را شناسایی کنند [۶]. این ویژگی باعث شده که یادگیری عمیق به‌طور گسترده در مسائلی مانند پیش‌بینی روند شاخص‌های بورس، تشخیص نوسانات بازار، و مدیریت ریسک مورد استفاده قرار گیرد.

۲-۱ شبکه‌های عصبی هم‌آمیختی (CNN) و نقش آنها در تحلیل داده‌های مالی

شبکه‌های عصبی هم‌آمیختی که ابتدا برای تحلیل تصاویر و پردازش داده‌های فضایی توسعه یافتند، به‌سرعت به حوزه سری‌های زمانی نیز وارد شدند. دلیل این موضوع آن است که CNN با اعمال فیلترهای هم‌آمیختی قادر است الگوهای محلی و کوتاه‌مدت موجود در داده‌های مالی را استخراج کند. در بازار سهام، این الگوها می‌توانند شامل تغییرات لحظه‌ای قیمت، رفتارهای تکرارشونده در بازه‌های زمانی خاص یا واکنش سریع بازار به اخبار

۲-۳ سازوکار توجه

سازوکار توجه یکی از نوآوری‌های کلیدی در یادگیری عمیق است که نخستین بار در پردازش زبان طبیعی معرفی شد [۱۰]. ایده اصلی آن شبیه به فرایند شناختی انسان است که هنگام پردازش اطلاعات، بر بخش‌های مهم تمرکز می‌کند و سایر بخش‌ها را نادیده می‌گیرد. در مدل‌های پیش‌بینی سری‌های زمانی، سازوکار توجه این امکان را فراهم می‌کند که مدل بتواند وزن بیشتری به نقاط عطف و بخش‌های کلیدی داده‌های قیمتی بدهد. به عنوان مثال، تغییرات شدید ناشی از اخبار مهم یا رفتارهای ناگهانی بازار می‌تواند وزن بالاتری دریافت کند و نقش پررنگ‌تری در پیش‌بینی آینده ایفا کند [۱۱].

۲-۴ معماری‌های ترکیبی

یکی از رویکردهای نوظهور در پیش‌بینی مالی، ترکیب CNN، BiLSTM و سازوکار توجه است. در این معماری، ابتدا لایه‌های CNN ویژگی‌های محلی کوتاه‌مدت را استخراج می‌کنند. سپس لایه‌های BiLSTM وابستگی‌های بلندمدت و دوطرفه را یاد می‌گیرند. در نهایت، سازوکار توجه بخش‌های مهم توالی را شناسایی کرده و وزن بیشتری به آنها اختصاص می‌دهد. این رویکرد چندمزیت کلیدی دارد: کاهش نوفه داده‌ها از طریق CNN.

درک وابستگی‌های بلندمدت توسط BiLSTM.

تمرکز بر سیگنال‌های کلیدی از طریق توجه.

پژوهش‌ها نشان داده‌اند که این ترکیب می‌تواند نسبت به مدل‌های منفرد (مانند CNN یا LSTM) خطای پیش‌بینی را در معیارهایی نظیر MAE، RMSE و MAPE به شکل معناداری کاهش دهد [۱۴].

۲-۵ کاربردهای عملی مدل‌های ترکیبی در بازارهای مالی

کاربردهای مدل CNN-BiLSTM-Attention تنها به پیش‌بینی قیمت محدود نمی‌شود. این مدل می‌تواند در طراحی راهبردهای معاملاتی خودکار، مدیریت ریسک، و

باشند. CNN به‌ویژه در مرحله پیش‌پردازش داده‌های سری زمانی و استخراج ویژگی‌های مقدماتی عملکرد قابل‌توجهی دارد. همچنین، لایه‌های Pooling در CNN امکان کاهش ابعاد داده‌ها و حذف نوفه‌های غیرضروری را فراهم می‌کنند که این امر در داده‌های مالی پرنوسان اهمیت زیادی دارد.

۲-۲ شبکه‌های بازگشتی و LSTM

شبکه‌های عصبی بازگشتی (RNN) برای پردازش داده‌های ترتیبی طراحی شده‌اند. ویژگی کلیدی آنها توانایی نگهداری اطلاعات گذشته در حالت‌های پنهان است. با این حال، شبکه‌های عصبی بازگشتی معمولی در یادگیری وابستگی‌های بلندمدت دچار مشکل هستند که این پدیده به نام «محو شدن یا انفجار گرادیان» شناخته می‌شود [۷]. برای رفع این مشکل، ساختار حافظه کوتاه‌مدت بلندمدت (LSTM) معرفی شد [۸]. با استفاده از سازوکار سلول‌های حافظه و دروازه‌های ورودی، خروجی و فراموشی قادر است وابستگی‌های بلندمدت در داده‌های مالی را یاد بگیرد. این ویژگی برای پیش‌بینی قیمت سهام اهمیت دارد زیرا تغییرات قیمتی اغلب تحت تأثیر روندها و چرخه‌های طولانی مدت قرار دارند.

اگر چه LSTM قادر به یادگیری وابستگی‌های بلندمدت است، اما جریان اطلاعات در آن یک‌طرفه (از گذشته به آینده) است. این محدودیت در مواردی که داده‌های مالی تحت تأثیر هم‌زمان روندهای گذشته و نشانه‌های آتی قرار دارند، باعث افت دقت می‌شود. شبکه‌های LSTM دوطرفه (BiLSTM) این مشکل را برطرف می‌کنند زیرا داده‌ها را هم در جهت رو به جلو و هم در جهت رو به عقب پردازش می‌کنند [۹]. در پیش‌بینی سهام، BiLSTM می‌تواند هم‌زمان تأثیر نوسانات گذشته و سیگنال‌های کنونی را درک کرده و الگوهای دقیق‌تری از پویایی بازار ارائه دهد. مطالعات اخیر نشان داده‌اند که BiLSTM در مقایسه با LSTM معمولی، دقت بیشتری در پیش‌بینی روندهای پیچیده مالی دارد [۱۳].

