E-ISSN: | P-ISSN:

Management Sciences and Decision Analysis

www.msda.reapress.com

Manag. Sci. Decis. Anal. Vol. 3, No. 2 (2025) 59–67.

Paper Type: Original Article

Enhancing Financial Decision-Making through BiLSTM Deep Learning: Evidence from Apple Stock Price Forecasting

Mobina Ghasemi[1], Marzieh Shiri[2,*], Omid Taheri[1]

[1] Department of Accounting, National University of Skills, Tehran, Iran; ghm62119@gmail.com; omid.taheri6080@gmail.com.

[2]Department of Computer Science, National University of Skills, Tehran, Iran; mshiri6631@gmail.com.

Citation:

Ghasemi, Gh., Shiri, M., & Taheri, O. (2025). Investigating the performance of the BiLSTM deep learning model in predicting Apple Stock Prices. *Management sciences and decision analysis, 3*(2), 59-67.

Received: 06/05/2024 Reviewed: 16/07/2024 Revised: 09/10/2024 Accepted: 27/10/2024

Abstract

Purpose: Accurate stock price forecasting is a critical challenge in financial decision-making, given the nonlinear and dynamic nature of financial markets. Although deep learning models have demonstrated promising capabilities in time-series prediction, the impact of activation functions and optimization algorithms on the predictive performance of BiLSTM models has not been sufficiently explored. This study aims to evaluate the effectiveness of the BiLSTM deep learning architecture in stock price forecasting and identify the optimal parameter configuration for improving predictive accuracy.

Methodology: The study employs a deep learning framework based on Bidirectional Long Short-Term Memory (BiLSTM) networks using historical Apple Inc. stock data from 1980 to 2024 obtained from Yahoo Finance. After preprocessing and normalization, the model was implemented in Google Colab using Bidirectional, Dropout, and Dense layers. Multiple combinations of activation functions and optimization algorithms, including ReLU and Adam, were evaluated to analyze their impact on forecasting performance. Model effectiveness was assessed by comparing predictive accuracy and trend-matching capability in time-series forecasting.

Findings: The results indicate that the BiLSTM model effectively captures long-term temporal dependencies and successfully predicts overall stock price trends. Among the evaluated configurations, the combination of the ReLU activation function and Adam optimizer achieved the highest predictive accuracy and demonstrated superior capability in modeling market fluctuations. The proposed architecture showed lower forecasting error and stronger trend-prediction performance than baseline configurations.

Originality/Value: This study provides a focused analytical investigation into the interaction between network configuration parameters and forecasting performance within BiLSTM-based financial prediction models. The findings highlight how optimization strategies and activation mechanisms influence predictive stability and accuracy, offering practical insights for investors, analysts, and researchers seeking more reliable AI-driven financial forecasting tools.

Keywords: Deep learning, Stock price prediction, BiLSTM, Time series analysis, Model optimization.

Corresponding Author: mshiri6631@gmail.com 10.22105/msda.v3i2.86

Licensee. **Management Sciences and Decision Analysis**. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

علوم مدیریت و تحلیل تصمیم

دوره ۳، شماره (۲)، (۱۴۰۴)، ۶۷-۵۹

www.msda.reapress.com

نوع مقاله: پژوهشی

ارتقای تصمیم‌گیری مالی با استفاده از یادگیری عمیق BiLSTM: شواهدی از پیش‌بینی قیمت سهام اپل

مبینا قاسمی[۱]، مرضیه شیری[۱،*]، امید طاهری[۲]

[۱]گروه حسابداری، دانشگاه ملی مهارت، تهران، ایران.

[۲]گروه کامپیوتر، دانشگاه ملی مهارت، تهران، ایران.

چکیده

هدف: پیش‌بینی دقیق قیمت سهام به دلیل ماهیت غیرخطی و پویای بازارهای مالی، یکی از چالش‌های مهم در تصمیم‌گیری مالی محسوب می‌شود. اگرچه مدل‌های یادگیری عمیق قابلیت‌های قابل‌توجهی در پیش‌بینی سری‌های زمانی نشان داده‌اند، تأثیر توابع فعال‌سازی و الگوریتم‌های بهینه‌سازی بر عملکرد پیش‌بینی مدل‌های *BiLSTM* به‌طور کافی بررسی نشده است. هدف این پژوهش، ارزیابی کارایی معماری یادگیری عمیق BiLSTM در پیش‌بینی قیمت سهام و شناسایی بهینه‌ترین ترکیب پارامترها برای افزایش دقت پیش‌بینی است.

روش‌شناسی: این پژوهش از یک چارچوب یادگیری عمیق مبتنی بر شبکه‌های حافظه بلند و کوتاه‌مدت دوطرفه با استفاده از داده‌های تاریخی سهام شرکت اپل در بازه ۱۹۸۰ تا ۲۰۲٤ بهره می‌گیرد. داده‌ها پس از پیش‌پردازش و نرمال‌سازی، در محیط *Google Colab* و با استفاده از لایه‌های *Bidirectional*، *Dropout* و *Dense* پیاده‌سازی شدند. ترکیب‌های مختلفی از توابع فعال‌سازی و الگوریتم‌های بهینه‌سازی از جمله *ReLU* و *Adam* برای تحلیل اثر آن‌ها بر عملکرد مدل مورد ارزیابی قرار گرفتند. کارایی مدل از طریق دقت پیش‌بینی و توانایی تطبیق روندها در پیش‌بینی سری‌های زمانی بررسی شد.

