-- Iranian Journal of Economic Research---------
Volume 30, Issue 103, Summer 2025, 101-137
ijer.atu.ac.ir
https://doi.org/10.22054/ijer.2025.82330.1315

ATU PRESS

Original Research Received: 7/10/2024 Revised: 9/9/2025 Accepted: 22/9/2025 ISSN: 1726-0728 eISSN:2476-6445

Exchange Rate Forecasting in Iran Using Data Fusion and a Comprehensive Machine Learning Model

Elmira Asle Roosta — Ph.D. Candidate in Economics, Semnan University, Semnan, Iran

Alireza Erfani* — Professor, Economics Department, Semnan University, Semnan, Iran

Abdolmohammad Kashian — Associate Professor, Economics Department, Semnan University, Semnan, Iran

Abstract

The exchange rate is recognized as a key economic indicator influenced by multiple factors. Some of these factors manifest as measurable economic variables, while others are reflected in political and financial news. A central, unresolved question is whether it is possible to develop a comprehensive and scalable model for exchange rate modeling and forecasting that accounts for all relevant variables and factors. Using a data fusion approach, the present study proposed a comprehensive deep learning–based model supporting multiple data types. To train the model, exchange rate–related news was collected from ten major national and international sources covering the period from 2014 to 2023 (1393–1402 in the Iranian calendar). The data was then combined with exchange rate figures and other economic indicators. To identify the best model, eight machine learning models, two statistical models, and one large language model were trained and evaluated under both regression and classification settings. To mitigate bias and random effects, the study applied time series–aware cross-validation along with repeated training and testing using different random initializations. The results demonstrated that the proposed approach, which directly incorporates all influential factors, significantly outperforms existing methods.

* Corresponding Author: aerfani@semnan.ac.ir

How to Cite: Asle Roosta, E., Erfani, A. & Kashian, A. (2025). Exchange Rate Forecasting in Iran Using Data Fusion and a Comprehensive Machine Learning Model. *Iranian Journal of Economic Research*, *30*(103), 101-137.

1. Introduction

Exchange rate fluctuations represent one of the most complex challenges in modern economic analysis, shaped by a dynamic interplay of macroeconomic fundamentals, policy decisions, and informational signals disseminated through the media. Traditional econometric approaches often fail to capture these multidimensional interactions, as they rely primarily on quantitative variables and lagged historical data. As a result, they tend to overlook the qualitative influence of news, market sentiment, and expectations that often precede measurable economic changes. Recent advances in artificial intelligence and machine learning have introduced powerful tools for integrating diverse forms of data—both numerical and textual—into unified predictive systems. The present research tried to propose a comprehensive and extensible model for forecasting exchange rates in Iran, combining structured economic indicators with unstructured news data through a data fusion approach.

2. Materials and Methods

This study employed a quantitative and applied methodology based on supervised machine learning techniques. The dataset spans the period from April 2014 to March 2023 (1393–1402 in the Iranian calendar). Daily free-market exchange rates were obtained from three verified sources: the National Exchange website, the Gold and Currency Information Network, and the Bonbast platform. Additionally, key macroeconomic indicators—including GDP growth, inflation rate, unemployment rate, trade balance, public debt, foreign reserves, and oil prices—were collected from official statistical repositories. Then the study went on to incorporate qualitative dimensions. In this respect, news articles related to exchange rate dynamics were gathered from ten major national and international media outlets, including Donya-e-Eqtesad, San'at-Madan-Tijarat, Asia Daily, ISNA, Khabaronline, Tabnak, BBC Persian, and Voice of America Persian. Each news item was labeled according to the contemporaneous changes in exchange rates. Data preprocessing involved normalization, outlier removal, and interpolation of missing values for numerical data. Textual data underwent cleaning, tokenization, and embedding using the ParsBERT model (Farahani et al., 2021), which was fine-tuned on domain-specific economic texts to improve contextual representation. Following preprocessing, approximately 388,354 fused samples were constructed. Eight machine learning models (Random Forest, XGBoost, LightGBM, CNN-LSTM, GRU, Bi-GRU, LSTM, and Bi-LSTM), two statistical models (ARIMA and Prophet), and one large language model (GPT-4) were trained and compared under both regression and classification

settings. Model evaluation was conducted through time-series–aware cross-validation and repeated random initialization to minimize bias. Performance metrics included Mean Absolute Error (MAE), Mean Squared Error (MSE), Accuracy, and F1-score.