بورس هند نشان داد که این مدل توانسته است همبستگی پیش‌بینی با داده‌های واقعی را افزایش دهد و نسبت به مدل‌های کلاسیک عملکرد بهتری داشته باشد [۱۹].

یک مدل CNN-BiLSTM-Attention همراه با تحلیل احساسات (Sentiment Analysis) را به کار گرفتند. نتایج نشان داد که ادغام داده‌های متنی (احساسات اخبار و شبکه‌های اجتماعی) با داده‌های قیمتی توانسته است بهبود قابل‌توجهی در دقت پیش‌بینی ایجاد کند [۱۸].

مطالعات داخلی

در پژوهشی با عنوان کاربرد و اثربخشی روش‌های یادگیری ماشین و یادگیری عمیق در پیش‌بینی شاخص بورس شانگهای، مدل CNN-BiLSTM ترکیبی برای شبیه‌سازی پویایی‌های پیچیده بازار استفاده شد و نتایج نشان داد که مدل ترکیبی نسبت به LSTM یا CNN به‌تنهایی برتری دارد [۴].

برای پیش‌بینی قیمت سهام در بازار بورس اوراق بهادار تهران از دو مدل پیشنهادی LSTM-CNN و LSTM-CNN استفاده نمودند [۲]. این مدل‌ها با مدل‌های LSTM، و CNN-LSTM مورد مقایسه قرار گرفته‌اند و مقایسه نتایج معیارهای ارزیابی نشان از عملکرد مطلوب دو مدل پیشنهادی با سایر مدل‌ها می‌دهد. علاوه بر این، مدل پیشنهادی LSTM-CNN توانسته بهترین عملکرد را به ثبت رساند. همچنین، بررسی نتایج دو مدل پیشنهادی، با و بدون الگوریتم PSO مشخص کرد که این الگوریتم به بهینه‌سازی مدل‌ها کمک می‌کند. همچنین مدل‌های پیشنهادی LSTM-CNN و LSTM-CNN توانسته‌اند در معیار بازده مالی و نسبت شارپ در مقایسه با سایر راهبردها موفقیت مطلوبی به‌دست آورند که مدل پیشنهادی LSTM-CNN مطلوب‌ترین عملکرد را به ثبت رسانده است [۳].

۳- روش شناسی مدل پیشنهادی

به طور خلاصه، وجه تمایز پژوهش حاضر با مطالعات

حتی تحلیل احساسات بازار با استفاده از داده‌های متنی (مانند اخبار و شبکه‌های اجتماعی) به کار رود. بهبود دقت پیش‌بینی به سرمایه‌گذاران کمک می‌کند تا زمان‌بندی بهینه معاملات را شناسایی کرده، سود خود را حداکثر و زیان را حداقل کنند [۱۲]. در بورس‌های نوظهور مانند بورس تهران که نوسانات بالا و عدم شفافیت اطلاعات از ویژگی‌های اصلی آن هستند، چنین مدل‌هایی می‌توانند ابزار ارزشمندی برای تصمیم‌گیری سرمایه‌گذاران و تحلیلگران باشند.

۲-۶ پیشینه پژوهش

مطالعات خارجی

به پیش‌بینی نوسانات بازار سهام با استفاده از مدل ترکیبی CNN-BiLSTM-Attention پرداخته است. داده‌ها شامل اطلاعات با فرکانس‌های متفاوت (دقیقه‌ای و روزانه) بودند و نتایج نشان داد که ترکیب لایه‌های هم‌آمیختی و بازگشتی همراه با سازوکار توجه توانسته است دقت بالاتری در مقایسه با مدل‌های سنتی LSTM و ARIMA ارائه دهد [۱۵].

یک مدل ترکیبی CEEMD-CNN-BiLSTM-AM معرفی کردند. این مدل ابتدا با تجزیه سری زمانی (CEEMD) داده‌های پرنویز بازار را تفکیک کرده و سپس با شبکه‌های عمیق (CNN و BiLSTM) الگوها را شناسایی می‌کند. لایه توجه (Attention Mechanism) باعث شد که ویژگی‌های مهم سری زمانی برای پیش‌بینی قیمت سهام استخراج شوند [۱۶].

مدل Attention-Augmented BiLSTM-CNN را برای پیش‌بینی قیمت پایانی روزانه سهام ارائه کرد. آزمایش روی داده‌های S&P500 نشان داد که این مدل در مقایسه با LSTM-CNN و GRU توانست خطای RMSE را به میزان قابل‌توجهی کاهش دهد [۱۷].

یک مدل ترکیبی MHABiLSTM-CNN پیشنهاد دادند که از سازوکار Multi-Head Attention برای استخراج ویژگی‌های چندبعدی استفاده کردند. نتایج تجربی بر روی داده‌های

انجام شده قبلی در ۵ بعد زیر است:

بومی سازی در بازار بورس تهران با ویژگی های خاص اقتصادی-سیاسی.

مقایسه با سه نمونه از مدل های مرجع.

تمرکز بر کاربردهای عملی در معاملات و مدیریت ریسک نخستین پژوهش داخلی با بهره گیری جدی از سازوکار توجه.

مدل پژوهش، ویژگی ها و شناخت آنها

در این مطالعه، داده های تاریخی قیمت سهام پنج شرکت منتخب بورس تهران که در وبگاه شرکت مدیریت فناوری بورس تهران در دسترس می باشد مورد استفاده قرار گرفت. ابتدا داده ها با حذف مقادیر گمشده و مقادیر پرت پاک سازی شدند و سپس با استفاده از روش های نرمال سازی و مقیاس بندی پیش پردازش شدند تا دامنه ویژگی های ورودی یکنواخت و مناسب مدل سازی شود. داده های پیش پردازش شده به دو بخش آموزش (مثلاً ۸۰ درصد داده های اولیه) و آزمون (۲۰ درصد پایانی) تقسیم شدند. تقسیم بندی زمانی به این صورت انجام شد که بازه زمانی ابتدایی برای آموزش مدل و بازه انتهایی برای ارزیابی پیش بینی نگه داشته شود؛ بدین ترتیب، مدل با داده های قدیمی تر آموزش دیده و با داده های جدیدتر سنجیده شد. برای سنجش کیفیت پیش بینی مدل ها از چهار معیار استاندارد RMSE، MAPE، MSE و R^2 استفاده شد. سه معیار اول میانگین درصد خطای مطلق، میانگین مربع خطا، ریشه میانگین مربع خطا و معیار آخر دقت یا ضریب تعیین را نمایش می دهد.