یافته‌ها: نتایج نشان داد مدل *BiLSTM* توانایی بالایی در استخراج وابستگی‌های زمانی بلندمدت و پیش‌بینی روند کلی قیمت سهام دارد. در میان ترکیب‌های بررسی‌شده، ترکیب تابع فعال‌سازی *ReLU* و الگوریتم *Adam* بالاترین دقت پیش‌بینی را به دست آورد و عملکرد بهتری در مدل‌سازی نوسانات بازار نشان داد. معماری پیشنهادی نسبت به پیکربندی‌های پایه دارای خطای پیش‌بینی کمتر و توانایی بالاتر در پیش‌بینی روندها بود.

اصالت/ارزش افزوده علمی: این پژوهش یک تحلیل متمرکز از تعامل میان پارامترهای پیکربندی شبکه و عملکرد پیش‌بینی در مدل‌های مالی مبتنی بر *BiLSTM* ارایه می‌دهد. یافته‌ها نشان می‌دهد راهبردهای بهینه‌سازی و سازوکارهای فعال‌سازی چگونه بر پایداری و دقت پیش‌بینی اثر می‌گذارند و بینش‌های کاربردی ارزشمندی برای سرمایه‌گذاران، تحلیلگران و پژوهشگران در توسعه ابزارهای پیش‌بینی مالی مبتنی بر هوش مصنوعی فراهم می‌سازند.

کلیدواژه‌ها: یادگیری عمیق، پیش‌بینی قیمت سهام، تحلیل داده‌های مالی، مدل *BiLSTM*، تابع فعال‌سازی، الگوریتم بهینه‌سازی.

۱- مقدمه

پیش‌بینی قیمت سهام، تکنیکی است که به منظور پیش‌بینی قیمت‌های آینده سهام با استفاده از داده‌های قبلی بازار سهام و اطلاعات مرتبط به کار می‌رود. اهمیت پیش‌بینی قیمت سهام از آن‌جاست که می‌تواند رویکردهای مفیدی برای سرمایه‌گذاران و معامله‌گران فراهم کند تا تصمیمات آگاهانه‌تری در بازار سهام اتخاذ کنند. مساله پیش‌بینی قیمت سهام به‌عنوان یکی از پیچیده‌ترین و مهم‌ترین وظایف در حوزه‌های مالی و اقتصادی

شناخته می‌شود. بازار سهام به دلیل پیچیدگی و دینامیک بودنش تحت تاثیر عوامل گوناگونی قرار دارد که شامل شاخص‌های اقتصادی، اخبار خاص شرکت‌ها، احساسات سرمایه‌گذاران و حتی شرایط بحرانی مانند پاندمی‌ها[1] هستند [1].
عملکرد بازار سهام معمولا به‌عنوان نمایانگر سلامت کلی یک اقتصاد درنظر گرفته می‌شود؛ بنابراین، پیش‌بینی قیمت سهام تاثیرات متعددی بر اقتصاد کلان دارد. پیش‌بینی‌های دقیق قیمت سهام می‌توانند به‌طور موثری عملکرد آینده شرکت‌ها و به‌طورکلی اقتصاد را پیش‌بینی کنند که این اطلاعات می‌تواند در تصمیم‌گیری‌های سرمایه‌گذاری و سیاست‌گذاری مورداستفاده قرار گیرد.

پیش‌بینی قیمت‌ها می‌تواند بر تصمیمات سرمایه‌گذاری تاثیر بگذارد و در نتیجه بر اقتصاد کلان اثرگذار باشد. اگر پیش‌بینی‌های قیمت سهام نشان‌دهنده عملکرد خوب یک صنعت یا گروهی از شرکت‌ها در آینده باشد، ممکن است سرمایه‌گذاری بیشتری در این شرکت‌ها صورت گیرد که می‌تواند به رشد اقتصادی در آن منطقه کمک کند. برعکس، اگر پیش‌بینی‌ها حاکی از عملکرد ضعیف یک صنعت باشد، ممکن است منجر به کاهش سرمایه‌گذاری در آن بخش و افت فعالیت‌های اقتصادی شود [2]، [3].

پیش‌بینی‌های قیمت سهام همچنین می‌توانند بر تصمیمات سیاست‌گذاری پولی بانک‌های مرکزی تاثیر بگذارند به‌عنوان‌مثال، اگر انتظارات بازار سهام نشان دهد که اقتصاد ممکن است دچار رکود شود، بانک مرکزی می‌تواند با کاهش نرخ‌های بهره به تحریک توسعه اقتصادی و تشویق به وام‌گیری و هزینه کرد بپردازد. از طرف دیگر، اگر پیش‌بینی‌ها نشان‌دهنده خطر تورم و رونق بیش‌ازحد اقتصاد باشد، ممکن است بانک مرکزی نرخ‌های بهره را افزایش دهد تا از این وضعیت جلوگیری کند [4].

پیش‌بینی قیمت سهام، به‌عنوان یکی از چالش‌های اساسی در حوزه مالی، اهمیت زیادی برای سرمایه‌گذاران، تحلیل‌گران و تصمیم‌گیرندگان اقتصادی دارد. بازارهای مالی به دلیل ناپایداری و پیچیدگی‌های ذاتی، نیازمند مدل‌های پیشرفته و دقیق برای تحلیل داده‌ها و پیش‌بینی روندهای آینده هستند. در سال‌های اخیر، استفاده از تکنیک‌های یادگیری ماشین[2] و یادگیری عمیق، به‌ویژه مدل‌های شبکه عصبی، به‌طور چشم‌گیری در این حوزه رشد کرده است.