3. Results and Discussion

The results revealed that models integrating textual and numerical data substantially outperform those trained solely on numerical inputs. Specifically, the inclusion of news embeddings reduced forecasting error by more than 5% across most deep learning architectures. Among the evaluated models, the fine-tuned GPT-4 achieved the highest overall accuracy and the lowest error metrics in both regression and classification tasks. However, considering constraints on interpretability and data security, the Bi-directional Gated Recurrent Unit (Bi-GRU) model was identified as the optimal choice for practical implementation. The Bi-GRU model exhibited strong learning capability in capturing temporal dependencies and contextual relationships between macroeconomic variables and market sentiment. In classification mode, it achieved an F1-score of 0.84 and an accuracy rate of 0.86 when textual data were incorporated. In contrast, traditional statistical models such as ARIMA and Prophet showed limited capacity to reflect short-term market shocks influenced by real-time news.

The findings highlighted the importance of data fusion in financial forecasting. Textual news data provide early signals of market sentiment that often precede observable changes in economic variables. By integrating these heterogeneous data sources, the proposed model can offer a more dynamic and responsive forecasting framework, particularly suited to volatile markets such as Iran's foreign exchange sector.

4. Conclusion

This study proposed a comprehensive machine learning–based model that successfully integrates textual and numerical data for exchange rate forecasting in Iran. The results confirmed that data fusion enhances predictive accuracy and robustness, outperforming both conventional econometric methods and single-modality deep learning models. Among the evaluated architectures, Bi-GRU offered the most practical balance between performance, interpretability, and computational efficiency. The findings underscored that incorporating news-driven sentiment and contextual information provides a timely advantage for policy formulation and risk management. Moreover, the modular structure of the proposed model allows for future extensions to other economic domains such as stock market analysis and inflation

forecasting. Future studies are recommended to expand the dataset to include social media sentiment and to adopt explainable AI (XAI) techniques to improve interpretability and transparency.

Keywords: Exchange Rate Forecasting, Data Fusion, Comprehensive Model, Machine Learning, Artificial Intelligence

JEL Classification: C45, C51, C53, C55, G17

مقاله پژوهشی

تاریخ دریافت: ۱۴۰۳/۷/۱۶

تاریخ بازنگری: ۱۴۰۴/۶/۱۸

تاریخ پذیرش: ۱۴۰۴/۶/۳۱

ISSN:۱۷۲۶-۰۷۲۸

eISSN:۲۴۷۶-۶۴۴۵

-- پژوهش‌های اقتصادی ایران ---------

دوره ۳۰، شماره ۱۰۳، تابستان ۱٤۰٤، ۱۰۱-۱۳۷

ijer.atu.ac.ir

https://doi.org/10.22054/ijer.2025.82330.1315

پیش‌بینی نرخ ارز در ایران با استفاده از تلفیق داده‌ها و مدل جامع مبتنی بر یادگیری ماشین

المیرا اصل روستا | دانشجوی دکتری رشته اقتصاد، دانشگاه سمنان، سمنان، ایران.

علیرضا عرفانی * | استاد گروه اقتصاد، دانشگاه سمنان، سمنان، ایران.

عبدالمحمد کاشیان | دانشیار گروه اقتصاد، دانشگاه سمنان، سمنان، ایران.

چکیده

نرخ ارز همواره از مهم‌ترین شاخص‌های اقتصادی بوده که عوامل مختلفی در تعیین آن مؤثرند. بعضی از این عوامل در قالب متغیرهای اقتصادی و برخی دیگر به شکل اخبار سیاسی-اقتصادی بازتاب دارند. پرسش مهمی که تاکنون پاسخ دقیقی به آن داده نشده، این است که آیا می‌توان مدلی جامع و توسعه‌پذیر به منظور مدل‌سازی و پیش‌بینی نرخ ارز داشت به‌نحوی که دربرگیرنده تمامی متغیرها و عوامل مؤثر باشد؟ در این پژوهش به‌عنوان پاسخی برای این پرسش، با استفاده از یادگیری ماشین و رویکرد تلفیق داده‌ها، مدلی جامع مبتنی بر یادگیری عمیق ارائه شده که از انواع داده پشتیبانی می‌کند. به منظور آموزش مدل اخبار مؤثر بر نرخ ارز از ۱۰ پایگاه اصلی داخلی و خارجی در بازه زمانی ۱۳۹۳ تا ۱۴۰۲ جمع‌آوری شده و به همراه داده‌های نرخ ارز و سایر شاخص‌های اقتصادی مستقیماً به مدل داده شده است. به منظور یافتن بهترین مدل، ۸ مدل یادگیری ماشین، ۲ مدل آماری و یک مدل زبانی بزرگ در هر دو حالت رگرسیون و کلاس‌بندی آموزش و آزموده شده‌اند. برای اجتناب از سوگیری و نتایج تصادفی، از تکنیک‌های اعتبارسنجی متقابل منطبق بر توالی زمانی و تکرار آموزش و آزمون مدل‌ها با مقادیر اولیه تصادفی متفاوت، استفاده شده است. نتایج به‌دست آمده حاکی از آن است که رویکرد جامع و توسعه‌پذیر پیشنهادی با لحاظ کردن تمامی عوامل مؤثر به صورت مستقیم، به‌طور قابل توجهی عملکرد بهتری در مقایسه با رویکردهای گذشته داشته است.