تمام مدل های مورد نظر با استفاده از کتابخانه های متن باز یادگیری عمیق پیاده سازی شدند. برای هر یک از سهام منتخب، ورودی مدل شامل توالی ای از ویژگی های قیمت (مثلاً قیمت باز، بسته، بالاترین، پایین ترین) بود. شبکه LSTM با تعریف لایه LSTM در Keras ساخته شد که به طور ویژه قادر به یادگیری وابستگی های زمانی بلندمدت است.

در مدل های CNN-LSTM و CNN-BiLSTM، ابتدا یک یا چند لایه هم آمیختی برای استخراج ویژگی های مکانی از داده های توالی اعمال شده و سپس خروجی این بخش هم آمیختی به لایه های LSTM یا Bidirectional LSTM وارد شد.

مدل های CNN-LSTM-Attention و CNN-BiLSTM-Attention مشابه مورد قبل بودند، با این تفاوت که یک سازوکار «توجه» به معماری اضافه شده بود تا بر گام های زمانی مهم تمرکز کند. در عمل، پس از لایه های LSTM یا BiLSTM، یک لایه توجه تعبیه می شد که وزن دهی به خروجی های توالی را بر اساس اهمیت آنها یاد می گرفت. به این ترتیب، مدل توانست به سادگی به جنبه های کلیدی اطلاعات ورودی توجه کند و دقت پیش بینی را افزایش دهد. مطابق نتایج مطالعات مشابه، استفاده از سازوکار توجه همراه با شبکه های ترکیبی CNN-BiLSTM موجب بهبود چشمگیر دقت پیش بینی شد؛ به عنوان مثال، مدل CNN-BiLSTM-At- tention کمترین مقدار RMSE و MAPE و بالاترین R^2 را در مقایسه با سایر معماری ها نشان داد.

۳-۱ معماری مدل ها

هر سلول LSTM شامل یک حافظه داخلی و چندین دروازه (ورودی، فراموشی و خروجی) است که جریان اطلاعات را تنظیم می کنند. این ساختار به LSTM امکان می دهد وابستگی های بلندمدت در داده های زمانی را به خوبی یاد بگیرد. لایه های LSTM در کنار لایه های RNN کلاسیک، مشکل گم شدن گرادیان را برطرف کرده و پیش بینی دقیق تر توالی های بلند را فراهم می کنند.

ساختار یک شبکه CNN استاندارد برای مثال معماری VGG-16 را نشان می دهد که متشکل از چند بلوک هم آمیختی و بلوک های تجمیع (Pooling) است. در مدل های پیش بینی قیمت سهام مبتنی بر CNN-LSTM، بلوک هم آمیختی در ابتدا قرار می گیرد تا ویژگی های فضایی یا ساختاری از داده های ورودی استخراج شود. سپس خروجی ها به یک

۴- برآورد مدل و نتایج پژوهش

برای سهم فولاد، که نوسانات قابل‌توجهی را نشان می‌دهد، مدل CNN-BiLSTM-Attention نتایج برتری نسبت به سایر مدل‌ها ارائه داد. به دلیل نوسانات بالا، مدل‌های ساده‌تر مانند LSTM یا CNN-LSTM با خطاهای بزرگ‌تری در پیش‌بینی مواجه شدند، در حالی که سازوکار توجه مدل پیشنهادی توانست تغییرات کوتاه‌مدت قیمت را بهتر شناسایی کند. در عمل، این مدل مقادیر کمتری از MAPE و RMSE را ثبت کرد و ضریب تعیین آن (نزدیک به ۱) بالاتر از مدل‌های پایه بود؛ برای مثال در مطالعات مقایسه‌ای مقدار R^2 مدل پیشنهادی حدود ۰/۹۸۵ گزارش شده است. نتایج آزمون فریدمن اختلاف معناداری را میان عملکرد مدل‌ها تأیید کرد ($p < 0.05$)، نشان می‌دهد که بهبود حاصل نسبت به مدل‌های پایه قطعی است. همچنین خطاهای مطلق پیش‌بینی (RMSE) و خطای نسبی (MAPE) در این سهم برای مدل CNN-BiLSTM-Attention کمینه بود که بیانگر دقت نسبی بالای این مدل در تحلیل رفتار سهام فولاد است.

جدول ۲. نتایج پیش‌بینی مدل‌های مختلف برای سهام فولاد

المترینم‌های یادگیری ماشین	r^2	mae	rmse	mse
LSTM	۰/۶۷۹۰	۰/۳۸۹	۰/۴۴۱	۰/۰۰۱۹
CNN-LSTM	۰/۸۷۹۷	۰/۰۱۹۶	۰/۰۲۷۲	۰/۰۰۰۷
CNN-LSTM-Attention	۰/۸۸۹۲	۰/۰۱۸۰	۰/۰۲۶۱	۰/۰۰۰۶
CNN-BiLSTM-Attention	۰/۹۰۴۴	۰/۰۱۶۶	۰/۰۲۴۳	۰/۰۰۰۵