یادگیری عمیق یکی از شاخه‌های مهم یادگیری ماشین است که به‌ویژه در پردازش داده‌های پیچیده و غیرخطی کاربرد فراوانی دارد. یکی از مدل‌های پیشرفته در این زمینه، شبکه‌های عصبی طولانی-کوتاه مدت[3] هستند که به‌طور خاص برای مدیریت و تحلیل داده‌های ترتیبی طراحی شده‌اند. *LSTMs* به خاطر قابلیت‌های خاص خود در حفظ و یادآوری اطلاعات در طول زمان، به‌ویژه در وظایفی مانند شناسایی گفتار، ترجمه زبان و پیش‌بینی سری‌های زمانی، محبوبیت زیادی پیدا کرده‌اند. این شبکه‌ها با استفاده از ساختارهای خاصی که شامل دروازه‌ها و حافظه‌های داخلی است، می‌توانند به‌خوبی با مشکلاتی مانند ناپایداری گرادیان مقابله کنند و به‌این‌ترتیب، کارایی بیشتری در تجزیه‌وتحلیل داده‌های توالی‌دار دارند.

مدل‌های حافظه بلندمدت و کوتاه‌مدت و حافظه بلندمدت و کوتاه‌مدت دوطرفه به‌عنوان ابزارهای موثر در پردازش داده‌های زمانی شناخته شده‌اند. *LSTM* به دلیل توانایی خود در حفظ اطلاعات در طول زمان، به ویژه برای داده‌های غیرخطی و وابسته به زمان، کاربرد زیادی دارد. *BiLSTM*، با پردازش اطلاعات به‌صورت دو طرفه، قابلیت بهتری در درک الگوهای پیچیده و وابستگی‌های طولانی‌مدت در داده‌ها دارد.

با توجه به اهمیت پیش‌بینی قیمت سهام در اقتصاد، هدف این مطالعه تمرکز بر تحلیل اکتشافی راهکارهای موثر برای این مساله است. این پژوهش به بررسی و مقایسه عملکرد مدل‌های *LSTM* و *BiLSTM* در پیش‌بینی قیمت سهام می‌پردازد و به دنبال شناسایی بهترین روش و تابع فعال‌سازی مناسب برای مدل *BiLSTM* است.

۲-۱- تعریف مساله

هدف این مقاله، بررسی عملکرد مدل *BiLSTM* در پیش‌بینی قیمت سهام و ارزیابی تاثیر تابع فعال‌سازی و روش‌های بهینه‌سازی بر روی آن است. با شناسایی بهترین ترکیب‌های این پارامترها، می‌توان به بهبود دقت پیش‌بینی[4] و افزایش کارایی مدل‌های مالی دست یافت. این تحقیق می‌تواند به

[1] Pandemics
[2] Machine learning
[3] Long Short-Term Memory Models (LSTM)
[4] Accuracy

سرمایه‌گذاران و تحلیل‌گران در اتخاذ تصمیم‌های آگاهانه کمک کند و به‌عنوان مبنایی برای تحقیقات آینده در زمینه پیش‌بینی قیمت سهام با استفاده از یادگیری عمیق عمل کند.

۳-۱- اهمیت موضوع

پیش‌بینی قیمت سهام به‌عنوان یک موضوع کلیدی در حوزه مالی و اقتصادی، اهمیت زیادی دارد. این اهمیت از جنبه‌های مختلف قابل‌بررسی است:

۱. تصمیم‌گیری سرمایه‌گذارها: پیش‌بینی‌های دقیق قیمت سهام می‌تواند به سرمایه‌گذاران کمک کند تا تصمیمات بهتری در خصوص خریدوفروش سهام اتخاذ کنند. اطلاعات دقیق درباره روندهای آینده می‌تواند ریسک سرمایه‌گذاری را کاهش دهد و منجر به بازدهی بالاتر شود.
۲. تحلیل بازار: درک نحوه عملکرد بازار سهام و عوامل موثر بر قیمت‌ها به تحلیل‌گران کمک می‌کند تا شرایط بازار را بهتر ارزیابی کنند. این تحلیل‌ها می‌توانند به پیش‌بینی نوسانات بازار و شناسایی الگوهای رفتاری سرمایه‌گذاران منجر شوند.
۳. تاثیر بر اقتصاد کلان: پیش‌بینی قیمت سهام می‌تواند تاثیرات قابل‌توجهی بر تصمیمات اقتصادی و سیاست‌گذاری‌های کلان داشته باشد. به‌عنوان‌مثال، پیش‌بینی مثبت از قیمت‌ها می‌تواند منجر به افزایش سرمایه‌گذاری در صنایع خاص و رشد اقتصادی شود، درحالی‌که پیش‌بینی منفی ممکن است به کاهش سرمایه‌گذاری و افت اقتصادی منجر گردد.
٤. مدیریت ریسک: شرکت‌ها و سرمایه‌گذاران می‌توانند از پیش‌بینی‌های قیمت سهام برای مدیریت ریسک‌های مالی خود استفاده کنند. با شناسایی نقاط قوت و ضعف در بازار، می‌توانند استراتژی‌های مناسبی برای محافظت از دارایی‌های خود ایجاد کنند.
۵. پیش‌بینی روندهای اقتصادی: پیش‌بینی قیمت سهام می‌تواند به‌عنوان معیاری برای ارزیابی سلامت کلی اقتصاد عمل کند. عملکرد بازار سهام معمولا به‌عنوان نشانه‌ای از شرایط اقتصادی درنظر گرفته می‌شود و پیش‌بینی‌های دقیق می‌تواند به شناسایی روندهای اقتصادی آینده کمک کند.

۲- مروری بر کارهای پیشین

در سال‌های اخیر، یادگیری ماشین، به ویژه یادگیری عمیق، به‌عنوان ابزاری قوی برای پیش‌بینی قیمت سهام شناخته شده است. این فناوری در انواع مختلفی از وظایف مربوط به پیش‌بینی قیمت سهام، از جمله پیش‌بینی قیمت‌های آینده، شناسایی روندها و کشف الگوهای داده‌ای مورداستفاده قرار گرفته است [5].