کلیدواژه‌ها: پیش‌بینی نرخ ارز، تلفیق داده‌ها، مدل جامع، یادگیری ماشین، هوش مصنوعی.

طبقه‌بندی JEL: C45, C51, C53, C55, G17.

مقاله حاضر برگرفته از رساله دکتری رشته اقتصاد دانشگاه سمنان است.

* نویسنده مسئول: aerfani@semnan.ac.ir

۱. مقدمه

نرخ ارز تحت تأثیر عوامل مختلفی ازجمله شاخص‌های کلان اقتصادی، سیاست‌های مالی و پولی و همچنین اخبار منتشره در رسانه‌ها قرار دارد به‌طوری‌که می‌توان گفت نرخ نهایی از برآیند تمامی این عوامل منتج می‌شود.

عدم پشتیبانی از کلیه عوامل مؤثر بر نرخ ارز، کارایی مدل‌ها در پیش‌بینی رخدادهای آنی همچون آنچه در خبرها منعکس می‌شود را از بین می‌برد.

پیشرفت‌های اخیر در حوزه یادگیری ماشین و یادگیری عمیق این امکان را فراهم کرده است که داده‌های متنی، نظیر اخبار و تحلیل‌های اقتصادی، به همراه داده‌های عددی در مدل‌های پیش‌بینی اقتصادی مورد استفاده قرار گیرند. بر این اساس، با استفاده از رویکرد یادگیری ماشین و به‌کارگیری ایده‌هایی چون بازنمایی انواع داده، می‌توان مدلی جامع ایجاد کرد که امکان پشتیبانی از انواع مختلف داده را داشته و با تلفیق آنها کارایی بالاتری نیز به‌دست دهد. این مسئله ایده اصلی این مقاله است.

این پژوهش شامل چندین مرحله کلیدی است. ابتدا، داده‌های عددی مربوط به نرخ ارز و شاخص‌های کلان اقتصادی و داده‌های متنی شامل اخبار منتشره از منابع معتبر در بازه زمانی مشخص نیز استخراج و سپس، داده‌های متنی با استفاده از تکنیک‌های پردازش زبان طبیعی[1] و مدل‌های جاسازی[2] مانند برت[3] به بردارهای عددی تبدیل شدند. بردارهای عددی حاصل از انواع داده با یکدیگر تلفیق شده و به مدل داده می‌شوند.

در مرحله بعد، از مدل‌های مختلف یادگیری ماشین برای آموزش و پیش‌بینی نرخ ارز استفاده شده است. در آخر و پس از ارزیابی کارایی مدل‌ها، بهترین مدل انتخاب شده است. نتایج نشان می‌دهد که تغذیه مدل با تلفیقی از انواع داده‌ها، به‌طور قابل‌توجهی پیش‌بینی‌های انجام شده را بهبود می‌بخشد.

سازمان‌دهی این مقاله به این صورت است: پس از مقدمه‌ای بر مسئله، پیشینه داخلی و خارجی مرتبط در این زمینه تشریح شده است. در ادامه، معماری و مدل پیشنهادی به همراه توضیحات هر یک از مراحل موجود در خط لوله پردازش داده، بخش اعظم مقاله را تشکیل

1. Natural Language Processing
2. Embedding
3. BERT

می‌دهد. در خلال توضیحات مدل و مراحل پردازش، جزئیات پیکربندی هر مدل نیز ذکر شده است. پس از تشریح مدل و مراحل پردازشی، بخش مرتبط با آزمون‌ها و نتیجه قرار دارد. در این بخش انواع مدل‌ها در دو حالت رگرسیون و کلاس‌بندی ارزیابی شده و نتایج آزمایشات گزارش شده است. بخش انتهایی به بحث و جمع‌بندی اختصاص یافته است.