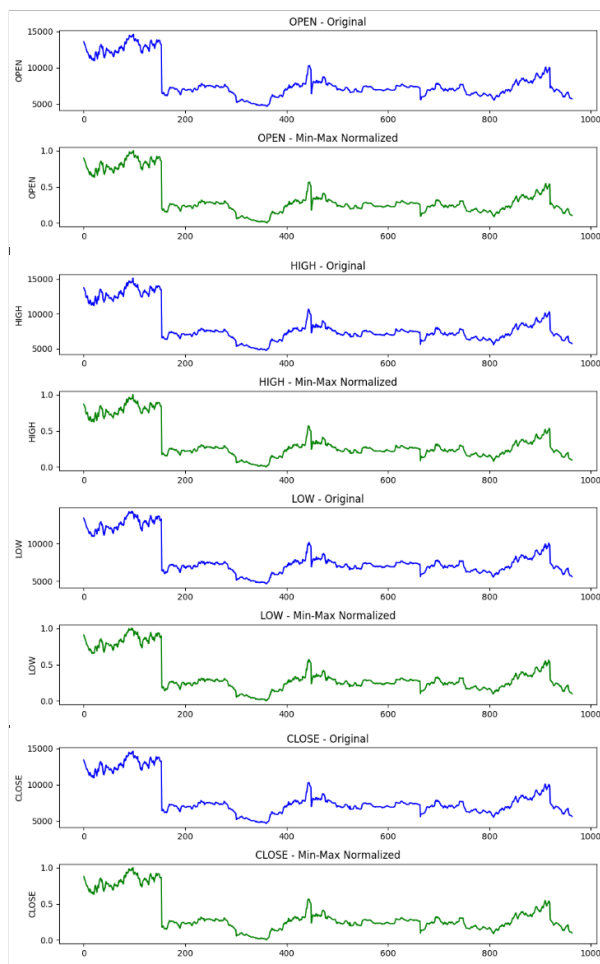
در مورد سهم فملی، مدلی مانند CNN-BiLSTM-Attention نیز دقت پیش‌بینی بالاتری نسبت به سایر مدل‌ها داشت. این سهم که تحت تأثیر نوسانات قیمت جهانی فلزات پایه قرار دارد، چالش‌هایی مشابه فولاد دارد، اما مدل پیشنهادی با بهره‌گیری از ساختار دوبرخی LSTM و سازوکار توجه، توانست نوبت کوتاه‌مدت را بهتر فیلتر کرده و روند اصلی قیمت را ثبت کند. به همین جهت، مقادیر RMSE و MAPE

یا چند لایه LSTM یا BiLSTM ارسال می‌شوند تا روابط زمانی بین داده‌ها مدل‌سازی گردد. معماری تلفیقی CNN-LSTM به این صورت است که ابتدا ویژگی‌های مهم با عملیات هم‌آمیختگی بر روی توالی داده‌ها شناسایی می‌شوند و سپس این ویژگی‌های مکانی توسط LSTM به عنوان ورودی زمانی استفاده می‌شوند. در شکل کلی چنین مدلی، لایه هم‌آمیختگی و تجمیع در ورودی قرار گرفته و پس از آن لایه‌های LSTM قرار می‌گیرند که خروجی نهایی را پیش‌بینی می‌کنند. مدل CNN-BiLSTM نیز مشابه است با این تفاوت که به جای LSTM، از لایه دوطرفه BiLSTM استفاده می‌کند تا اطلاعات هر توالی هم از گذشته و هم از آینده کاوش شود. اضافه کردن سازوکار به این معماری‌ها مثلاً در «CNN-LSTM-Attention» و «CNN-BiLSTM-Attention» باعث شد تا مدل بتواند گام‌های زمانی پر اهمیت‌تر را برجسته کند و دقت پیش‌بینی را بیشتر نماید.

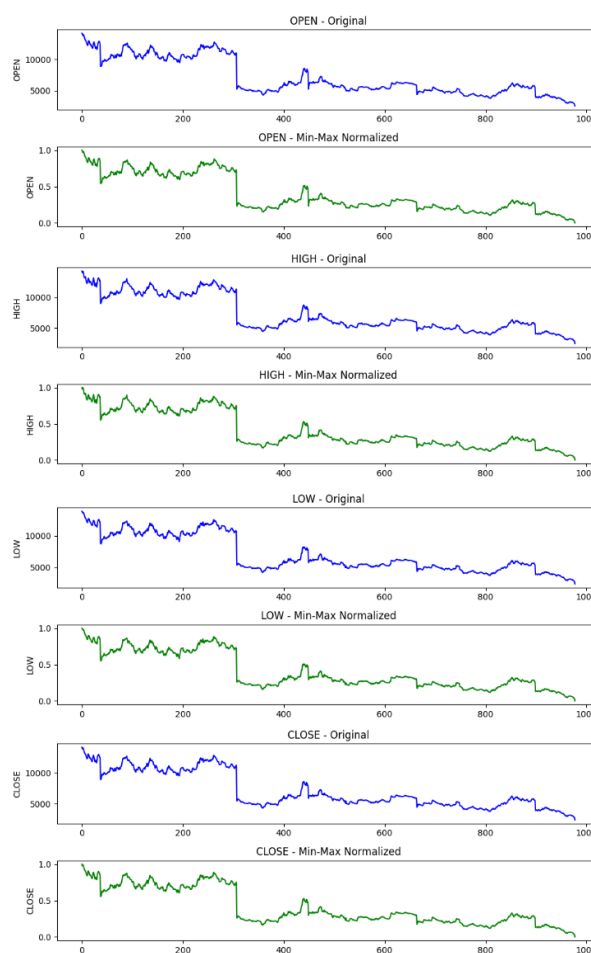
هر یک از این معماری‌ها با تنظیم مناسب تعداد لایه‌های هم‌آمیختگی، فیلترها و نوروهای LSTM پیاده‌سازی شدند. به عنوان مثال، می‌توان از دو لایه ۱-D هم‌آمیختگی با تابع فعال‌سازی ReLU، یک لایه MaxPooling و سپس یک یا دو لایه LSTM یا BiLSTM استفاده کرد. علاوه بر این، به منظور جلوگیری از بیش‌برازش، در برخی مدل‌ها از لایه‌های Dropout و تنظیم‌کننده‌های وزن استفاده شد. تمامی مدل‌ها با توابع هزینه‌ی مناسب (مثلاً میانگین مربع خطا برای پیش‌بینی و بهینه‌سازهای استاندارد آموزش داده شدند. در نهایت، عملکرد هر مدل با استفاده از معیارهای اشاره شده مورد مقایسه قرار گرفت.