پاتل و همکاران [6] دو روش را با استفاده از چهار مدل سنتی یادگیری ماشین بررسی کردند. در روش اول، داده‌های معاملاتی سهام برای شناسایی و ۱۰ محدودیت تکنیکی مورداستفاده قرار می‌گیرد، درحالی‌که روش دوم بر بازتولید این ویژگی‌ها با استفاده از داده‌های تصمیم‌گیری ثابت تمرکز دارد. دقت هر یک از این مدل‌های پیش‌بینی برای هر کدام از این دو روش ورودی ارزیابی می‌شود.

ژانگ و همکاران [7]، معماری شبکه‌های مولد تخاصمی[1] را معرفی کردند که در آن از حافظه بلندمدت کوتاه‌مدت[2] به‌عنوان مولد و از پرسپترون چندلایه[3] به‌عنوان تمایزدهنده استفاده شده است. این مدل *GAN* با *LSTM*، شبکه عصبی مصنوعی[4] و رگرسیون برداری پشتیبان[5] مقایسه شده است. برای ارزیابی این مدل‌ها از معیارهای متعددی بهره‌برداری شده و نتایج نشان می‌دهد که مدل *GAN* پیشنهادی در مقایسه با سایر مدل‌ها برتری دارد، بر اساس تمامی معیارهایی که در این مقاله استفاده شده است.

پتیل و همکارن [8] سه مدل شامل رگرسیون برداری پشتیبان، رگرسیون خطی[6] و حافظه بلندمدت کوتاه‌مدت را معرفی کردند. نتایج نشان می‌دهد که *LSTM* به‌طرز قابل‌توجهی از سایر مدل‌ها بهتر عمل کرده و امتیاز ۰٫۰۱۵۱ را کسب کرده است؛ درحالی‌که *LR* با امتیاز ۱۳٫۸۷۲ در مقام دوم قرار دارد و *SVR* با امتیاز ۳٤٫٦۲۳ در آخرین جایگاه قرار دارد.

---

[1] Generative Adversarial Network (GAN)
[2] Long Short-Term Memory (LSTM)
[3] Multilayer Perceptron (MLP)
[4] Artificial Neural Network (ANN)
[5] Support Vector Regression (SVR)
[6] Linear Regression (LR)

رانا و همکاران [9] دو مدل مبتنی بر نظریه گراف را پیشنهاد کردند. مدل اول بر اساس همبستگی بین قیمت‌های تاریخی و مدل دوم بر اساس علیت طراحی شده است. نتایج این تحقیق نشان داد که مدل‌های مبتنی بر گراف نسبت به روش‌های سنتی برتری دارند و مدل مبتنی بر علیت نتایج بهتری نسبت به مدل همبستگی به دست آورد.

دی پرسیو و همکارن [10] مدل‌های پایه *RNN*،*LSTM* و *GRU*[1] معرفی کردند. مدل *GRU* به دقت ۰٫۶۷ و ضرر لاگ ۰٫۶۲۹ دست یافت، درحالی‌که *LSTM* دقت ۰٫۶۶۵ و ضرر لاگ ۰٫۶۲۹ داشت و *RNN* با دقت ۶۲٫۵ و ضرر لاگ[2] ۰٫۷۲۵ در مقام آخر قرار گرفت. هر دو مدل *LSTM* و *GRU* با افزودن یک لایه دراپ اوت تغییراتی را تجربه کردند؛ اما مدل *GRU* هیچ بهبودی نشان نداد، درحالی‌که *LSTM* بهبود جزیی در عملکرد به میزان ۲% داشت.

روندی والا و همکاران [11] مدل *LSTM* را برای پیش‌بینی قیمت‌های سهام *Nifty 50* پیشنهاد کردند. *LSTM* یک معماری *RNN* است که به‌طور گسترده‌ای در پردازش زبان طبیعی (*NLP*) مورداستفاده قرار می‌گیرد. نتایج این تحقیق نشان داد که با افزایش تعداد پارامترها و دوره‌های آموزشی[3]، عملکرد مدل بهبود می‌یابد و *LSTM* توانست بهترین عملکرد را با مقدار ۰٫۰۰۸۵۹ در معیار *RMSE* با استفاده از مجموعه پارامترهای *High*، *Low*، *Open* و *Close* و ۵۰۰ دوره آموزشی به‌دست آورد.

هیرانشا و همکارن [12] چهار مدل یادگیری عمیق شامل *MLP*، *RNN*،*CNN* و *LSTM* معرفی کردند که این مدل‌ها بر روی داده‌های شرکت *TATA MOTORS* آموزش دیده‌اند. پس از فرآیند آموزش، مدل‌ها با پیش‌بینی قیمت‌های سهام مورد ارزیابی قرار گرفتند و نتایج رضایت‌بخش به‌دست آمد. این مدل‌ها حتی توانستند الگوهای حرکات سهام را در بازارهای دیگر شناسایی کنند که نشان‌دهنده توانایی مدل‌های یادگیری عمیق در شناسایی دینامیک‌های نهفته است. در این میان، مدل *CNN* از سایر مدل‌ها برتری بیشتری نشان داد. همچنین، این مقاله مدل *ARIMA*[4] را آزمایش کرد، اما نتوانست دینامیک‌های موجود بین چندین سری زمانی را به‌خوبی یاد بگیرد.