۲. پیشینه پژوهش

به منظور حفظ اختصار، مهم‌ترین و مرتبط‌ترین پژوهش‌ها در این بخش ذکر شده‌اند. پژوهش‌های مرتبط با این مقاله، شامل تمامی پژوهش‌هایی است که یا مستقیماً درگیر پیش‌بینی نرخ ارز در بازار داخلی یا خارجی بوده‌اند یا روش پیشنهادی حاوی ایده و نکته ارزشمندی از جهت به‌کار بردن انواع داده‌ها در پیش‌بینی است.

۲-۱. پیشینه داخلی

در پژوهش‌های موجود رویکردهای مختلفی در مدل‌سازی نرخ و تحلیل بازار ارز به‌کار بسته شده است. مستقل از نوع نگاه و مدل استفاده شده، تا زمان نگارش این مقاله، در هیچ‌یک از پژوهش‌ها سایر انواع داده مانند داده‌های متنی مستقیماً در مدل‌سازی و پیش‌بینی نرخ ارز به‌کار نرفته‌اند.

جدول ۱. پیشینه پژوهش داخلی

پژوهش(ها)	شرح
امیری و همکاران (۱۳۹۹)	استفاده از مدل‌سازی سیستم‌های پویا[1] و معادلات تفاضلی[2] به منظور بررسی رفتار نرخ ارز و تأثیر متغیرهای کلیدی اقتصادی مانند تورم، نرخ بهره، صادرات و ذخایر ارزی؛ گزارش نتایج براساس معیار آرام‌اس‌ای[3].
گودرزی فراهانی و همکاران (۱۳۹۹)	بررسی اثرات نااطمینانی سیاست‌های اقتصادی بر نوسانات نرخ ارز در ایران با به‌کارگیری مدل ناردی‌ال[4]؛ تصدیق تأثیر نامتقارن شوک‌های سیاستی مثبت و منفی بر نوسانات نرخ ارز.

1. Dynamic Systems
2. Differential Equations
3. RMSE
4. NARDL

ادامه جدول ۱. پیشینه پژوهش داخلی

پژوهش(ها)	شرح
هاشمی دیزج و همکاران (۱۳۹۹)، منصوری گرگری و خداویسی (۱۳۹۸)، شریف‌مقدم و هاشمی (۱۳۹۷)، بیات (۱۳۹۷)، خاشعی و همکاران (۱۳۹۲)، زرانژاد و همکاران (۱۳۸۷)، طیبی و همکاران (۱۳۸۷) و مرزبان و همکاران (۱۳۸٤)	بررسی و به‌کارگیری انواع مختلف شبکه‌های عصبی[1] مانند پرسپترون چندلایه[2]، احتمالاتی[3] به منظور پیش‌بینی نرخ ارز در ایران. حصول دقت بالاتری نسبت به مدل‌های خطی مانند میانگین متحرک یکپارچه خودرگرسیونی[4].
یارمحمدی و محمودوند (۱۳۹٤)	استفاده از تحلیل مجموعه مقادیر به‌عنوان یک روش ناپارامتری در پیش‌بینی نرخ دلار؛ گزارش دقت بالاتر نسبت به مدل‌های میانگین متحرک یکپارچه خودرگرسیونی.
شیرازی و نصرالهی (۱۳۹۲)	استفاده از مدل‌های پولی دورنبوش–فرانکل[5] و بیلسون–فرانکل[6] با در نظر گرفتن انتظارات عقلایی و مقایسه آن‌ها با مدل گام تصادفی؛ گزارش عملکرد بهتر مدل دورنبوش–فرانکل در پیش‌بینی نرخ ارز.
فتاحی و همکاران (۱۳۹۲)	استفاده از الگوریتم ژنتیک برای بهینه‌سازی مدل‌های پیش‌بینی نرخ ارز؛ گزارش دقت بالاتر در مقایسه با میانگین متحرک یکپارچه خودرگرسیونی در پیش‌بینی‌های کوتاه‌مدت.
خداویسی و ملابهرامی (۱۳۹۱)	استفاده از معادلات دیفرانسیل تصادفی مانند مدل حرکت براونی هندسی و مدل پرش مرتن برای پیش‌بینی نرخ ارز؛ عملکرد بهتر مدل اول نسبت به مدل میانگین متحرک یکپارچه خودرگرسیونی.
تقوی و خدام (۱۳۹۰)	استفاده از رویکردهای کلاسیک تطبیقی در پیش‌بینی نرخ ارز شامل نظریه‌های برابری قدرت خرید، تئوری بازار دارایی‌ها، مدل پولی با قیمت‌های انعطاف‌پذیر و مدل ماندل–فلمینگ[7]؛ برتری مدل ماندل–فلمینگ نسبت به سایر مدل‌ها.