جدول ۱. سهم‌ها و بازه زمانی انتخابی

مجموعه داده	زمان شروع	زمان پایان
فولاد	۱۴۰۰/۱/۱	۱۴۰۴/۵/۱۵
فملی	۱۴۰۰/۱/۱	۱۴۰۴/۵/۱۵
وبملت	۱۴۰۰/۱/۱	۱۴۰۴/۵/۱۵
فارس	۱۴۰۰/۱/۱	۱۴۰۴/۵/۱۵
خودرو	۱۴۰۰/۱/۱	۱۴۰۴/۵/۱۵



نمودار ۲. نمودار داده های سری زمانی سهام فملی



نمودار ۱. نمودار داده های سری زمانی سهام فولاد

سهام وبملت (بانک ملت) که معمولاً نوسانات کمتری نسبت به بازار فلزات دارد، در سطح عمومی مدل های مختلف خطاهای پیش بینی پایین تری را نشان دادند. با این حال، مدل CNN-BiLSTM-Attention همچنان کمترین مقادیر MAPE و RMSE را نسبت به مدل های پایه کسب کرد [۱۲]. به طور مثال، در سایر مطالعات مشابه پیش بینی شاخص های بانکی، مدل پیشنهادی منجر به مقادیر خطا و درصد خطای نسبی پایین تری شد. ضریب تعیین این مدل برای وبملت بسیار نزدیک به ۱ بود که نشان دهنده دقت بالای پیش بینی است. آزمون فریدمن یکبار دیگر اختلاف معنادار میان مدل ها را مورد تأیید قرار داد و بیان کرد که عملکرد بالاتر مدل CNN-BiLSTM-Attention تصادفی نبوده است. در این سهم نسبتاً با ثبات، مدل پیشنهادی به دلیل قابلیت یادگیری الگوهای میان مدت قیمت کارآیی بیشتری داشت.

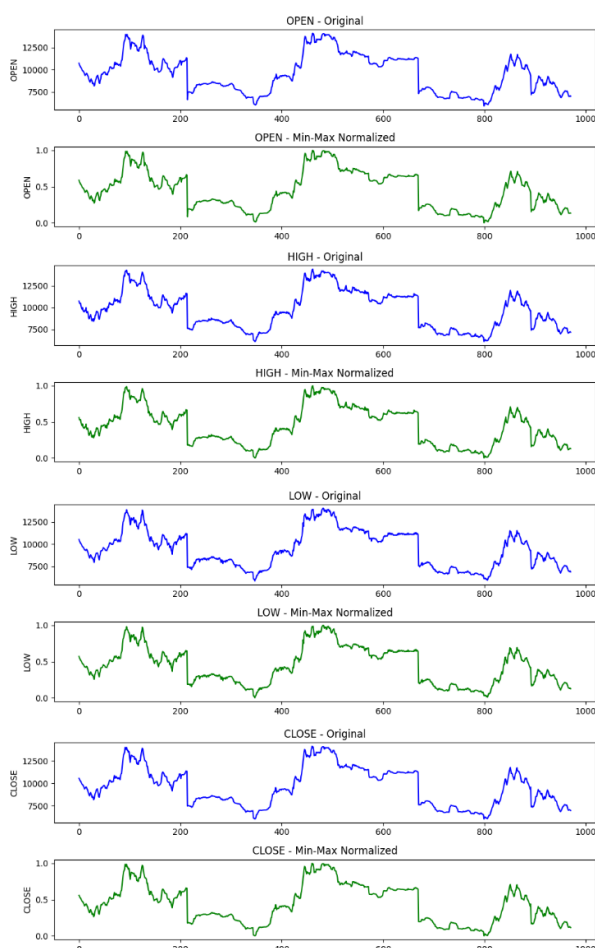
جدول ۳. نتایج پیش بینی مدل های مختلف برای سهام فملی

الگوریتم های یادگیری ماشین	r^2	mae	rmse	mse
LSTM	۰/۵۶۷۳	۰/۰۶۵۸	۰/۰۷۷۲	۰/۰۰۵۹
CNN-LSTM	۰/۸۴۵۶	۰/۰۳۲۰	۰/۰۴۶۵	۰/۰۰۲۱
CNN-LSTM-Attention	۰/۸۴۸۱	۰/۰۳۲۱	۰/۰۴۶۱	۰/۰۰۲۱
CNN-BiLSTM-Attention	۰/۸۷۰۵	۰/۰۲۹۷	۰/۰۴۲۶	۰/۰۰۱۸

این مدل برای سهم فملی به طور قابل توجهی از مقادیر متناظر مدل های LSTM و CNN-LSTM کمتر بود. ضریب تعیین به دست آمده نیز بالاتر بوده و نشان دهنده تناسب بهتر پیش بینی با داده های واقعی است. نتایج آزمون فریدمن مجدداً تفاوت عملکرد مدل ها را در این سهم معنادار یافت، که به برتری آماری مدل پیشنهادی اشاره می کند. در مجموع، ویژگی رفتاری فملی - به ویژه نوسانات مرتبط با روند جهانی فلزات توسط مدل CNN-BiLSTM-Attention با دقت بیشتری پیش بینی شد.

جدول ۵. نتایج پیش‌بینی مدل‌های مختلف برای سهام فارس

mse	$rmse$	mae	r^2	الگوریتم‌های یادگیری ماشین
۰/۰۰۸۹	۰/۰۹۴۵	۰/۰۸۰۴	۰/۷۵۸۹	LSTM
۰/۰۰۴۶	۰/۰۶۷۹	۰/۰۵۳۵	۰/۸۷۴۳	CNN-LSTM
۰/۰۰۴۲	۰/۰۶۵۴	۰/۰۵۰۳	۰/۸۸۳۵	CNN-LSTM-Attention
۰/۰۰۲۹	۰/۰۵۴۵	۰/۰۴۰۰	۰/۹۱۹۱	CNN-BiLSTM-Attention