مک نالی و همکاران [13] مدل جدیدی در سال ۲۰۱۸ برای پیش‌بینی قیمت بیت‌کوین پیشنهاد کردند که مشابه پیش‌بینی قیمت سهام عمل می‌کند. در این تحقیق، سه مدل یادگیری عمیق شامل *RNN* ساده، *LSTM* و *ARIMA* موردبررسی قرار گرفته‌اند. نتایج نشان می‌دهد که این سه مدل در زمینه دقت عملکرد مشابهی دارند؛ به‌طوری‌که *LSTM* با دقت ۵۲/۷۸%، *RNN* با دقت ۵۰/۲۵% و *ARIMA* با دقت ۵۰/۰۵% عمل کرده‌اند. بااین‌حال، در معیار *RMSE*[5]، دو مدل یادگیری عمیق به طرز چشمگیری از مدل *ARIMA* پیشی گرفته‌اند؛ به‌طوری‌که *LSTM* و *RNN* به‌ترتیب امتیازهای ۶/۸۷% و ۵/۴۵% را کسب کردند، درحالی‌که *ARIMA* با *RMSE* معادل ۵۳/۷۴% عملکرد ضعیفی داشت.

چانگ و همکاران [14] از یک مدل جدید یادگیری عمیق با استفاده از *CNN* ساده استفاده کردند، شبکه عصبی مصنوعی و یک مدل *CNN* که با استفاده از الگوریتم ژنتیک بهبود یافته است. این مدل جدید به‌منظور بهبود دقت پیش‌بینی قیمت بیت‌کوین طراحی شده و انتظار می‌رود نتایج بهتری نسبت به مدل‌های قبلی ارایه دهد.

ژو و همکاران [15] مدل جدیدی به نام *WAE-BLSTM* در سال ۲۰۲۰ معرفی کردند که این مدل دارای یک روند کاری سه مرحله‌ای است که شامل حذف نویز، کاهش ابعاد و پیش‌بینی با استفاده از *BLSTM* می‌باشد. برای نشان دادن قابلیت‌های این مدل، آن را با چهار مدل دیگر مقایسه کردند: *W-BLSTM*، *W-LSTM*، *BLSTM* و *LSTM* نتایج نشان می‌دهد که *WAE-BLSTM* در هر دو معیار *MAE* و *RMSE* از سایر مدل‌ها برتر عمل کرده است.

لو و همکاران [16] مدل *CNN-BiLSTM-AM* را معرفی کردند که از *CNN*، *BiLSTM* و مکانیزم توجه بهره می‌برد. در این مدل،*CNN* ویژگی‌ها را استخراج می‌کند، *BiLSTM* برای پیش‌بینی با استفاده از این ویژگی‌ها به‌کار گرفته می‌شود و مکانیزم توجه تاثیر ویژگی‌های استخراج‌شده را

[1] Gated recurrent unit
[2] Log loss
[3] Training epochs
[4] Auto regressive integrated moving average
[5] Mean absolute error

موردبررسی قرار می‌دهد. این مدل در مقایسه با *Bi-LSTM-AM*، *CNN-BiLSTM*، *CNN-LSTM*، *BiLSTM*، *LSTM*، *CNN*، *RNN* و *MLP* برتری خود را بر اساس معیارهای *MAE*[1] و *RMSE* نشان داده است.

وانگ و همکاران [17] شبکه عصبی *Elman* را پیشنهاد کردند که یک شبکه عصبی مبتنی بر *RNN* است. این شبکه به‌طور خاص برای کاربردهای مختلف در یادگیری عمیق و پردازش داده‌ها طراحی شده و قابلیت‌های منحصربه‌فردی را در تحلیل داده‌های زمانی ارایه می‌دهد.

دوتا و همکاران [18] مدل‌های مختلف *GRU* را برای پیش‌بینی قیمت بیت‌کوین معرفی کردند که این مدل‌ها با *LSTM* و شبکه عصبی مصنوعی مقایسه شده‌اند. مدل‌های *GRU* شامل *GRU* پایه، مدل *GRU-Dropout* و مدل *GRU-Dropout-GRU* بودند. نتایج نشان می‌دهد که *GRU* پایه در مقایسه با دو مدل دیگر GRU،*LSTM* و *ANN* عملکرد بهتری داشته و توانسته *RMSE* کم‌تری را به‌دست آورد.

کیو و همکاران [19] مدل *LSTM* مبتنی بر توجه که به آن *AWLSTM* گفته می‌شود را معرفی کردند که این مدل از تبدیل‌ها برای حذف نویز داده‌ها استفاده می‌کند و با *WLSTM*، *LSTM* و مدل‌های *GRU* مقایسه شده است. برای ارزیابی مدل‌ها، از سه مجموعه داده و چهار معیار استفاده شده که شامل *S&P 500*،*DIJA* و *HSI* می‌شود. نتایج نشان می‌دهد که *AWLSTM* در مقایسه با سایر مدل‌ها برتری دارد و از نظر چهار معیار موردبررسی، عملکرد بهتری را ارایه کرده است.

۳- مجموعه داده

در این مقاله از مجموعه داده سهام اپل که شامل اطلاعات تاریخی از سال ۱۹۸۰ تا ۲۰۲۴ است، استفاده شده است. این مجموعه داده شامل قیمت‌های آغازین، بالاترین قیمت، پایین‌ترین قیمت، قیمت پایانی تعدیل‌شده و همچنین حجم سهام این شرکت است. داده‌ها از *Yahoo Finance* استخراج و در قالب فایل *CSV* ذخیره شده‌اند. اطلاعات به‌روزرسانی‌شده به‌صورت روزانه در طول این سال‌ها جمع‌آوری شده است.

قیمت آغازین: این قیمت نشان‌دهنده شروع معاملات یک سهام در روز جاری است و بر اساس قیمت پایانی روز قبل و عوامل دیگر بازار تعیین می‌شود.