1. Neural Networks
2. MLP - Multi Layer Perceptron
3. PNN - Probabilistic Neural Network
4. ARIMA - Autoregressive Integrated Moving Average
5. Dornbusch–Frankel
6. Bilson–Frankel
7. Mundell–Fleming

ادامه جدول ۱. پیشینه پژوهش داخلی

پژوهش(ها)	شرح
ابونوری و همکاران (۱۳۸۸)	استفاده از خانواده مدل‌های اقتصادسنجیِ خودرگرسیونی مشروط ناهمسان واریانس[1] به منظور بررسی اثرات متقارن و نامتقارن اخبار بر نوسانات نرخ ارز؛ گزارش تأثیر بیشتر اخبار منفی بر نوسانات نرخ ارز و تأیید اثر نامتقارن اخبار مثبت و منفی؛ عدم استفاده مستقیم از متن اخبار و استفاده از نتیجه آنها بر بازار به منظور تحلیل.
رحیمی‌بروجردی (۱۳۷۹)	قدیمی‌ترین پژوهش رسمی در دسترس مرتبط با پیش‌بینی نرخ ارز در ایران؛ پیش‌بینی براساس مطالعه شاخص‌های اقتصادی، الگوهای اقتصادسنجی، روش رگرسیون و مدل خودرگرسیون برداری[2].

مأخذ: یافته‌های پژوهش

۲-۲. پیشینه خارجی

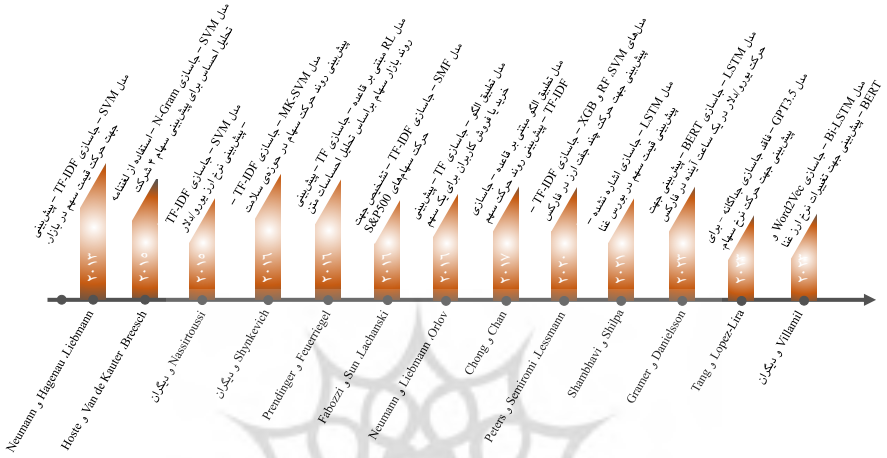
اغلب پژوهش‌های خارجی در سه دسته ارزیابی مدل‌های پولی، ارزیابی مدل‌های اقتصادسنجی و درنهایت ارزیابی مدل‌های مبتنی بر یادگیری ماشین در پیش‌بینی نرخ ارزهای مرجع مانند دلار، یورو، پوند و ین متمرکز بوده‌اند. همچنین بازار تبادل ارزهای خارجی[3] یا همان فارکس[4] مورد توجه بیشتری قرار گرفته است. در انتخاب این مقالات علاوه بر اعتبار علمی و به‌روز بودن، استفاده از سایر انواع داده‌ها (به‌طور خاص داده‌های متنی) در کنار داده‌های عددی نیز در نظر گرفته شده است. فارغ از داده‌های مورد استفاده، این روش‌ها را می‌توان با دو شاخص نحوه تبدیل متون به بردارهای عددی و همچنین مدل یادگیری ماشین استفاده شده طبقه‌بندی کرد.

در شکل ۱، لیستی از پژوهش‌هایی که برای پیش‌بینی مقادیر و متریک‌های مختلف در بازارهای مختلف به نحوی از داده‌های متنی نیز استفاده کرده‌اند، نمایش داده شده است. مشاهدات نشان می‌دهد در سال‌های اخیر روش‌ها در دو جهت پیشرفت داشته‌اند؛ بخش اول مربوط به استفاده از جاسازی‌های پیشرفته‌تر و بخش دوم به‌کارگیری مدل‌های پیچیده‌تر خصوصاً مدل‌های ترکیبی و مبتنی بر یادگیری عمیق است (Villamil, et al., 2023). در

1. ARCH - Autoregressive Conditional Heteroskedasticity
2. VAR - Vector Autoregression
3. Foreign Exchange Market
4. FOREX

بخش مدل، جزئیات بیشتری از این روش‌ها در خلال روش پیشنهادی این مقاله تشریح شده است.