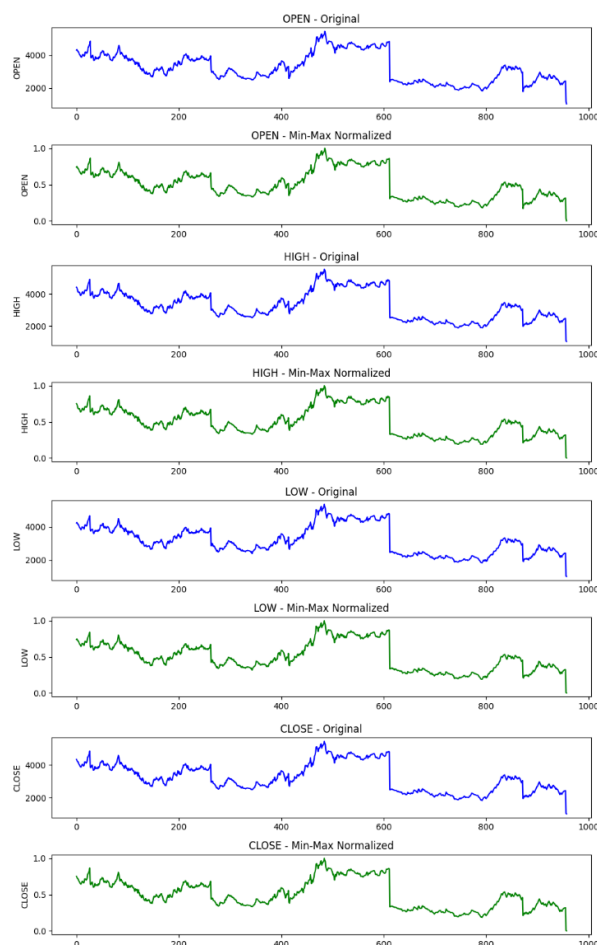


نمودار ۴. نمودار داده‌های سری زمانی سهام فارس

پیش‌بینی دقیق‌تر روندهای قیمت بود. بر اساس آزمون فریدمن، اختلاف عملکرد مدل CNN-BiLSTM-Attention نسبت به مدل‌های پایه معنادار اعلام شد که این امر بهبود این مدل را معتبر می‌سازد. با توجه به نوسانات نسبتاً کم در فارس، مدل‌های عمیق در استخراج الگوهای احتمالی درازمدت بهتر عمل کردند و در نتیجه خطای کلی پیش‌بینی کاهش یافت.

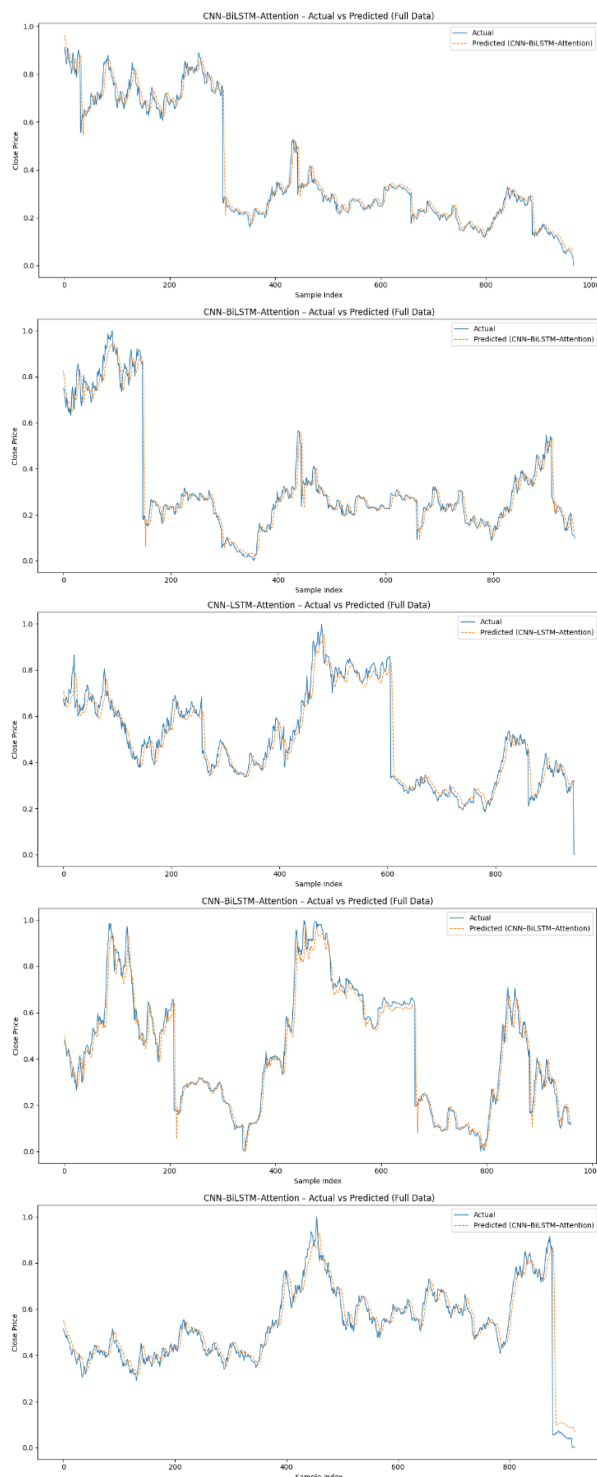
جدول ۴. نتایج پیش‌بینی مدل‌های مختلف برای سهام وبملت

mse	$rmse$	mae	r^2	الگوریتم‌های یادگیری ماشین
۰/۰۰۵۱	۰/۰۷۱۸	۰/۰۵۹۴	۰/۵۱۸۶	LSTM
۰/۰۰۳۰	۰/۰۵۴۹	۰/۰۳۳۱	۰/۷۱۸۸	CNN-LSTM
۰/۰۰۲۸	۰/۰۵۳۴	۰/۰۳۲۶	۰/۷۳۳۶	CNN-LSTM-Attention
۰/۰۰۲۷	۰/۰۵۳۳	۰/۰۳۲۵	۰/۷۳۴۱	CNN-BiLSTM-Attention