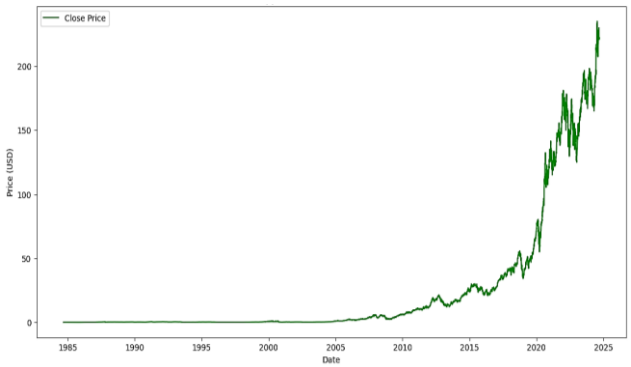
قیمت پایانی: این قیمت آخرین معامله‌ای است که در طول یک روز معاملاتی انجام می‌شود. این قیمت در انتهای روز و بر اساس آخرین معامله ثبت‌شده محاسبه می‌شود و برای سنجش تغییرات خالص و درصدی ارزش سهام در طول روز استفاده می‌شود.

بالاترین قیمت: به بالاترین قیمتی اشاره دارد که یک سهام در یک روز خاص یا در یک دوره زمانی مشخص به ثبت رسیده است. این قیمت به معامله‌گران و تحلیلگران کمک می‌کند تا عملکرد سهام را بررسی کنند.

قیمت پایانی تعدیل‌شده: این قیمت به منظور درنظر گرفتن تغییرات ناشی از رویدادهای شرکتی نظیر تقسیم سهام و سود تقسیمی محاسبه می‌شود. این نوع قیمت معمولا نمای دقیق‌تری از عملکرد سهام نسبت به قیمت پایانی عادی ارایه می‌دهد.

[1] Mean absolute error

حجم سهام: به تعداد سهام معامله‌شده در یک روز معاملاتی خاص یا در یک بازه زمانی اشاره دارد. حجم سهام برای نشان دادن میزان فعالیت و نقدینگی یک سهام استفاده می‌شود و می‌تواند یک معیار مفید برای ارزیابی عملکرد سهام و پیش‌بینی حرکات قیمت آینده باشد.



شکل ۱- قیمت پایانی تعدیل‌شده سهام اپل (Apple Inc.) از سال ۱۹۸۰ تا ۲۰۲۴.
Figure 1- Adjusted closing price of Apple Inc. stock from 1980 to 2024.

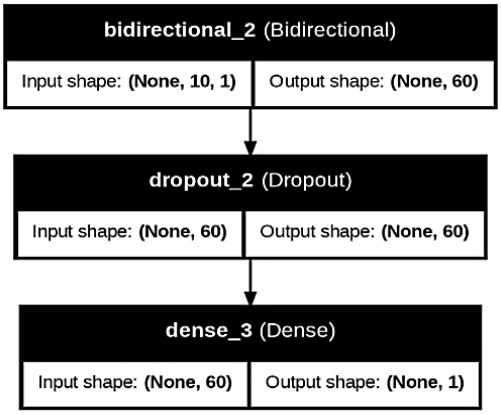
۴- پیاده‌سازی

مجموعه داده موردنظر دارای حجم بسیار بالا و ویژگی‌های دارای اهمیت بالا است. برای انجام آزمایشات مختلف، نیاز به یک سیستم قدرتمند و منابع پردازشی بالا داریم. به همین دلیل، ویژگی‌های مهم که شامل تعداد مناسبی از رکوردها است را حفظ می‌کنیم. برای اجرای عملیات لازم، از محیط *Google Colab* استفاده می‌کنیم. همچنین، در پیش‌پردازش داده‌ها از کتابخانه‌های مختلفی استفاده می‌کنیم. مراحل عملیاتی برای پیش‌پردازش داده‌ها را می‌توان به شرح زیر خلاصه کرد:

۱. نمایش داده‌ها: با استفاده از کتابخانه‌های موجود، چند سطر از دیتافریم نمایش داده می‌شود تا به‌راحتی ساختار و مقادیر اولیه موجود در داده‌ها را مشاهده کنیم.
۲. انتخاب ستون برای نرمال‌سازی: ستونی که شامل قیمت بسته شدن سهام است (معمولا با نام *Close* شناخته می‌شود) انتخاب می‌شود و به یک آرایه عددی تبدیل می‌گردد. این مرحله برای آماده‌سازی داده‌ها جهت پردازش‌های بعدی انجام می‌شود.
۳. ایجاد شی نرمال‌سازی: یک شی از نوع نرمال‌ساز (*MinMaxScaler*) ایجاد می‌شود که وظیفه نرمال‌سازی داده‌ها را بر عهده دارد. این نرمال ساز داده‌ها را به یک بازه مشخص (معمولا [۰,۱]) تغییر می‌دهد.
۴. نرمال‌سازی داده‌ها: داده‌های انتخاب‌شده با استفاده از نرمال‌ساز به بازه مشخصی نرمال‌سازی می‌شوند. این عمل به بهبود عملکرد مدل‌های یادگیری ماشین کمک می‌کند و باعث می‌شود داده‌ها در مقیاس یکسانی قرار بگیرند.

در سال ۱۹۹۷، مقاله‌ای با عنوان «شبکه‌های عصبی بازگشتی دوطرفه» منتشر شد که برای نخستین بار مفهوم *LSTM* دوطرفه *BiLSTM* را معرفی کرد. نویسندگان این تحقیق، معماری *BiLSTM* را به‌عنوان یک راهکار برای بهبود عملکرد در شناسایی صدا پیشنهاد کردند. این مدل بر پایه استفاده از دو شبکه *LSTM* بنا شده است، یکی که به جلو حرکت می‌کند و دیگری که به عقب. درحالی‌که *LSTM* رو به عقب، توالی ورودی را از راست به چپ می‌خواند، *LSTM* رو به جلو آن را از چپ به راست می‌خواند. این قابلیت به *BiLSTM* اجازه می‌دهد تا با ترکیب خروجی‌های هر دو *LSTM*، پیش‌بینی‌هایی انجام دهد و در نتیجه هم زمینه‌های گذشته و هم آینده را درنظر بگیرد [20].