شکل ۱. پژوهش‌های خارجی مرتبط با پیش‌بینی نرخ ارز



مأخذ: یافته‌های پژوهش

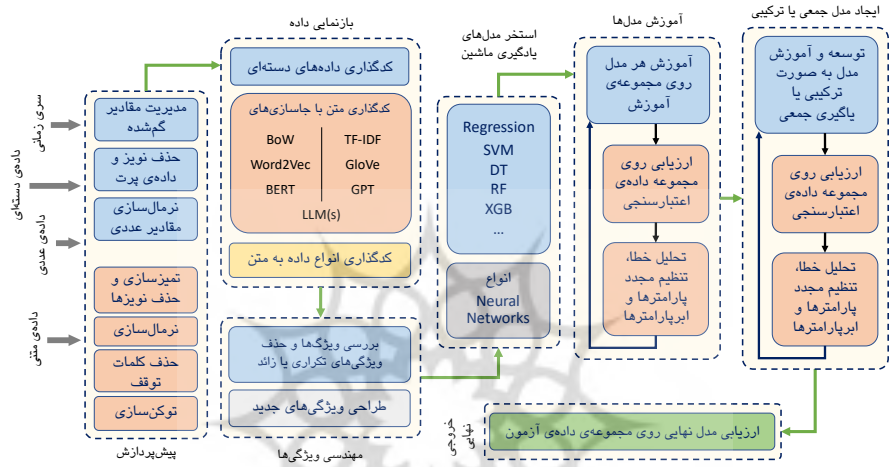
۳. معماری و مدل پیشنهادی

به‌عنوان یک رویکرد استاندارد حل مسئله در یادگیری ماشین، یک خط لوله[1] پردازش داده طراحی کرده‌ایم (شکل ۲). این خط لوله دارای پنج مرحله اصلی و تعدادی زیرمرحله است. این ساختار امکان توسعه و بهبود هر یک از مراحل را میسر ساخته و توسعه‌پذیری و سفارشی‌سازی مدل برای کاربردهای مختلف را تضمین می‌نماید. مواردی که در هر مرحله ذکر شده نمونه‌هایی از انتخاب‌هایی است که می‌توان برای آن مرحله داشت. پیش از ارزیابی نهایی مدل، مرحله‌ای با عنوان یادگیری جمعی وجود دارد. این مرحله الزامی نیست و بسته به کاربرد می‌توان آن را داشت و یا حذف کرد. در این مقاله، به دلیل ماهیت پیچیده خود مدل‌ها و آزمایش‌های انجام شده مشخص شد که ترکیب مدل‌ها منجر به بهبود نتیجه نشده و لذا از آن صرفنظر شده است.

1. Pipeline

به منظور داشتن مدلی جامع که قابلیت تلفیق انواع داده را داشته باشد، اقدامات لازم به منظور پردازش هر نوع داده در مراحل مربوطه انجام شده و درنهایت بردارهای عددی ساخته شده از انواع داده به مدل یادگیری ماشین داده می‌شود. در ادامه تمامی اقدامات انجام شده در این فرآیند تشریح شده است.