نمودار ۳. نمودار داده‌های سری زمانی سهام وبملت

در تحلیل سهام فارس (شرکت فارس و خوزستان)، مدل CNN-BiLSTM-Attention باز هم بهترین عملکرد را نشان داد. این سهم که با ثبات نسبی صنعتی همراه است، مدل‌های پایه دقت نسبتاً خوبی داشتند اما CNN-BiLSTM-Attention باز هم خطاهای اندکی پایین‌تر ارائه نمود. مقایسه آماری نشان داد MAPE و RMSE برای این مدل کمتر از سایرین است. همچنین R^2 به‌ویژه بهبود یافته و نشان‌دهنده

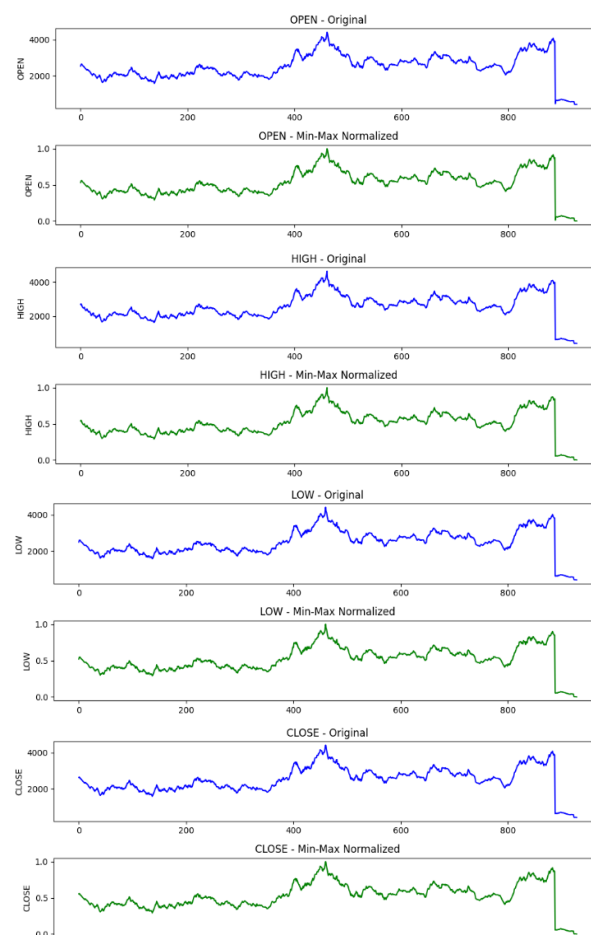


نمودار ۶. نمودار مقادیر واقعی و پیش‌بینی شده سهام‌های فولاد، فملی، وبملت، فارس و خودرو با استفاده از مدل CNN-BiLSTM-Attention

مدل پیشنهادی با ساختار دوطرفه LSTM و توجه توانست تغییرات پیش‌گفته را به خوبی شناسایی کند. بدین‌ترتیب، این مدل در هر دو شاخص MAPE و RMSE کمترین مقدار

جدول ۶. نتایج پیش‌بینی مدل‌های مختلف برای سهام خودرو

الموریت‌های یادگیری ماشین	r^2	mae	rmse	mse
LSTM	۰/۶۴۸۹	۰/۱۲۹۱	۰/۱۶۷۰	۰/۰۲۷۹
CNN-LSTM	۰/۷۰۴۸	۰/۱۱۹۲	۰/۱۵۵۵	۰/۰۲۴۱
CNN-LSTM-Attention	۰/۸۰۳۶	۰/۰۵۳۷	۰/۱۲۶۶	۰/۰۱۶۰
CNN-BiLSTM-Attention	۰/۸۱۰۴	۰/۰۵۲۵	۰/۱۲۴۴	۰/۰۱۵۴



نمودار ۵. نمودار داده‌های سری زمانی سهام خودرو

برای سهم خودرو نیز نتایج مشابهی مشاهده شد؛ یعنی CNN-BiLSTM-Attention بهترین نتایج پیش‌بینی را به‌دست آورد. این سهم که به‌طور معمول تحت تأثیر عوامل اقتصادی کلان نوسانات متوسطی دارد، باعث شد مدل‌های ساده‌تر خطای پیش‌بینی بیشتری داشته باشند. اما

جدول ۷. نتایج ارزیابی پیش‌بینی مدل‌های مختلف با
آزمون Friedman

p-value	فولاد	فملی	وبملت	فارس	خودرو
MSE	۰/۰۲	۰/۰۰۰۰۰۸	۰/۰۱	۰/۰۱	۰/۰۰۰۰۰۱
RMSE	۰/۰۲	۰/۰۰۰۰۰۸	۰/۰۱	۰/۰۱	۰/۰۰۰۰۰۱
MAE	۰/۰۳	۰/۰۰۰۲۲۸	۰/۰۱	۰/۰۱	۰/۰۰۰۰۰۵
R²	۰/۰۲	۰/۰۰۰۰۰۸	۰/۰۱	۰/۰۱	۰/۰۰۰۰۰۱

را نشان داد و ضریب تعیین آن بالاترین بود. آزمون فریدمن نیز تاکید کرد که این تفاوت‌ها در عملکرد کاملاً معنادارند. در مجموع می‌توان گفت که برتری CNN-BiLSTM-Attention در مدل‌سازی الگوهای پیچیده و نوسانات کوتاه‌مدت قیمت، منجر به کاهش خطاهای مطلق و افزایش دقت نسبی در پیش‌بینی قیمت خودرو شده است.

۵- نتیجه‌گیری و ارائه‌ی پیشنهادات

نتایج این پژوهش به وضوح نشان داد که مدل پیشنهادی CNN-BiLSTM-Attention نسبت به سایر ساختارهای آزمایش‌شده، دقت پیش‌بینی بالاتری ارائه می‌کند. معیارهای ارزیابی مدل‌های پیش‌بینی مانند RMSE، MAE و MAPE برای مدل ترکیبی پیشنهادی به طور چشمگیری بهبود یافته است. این برتری در دقت مخصوصاً در مقایسه با مدل LSTM معمولی قابل توجه بود؛ به‌طوری که مطالعات پیشین نیز نشان می‌دهد دقت پیش‌بینی مدل CNN-BiLSTM-Attention به‌طور معنی‌داری بیشتر از مدل‌های LSTM ساده است.