برای پیش‌بینی قیمت سهام، می‌توان داده‌های تاریخی بازار را به‌عنوان ورودی به شبکه *BiLSTM* وارد کرد. ابتدا، داده‌های تاریخی شامل قیمت‌های سهام، حجم معاملات و سایر شاخص‌های مالی جمع‌آوری می‌شود. پس از پیش‌پردازش این داده‌ها، شبکه *BiLSTM* به تجزیه‌وتحلیل آن‌ها در دو جهت جلو و عقب می‌پردازد.



شکل ۲- معماری مدل پیشنهادی.
Figure 2- Architecture of the proposed model.

همان‌طور که در **شکل ۲** مشخص است، یک معماری شبکه عصبی موثر را نشان می‌دهد که برای مساله پیش‌بینی تک متغیره طراحی شده است.

اجزای مدل و عملکرد آن‌ها

لایه *Bidirectional*

۱. نقش: این لایه یک لایه بازگشتی دو طرفه[1] است که ورودی را هم از چپ به راست و هم از راست به چپ پردازش می‌کند. این کار باعث می‌شود که مدل بتواند هم الگوهای زمانی و هم الگوهای مکانی را در داده‌ها شناسایی کند.
۲. شکل ورودی (*None, 10, 1*) :به این معنی است که ورودی یک دنباله با طول ۱۰ است و هر عنصر در دنباله یک ویژگی دارد.
۳. شکل خروجی (None, 60): نشان می‌دهد که خروجی این لایه یک بردار با ۶۰ ویژگی است که حاوی اطلاعات استخراج‌شده از ورودی است.

لایه *Dropout*

۱. نقش: این لایه با حذف تصادفی برخی از نورون‌ها در طول آموزش، از بیش برازش مدل جلوگیری می‌کند. این کار باعث می‌شود که مدل به ویژگی‌های کلی داده‌ها حساس‌تر شود و کمتر به نویزهای تصادفی حساس باشد.
۲. شکل ورودی و خروجی: همان شکل لایه قبلی است زیرا این لایه صرفا برخی از نورون‌ها را حذف می‌کند و به ساختار داده‌ها تغییری نمی‌دهد.

لایه *Dense*

۱. نقش: این یک لایه کاملا متصل است که به‌عنوان لایه خروجی عمل می‌کند. این لایه وزن‌های قابل تنظیم دارد که به مدل اجازه می‌دهد تا یک نگاشت خطی بین ورودی و خروجی ایجاد کند.
۲. شکل ورودی (*None, 60*): همان شکل خروجی لایه قبلی است.

[1] Recurrent Neural Network (RNN)

۳. شکل خروجی (*None, 1*): نشان می‌دهد که خروجی نهایی یک مقدار اسکالر است که احتمالا یک پیش‌بینی برای متغیر هدف است.

این مدل برای مسایلی مانند پیش‌بینی سری‌های زمانی، طبقه‌بندی متون کوتاه یا هر مساله دیگری که در آن ترتیب داده‌ها مهم است، مناسب می‌باشد. لایه *Bidirectional* به مدل کمک می‌کند تا اطلاعات مهمی را از کل دنباله استخراج کند و لایه *Dropout* از بیش برازش مدل جلوگیری می‌کند. لایه *Dense* نهایی نیز پیش‌بینی نهایی را بر اساس ویژگی‌های استخراج‌شده تولید می‌کند.

ابعاد لایه‌ها: ابعاد لایه‌ها (مانند ۱۰ در لایه ورودی و ٦٠ در لایه‌های بعدی) معمولا بر اساس اندازه داده‌ها و پیچیدگی مساله انتخاب می‌شوند.
توابع فعال‌سازی: در این شکل، توابع فعال‌سازی لایه‌ها ذکر نشده است. انتخاب تابع فعال‌سازی مناسب می‌تواند بر عملکرد مدل تاثیر زیادی داشته باشد.
الگوریتم بهینه‌سازی: برای آموزش این مدل، باید یک الگوریتم بهینه‌سازی مانند گرادیان نزولی تصادفی[1] یا *Adam* استفاده شود.

شکل ۳- مقایسه مدل پیشنهادی با میانگین متحرک سهام.

Figure 3- Comparison of the proposed model with the stock moving average.

۵- نتیجه‌گیری

شبکه‌های عصبی بازگشتی دو جهته به‌عنوان یکی از ابزارهای پیشرفته در پیش‌بینی سری‌های زمانی، به‌ویژه قیمت سهام، نتایج قابل‌توجهی را به نمایش گذاشته‌اند. این مدل‌ها با توانایی پردازش توالی‌های زمانی و در نظر گرفتن همبستگی‌های موجود در داده‌های گذشته و آینده، قادر به شناسایی الگوهای نهفته در داده‌های تاریخی هستند و می‌توانند روندهای آینده قیمت‌ها را با دقت قابل قبولی پیش‌بینی کنند. ارزیابی نتایج مدل *BiLSTM* نشان می‌دهد که این مدل توانسته است روند کلی قیمت سهام اپل را به‌خوبی شبیه‌سازی کند و در دوره‌های رشد قیمتی، عملکرد موثری از خود نشان داده است.