شکل ۲. خط لوله پردازش داده و آموزش مدل نهایی



مأخذ: یافته‌های پژوهش

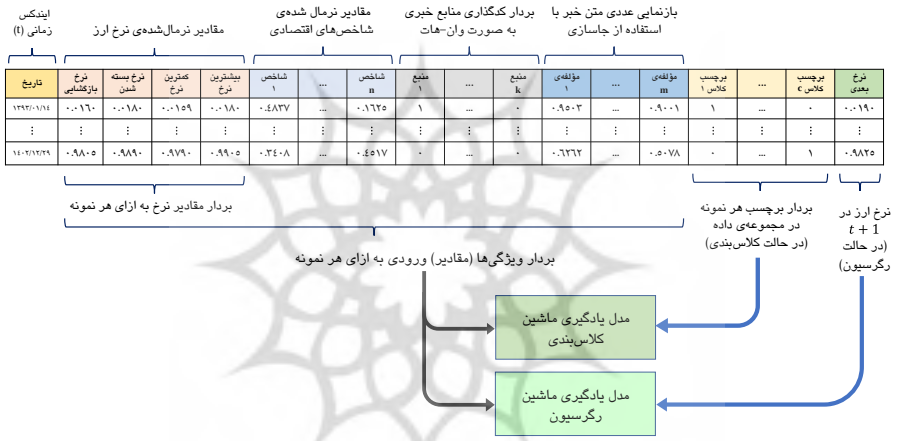
۱-۳. داده‌های مورد استفاده

داده‌های مورد استفاده شامل داده‌های مربوط به نرخ واقعی ارز (ارز معامله شده در بازار آزاد و نه ارز دولتی و سهمیه‌ای) در بازه زمانی فروردین ۱۳۹۳ تا اسفند ۱۴۰۲ است. این داده‌ها از وب‌سایت صرافی ملی، وب‌سایت شبکه اطلاع‌رسانی نرخ طلا و ارز و وب‌سایت بن‌بست به‌عنوان سه مرجع با دیدگاه‌های مختلف جمع‌آوری و مطابقت داده شده است. هر سطر از این داده‌ها، شامل نرخ باز شدن، نرخ بسته شدن، کمترین نرخ معامله و بیشترین نرخ معامله در هر روز است.

علاوه بر داده‌های عددی فوق، تمامی داده‌های متنی مربوط به اخبار منتشره توسط بانک مرکزی، روزنامه‌ها، پایگاه‌ها و خبرگزاری‌های اقتصادی فارسی زبان داخلی شامل روزنامه دنیای اقتصاد، روزنامه صنعت-معدن-تجارت، روزنامه آسیا، روزنامه ابرار اقتصادی، بخش

سیاسی-اقتصادی خبرگزاری ایسنا، بخش سیاسی-اقتصادی پایگاه خبرآنلاین، بخش اقتصادی پایگاه خبری تابناک و پایگاه‌های خارجی شامل بخش سیاسی-اقتصادی بی‌بی‌سی فارسی و همچنین بخش سیاسی-اقتصادی صدای امریکای فارسی در این بازه زمانی جمع‌آوری شده است. ازآنجاکه روش‌های یادگیری ماشین مورد استفاده در اینجا از نوع روش‌های یادگیری بانظارت است، باید به ازای هر خبر یک برچسب[1] نیز به آن الصاق گردد. الصاق برچسب براساس مشاهدات بازار صورت می‌گیرد؛ به این معنا که معاصر با زمان انتشار هر خبر، نرخ ارز چه تغییراتی در بازار داشته است.

شکل ۳. ساخت بردارهای عددی به ازای هر نمونه داده.



مأخذ: یافته‌های پژوهش

ایده مدل جامع پیشنهادی، تلفیق انواع داده است. ساختار مجموعه داده ساخته شده به صورت شکل ۳ است. در مورد بازنمایی‌های عددی، نرمال‌سازی‌ها و برچسب‌ها در ادامه توضیحات لازم داده شده است.

1. Label

۲-۳. پیش‌پردازش داده‌ها

در بخش پیش‌پردازش، اقدامات آماده‌سازی مربوط به هر نوع داده انجام می‌شود. این اقدامات برای سری‌های زمانی و داده‌های عددی مشابه بوده اما برای داده متنی، متفاوت است.

۱-۲-۳. پیش‌پردازش داده‌های عددی و سری‌های زمانی

پردازش داده‌های عددی در سه دسته کلی جای می‌گیرد: ۱- نرمال‌سازی بازه مقادیر ۲- مدیریت مقادیر ناموجود[1] ۳- حذف نویز و مقادیر پرت[2] (Singh & Singh, 2020). در این مقاله از روش استانداردسازی با استفاده از رابطه (۱)، مقادیر هر ویژگی به نحوی نرمال شده‌اند که هر ویژگی میانگین صفر و انحراف معیار استاندارد ۱ دارد.

$$Z = \frac{x-\mu}{\sigma} \quad (۱)$$

در این رابطه x همان مشاهدات یا مقدار ویژگی، μ میانگین و σ انحراف معیار است. با توجه به اینکه بسیاری از الگوریتم‌های یادگیری ماشین به‌صورت توابع احتمال توسعه داده شده و انتظار آنها از ورودی‌ها در بازه ۰ تا ۱ است، این رویکرد بسیار مؤثر است.

در بسیاری از موارد، ممکن است برای یک یا محدوده‌ای از تاریخ، داده‌ها در دسترس نباشند. برای مدیریت این داده‌ها، ایده‌های مختلفی وجود دارد (Josse & Husson, 2012)؛ ایده اول حذف این داده‌ها از مجموعه داده در صورت عدم آسیب به کلیت مسئله است. ایده دوم تقریب این داده‌ها براساس داده‌های مجاور آن با استفاده از روش‌هایی چون میانگین‌گیری و یا درون‌یابی است. در این مقاله در مجموع از ۳۶۵۳ روز، ۱۸۴ داده ناموجود داشتیم که حول آن‌ها نوسانات خاصی در نرخ ارز نبود. لذا این داده‌ها را از مجموعه داده حذف کردیم.