دلیل عملکرد بهتر مدل پیشنهادی را می‌توان در طراحی معماری آن جستجو کرد. لایه‌های هم‌آمیختگی (CNN) ویژگی‌های محلی و کوتاه‌مدت در داده‌های قیمتی را به خوبی استخراج می‌کنند، در حالی که BiLSTM قابلیت یادگیری وابستگی‌های بلندمدت زمانی دوطرفه را دارد. همچنین، سازوکار توجه در لایه‌های بالاتر به مدل اجازه می‌دهد بخش‌هایی از توالی قیمت که بیشترین تأثیر را بر روند آتی دارند، شناسایی شده و وزن بیشتری به آنها

اختصاص یابد. ترکیب این سه جز معماری باعث شده مدل بتواند داده‌های زائد را کنار بگذارد و روی سیگنال‌های کلیدی بازار تمرکز کند که نتیجه آن بهبود دقت پیش‌بینی در شرایط مختلف بازار است.

از نظر کاربردی، بهبود دقت پیش‌بینی قیمت سهام توسط مدل پیشنهادی می‌تواند فواید متعددی داشته باشد. همان‌طور که پیش‌تر گزارش شده است، پیش‌بینی دقیق قیمت سهام به سرمایه‌گذاران کمک می‌کند با زمان‌بندی مناسب معاملات، سود خود را حداکثر کرده و زیان را به حداقل برسانند. بنابراین، به‌کارگیری مدل CNN-BiLSTM-Attention در راهبردهای معاملاتی خودکار می‌تواند منجر به تصمیم‌گیری‌های بهینه‌تر و مدیریت ریسک بهتری در بورس تهران شود. همچنین، این مدل می‌تواند به عنوان ابزاری کمکی در تحلیل فنی و کمی سهام در کارگزاری‌ها و مؤسسات مالی مورد استفاده قرار گیرد.

در پایان، پیشنهادهایی برای مطالعات آتی ارائه می‌شود

- استفاده از داده‌ها و ویژگی‌های اضافی همچون متغیرهای کلان اقتصادی (نرخ تورم، قیمت دلار و ...)
- و شاخص‌های بنیادی شرکت‌ها به منظور غنی‌سازی مجموعه داده ورودی مدل

- افزودن سازوکارهای مکمل جهت شناسایی سریع افت‌های ناگهانی قیمت سهام مطابق مشاهدات مطالعات گذشته و هشداردهی به سرمایه‌گذاران در شرایط بحرانی.
- گسترش ارزیابی مدل پیشنهادی به سایر بازارها و شاخص‌های مالی و همچنین آزمون مدل در بازه‌های زمانی متفاوت (مانند پیش‌بینی کوتاه‌مدت در مقابل بلندمدت).

- آزمایش معماری‌های جدید سازوکار توجه مانند توجه چندسر (Multi-Head Attention) و ترکیب مدل‌های یادگیری عمیق با منابع داده‌ای دیگر (مثلاً اخبار اقتصادی یا شبکه‌های اجتماعی) برای افزایش دقت و قابلیت تبیین مدل.

مراجع

1. Sharma, P. (2016). Forecasting stock market volatility using realized GARCH model: International evidence. The

17. Syed, F., & Xuan, J. (May 2025). An Attention-Augmented BiLSTM-CNN Model for Stock Price Prediction. In 2025 IEEE Conference on Artificial Intelligence (CAI), pp. 139–144, IEEE.
18. Bhanujyothi, H. C., & Jacob, I. J. (2025). A Hybrid CNN-LSTM Attention-Based Deep Learning Model for Stock Price Prediction Using Technical Indicators. *Engineering, Technology & Applied Science Research*, 15(5), 28012–28017.
19. Fatiha, L., Sarkar, S., Bashir, G. M. M., & Sarker, M. (February 2025). Stock Trend Prediction Using a Multivariate Hybrid Attention-Based BiLSTM with 1D Convolution Model. In 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–6. IEEE.
۲. غلامی، نیما؛ و شمس قارنه، ناصر (۱۴۰۳). ارائه مدلی برای پیش بینی قیمت سهام مبتنی بر CNN-LSTM بهینه شده در بورس اوراق بهادار تهران. چشم انداز مدیریت مالی، ۱۴(۴۵)، ۱۲۳–۱۴۷. doi: ۱۰.۴۸۳۰۸/jfmp/۱۰.۴۸۹۲.۲۰۲۴.
3. Wang, X. (2024, September). CNN-BiLSTM-Attention Algorithm-Based Stock Prices Prediction During COVID-19. In 2024 International Conference on Artificial Intelligence and Communication (ICAIC 2024) (pp. 441–452). Atlantis Press.
4. Rafibakhsh, R.; Khorasani, A. M.; Rezazadeh, M. (????). The Application and Effectiveness of Machine Learning and Deep Learning Methods in Analyzing and Predicting the Shanghai Stock Index. (publication details not provided).
5. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175.
6. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge: MIT Press.
7. Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157–166.
8. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
9. Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11), 2673–2681.
10. Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
11. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ..., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
12. 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.
13. Liu, S., & Lan, Q. (2024, December). Long-term stock correlation prediction based on CNN-BiLSTM-Attention models. In Proceedings of the 4th Asia-Pacific Artificial Intelligence and Big Data Forum (pp. 1111–1117).
14. Zhang, J., Ye, L., & Lai, Y. (2023). Stock Price Prediction Using CNN-BiLSTM-Attention Model. *Mathematics*, 11(9), 1985.
15. Zhang, Y., Zhang, T., & Hu, J. (2025). Forecasting Stock Market Volatility Using CNN-BiLSTM-Attention Model with Mixed-Frequency Data. *Mathematics*, 13(11), 1889.
16. Wei, E., & Wang, Z. (March 2025). Stock price prediction using CEEMD-CNN-BiLSTM-AM: A hybrid deep learning approach. In Proceedings of the 2025 5th International Conference on Applied Mathematics, Modelling and Intelligent Computing, pp. 287–295.