با توجه به پتانسیل بالای مدل‌های یادگیری عمیق مانند *BiLSTM*، این تکنیک‌ها ابزارهای ارزشمندی برای سرمایه‌گذاران و تحلیلگران مالی به شمار می‌آیند. استفاده از این مدل‌ها می‌تواند به تسهیل تصمیم‌گیری‌های مالی و بهبود استراتژی‌های سرمایه‌گذاری کمک کند. همچنین، استمرار تحقیقات در این زمینه و توسعه مدل‌های پیچیده‌تر می‌تواند به ارتقا دقت پیش‌بینی‌ها و ارایه بینش‌های عمیق‌تر در بازارهای مالی منجر شود. به‌این‌ترتیب، مدل‌های یادگیری عمیق می‌توانند به‌عنوان ابزاری کارآمد در جهت شناسایی روندهای قیمتی و بهبود شفافیت در بازارهای مالی عمل کنند.

کارهای آتی

توسعه مدل برای پیش‌بینی چندین گام جلوتر: در حال حاضر، مدل *BiLSTM* تنها یک گام جلوتر را پیش‌بینی می‌کند. می‌توان مدل را توسعه داد تا بتواند چندین گام جلوتر را با دقت بیشتری پیش‌بینی کند.

[1] Stochastic Gradient Descent (SGD)

استفاده از مدل برای تصمیم‌گیری در سرمایه‌گذاری: می‌توان از خروجی مدل *BiLSTM* به‌عنوان سیگنالی برای تصمیم‌گیری در مورد خرید یا فروش سهام استفاده کرد.

تطبیق مدل برای بازارهای مختلف: می‌توان مدل را برای پیش‌بینی قیمت سهام در بازارهای مختلف و صنایع مختلف تطبیق داد.

منابع

[1] Hondroyiannis, G., & Papapetrou, E. (2001). Macroeconomic influences on the stock market. *Journal of economics and finance*, *25*(1), 33–49. https://doi.org/10.1007/BF02759685

[2] Mazur, M., Dang, M., & Vega, M. (2021). COVID-19 and the march 2020 stock market crash. Evidence from S&P1500. *Finance research letters*, *38*, 101690. https://doi.org/10.1016/j.frl.2020.101690

[3] Masoud, N. M. H. (2013). The impact of stock market performance upon economic growth. *International journal of economics and financial issues*, *3*(4), 788–798. https://www.econjournals.com/index.php/ijefi/article/view/557

[4] Li, Y. D., Işcan, T. B., & Xu, K. (2010). The impact of monetary policy shocks on stock prices: Evidence from Canada and the United States. *Journal of international money and finance*, *29*(5), 876–896. https://doi.org/10.1016/j.jimonfin.2010.03.008

[5] Ampomah, E. K., Qin, Z., & Nyame, G. (2020). Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement. *Information*, *11*(6), 332. https://doi.org/10.3390/info11060332

[6] Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert systems with applications*, *42*(1), 259–268. https://doi.org/10.1016/j.eswa.2014.07.040

[7] 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. https://doi.org/10.1016/j.procs.2019.01.256

[8] Patil, P., Wu, C. S. M., Potika, K., & Orang, M. (2020). Stock market prediction using ensemble of graph theory, machine learning and deep learning models. *Proceedings of the 3rd international conference on software engineering and information management*. (pp. 85–92) Sydney, Australia. https://doi.org/10.1145/3378936.3378972

[9] Rana, M., Uddin, M. M., & Hoque, M. M. (2019). Effects of activation functions and optimizers on stock price prediction using lstm recurrent networks. *Proceedings of the 2019 3rd international conference on computer science and artificial intelligence* (pp. 354–358) Beijing, China. https://doi.org/10.1145/3374587.337462

[10] Di Persio, L., & Honchar, O. (2017). Recurrent neural networks approach to the financial forecast of google assets. *International journal of mathematics and computers in simulation*, *11*, 7–13. https://iris.univr.it/handle/11562/959057

[11] Roondiwala, M., Patel, H., & Varma, S. (2015). Predicting stock prices using LSTM. *International journal of science and research*, *6*(4), 2319–7064. https://www.quandl.com/data/NSE

[12] Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2018). NSE stock market prediction using deep-learning models. *Procedia computer science*, *132*, 1351–1362. https://doi.org/10.1016/j.procs.2018.05.050

[13] McNally, S., Roche, J., & Caton, S. (2018). Predicting the price of bitcoin using machine learning. *2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP)* (pp. 339–343). IEEE. https://doi.org/10.1109/PDP2018.2018.00060

[14] Chung, H., & Shin, K. (2020). Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction. *Neural computing and applications*, *32*(12), 7897–7914. https://doi.org/10.1007/s00521-019-04236-3

[15] Xu, Y., Chhim, L., Zheng, B., & Nojima, Y. (2020). Stacked deep learning structure with bidirectional long-short term memory for stock market prediction. *Neural computing for advanced applications: First international conference, ncaa 2020, shenzhen, china, july 3-5, 2020, proceedings 1* (pp. 447–460). Springer. https://doi.org/10.1007/978-981-15-7670-6_37

[16] Lu, W., Li, J., Wang, J., & Qin, L. (2021). A CNN-BiLSTM-AM method for stock price prediction. *Neural computing and applications*, *33*(10), 4741–4753. https://doi.org/10.1007/s00521-020-05532-z

[17] Wang, Y., Wang, L., Yang, F., Di, W., & Chang, Q. (2021). Advantages of direct input-to-output connections in neural networks: The Elman network for stock index forecasting. *Information sciences*, *547*, 1066–1079. https://doi.org/10.1016/j.ins.2020.09.031

[18] Dutta, A., Kumar, S., & Basu, M. (2020). A gated recurrent unit approach to bitcoin price prediction. *Journal of risk and financial management*, *13*(2), 23. https://doi.org/10.3390/jrfm13020023

[19] Qiu, J., Wang, B., & Zhou, C. (2020). Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PLoS one*, *15*(1), e0227222. https://doi.org/10.1371/journal.pone.0227222

[20] Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on signal processing*, *45*(11), 2673–2681. https://doi.org/10.1109/78.650093