همچنین در بسیاری از مجموعه‌های داده ممکن است به دلایل مختلف اشکال و مقادیر نامربوط وجود داشته باشند. در نگاه اول، وجود این داده‌ها اعتبارسنجی می‌شود. به‌عنوان مثال وقتی ارزش بیشترین معامله در یک روز ۱۰ تومان باشد، اگر قیمت میانگین ۲۰ تومان

1. Missing Values
2. Outliers

درج شده باشد، نویز و داده پرت بوده و باید از مجموعه داده‌ها حذف شود. در مجموع ۵ داده دارای اشکال از مجموعه داده‌ها حذف شد.

علاوه بر نرخ ارز در هر روز، شاخص‌های اقتصادی نیز به‌صورت داده‌های عددی در مجموعه داده وجود دارند. شاخص‌های رشد تولید ناخالص داخلی، نرخ تورم، نرخ بیکاری، تراز تجاری، بدهی عمومی، قیمت نفت، ذخایر ارزی و رشد نقدینگی به‌عنوان شاخص‌های اقتصادی در نظر گرفته شده‌اند. در بخش مهندسی ویژگی‌ها میزان همبستگی این شاخص‌ها با نرخ ارز و با یکدیگر بررسی شده و در نهایت شاخص رشد نقدینگی به دلیل همبستگی بالا با بدهی عمومی از لیست شاخص‌ها حذف شده است.

همچنین به دلیل آنکه نرخ ارز به‌صورت روزانه در دسترس بوده و سایر شاخص‌ها به‌صورت بازه‌های زمانی بزرگ‌تر در دسترس هستند، از یک نگاشت خطی ساده از ابتدا به انتهای هر بازه برای تولید مقدار شاخص به ازای هر روز استفاده شده است. این نگاشت را می‌توان مانند یک خط مستقیم بین مقدار شاخص در ابتدای یک ماه و مقدار آن در انتهای ماه در نظر گرفت. هر نقطه روی این خط، بیانگر مقدار شاخص در آن روز خواهد بود. لازم به ذکر است که این فرض از نظر تئوری یادگیری ماشین، خللی در فرآیند یادگیری و اعتبار مدل ایجاد نمی‌کند.

۲-۲-۳. پیش‌پردازش داده‌های دسته‌ای

منظور از داده‌های دسته‌ای[1] داده‌هایی است که مقدار آن‌ها از یک مجموعه محدود گسسته انتخاب می‌شود. برای این داده‌ها کدگذاری‌های مختلفی پیشنهاد شده است. ایده‌هایی چون استفاده از اعداد ترتیبی نرمال شده حول صفر، کدگذاری وان-هات[2]، کدگذاری دودویی[3] و کدگذاری همینگ[4] هر یک مزایا و معایب خاص خود را در مسائل مختلف دارند (Potdar, et al., 2017). در این مقاله دو نوع داده دسته‌ای وجود دارد. نخست برچسب‌های خروجی مدل‌ها در حالت کلاس‌بندی است که شامل ۷ حالت است. دوم متغیر مشخص‌کننده منبع یک خبر است که ۱۰ مقدار مختلف می‌تواند به خود بگیرد. در اینجا از

1. Categorical
2. One-Hot Coding
3. Binary Coding
4. Hamming Coding

کدگذاری وان-هات برای هر دو مورد استفاده شده که برای این مسئله کاملاً مناسب بوده و کفایت می‌کند. در این کدگذاری، به تعداد حالت‌های مقادیر مختلف یک متغیر دسته‌ای ستون در نظر گرفته می‌شود. به ازای هر نمونه، مقدار آن ویژگی هرچه باشد، ستون متناظر با آن یک شده و مابقی ستون‌ها صفر می‌شوند. در شکل ۳ این کدگذاری قابل مشاهده است.

۳-۲-۳- پیش‌پردازش داده‌های متنی

پیش‌پردازش داده‌های متنی شامل مجموعه اقداماتی است که هدف آن آماده‌سازی داده به منظور ساخت جاسازی[1] برداری بهتر از متن است. عملیات پیش‌پردازش متنی عموماً عبارتند از: تمیزسازی و حذف داده‌های مشکل‌دار، نرمال‌سازی، توکن‌سازی[2]، ریشه‌یابی و حذف کلمات توقف.

منظور از نرمال‌سازی متن، مجموعه عملیاتی است که انواع مختلف حروف موجود در متون مختلف را یکدست و متحدالشکل می‌کند. در اینجا منظور از نرمال‌سازی، یکدست‌سازی قلم[3]ها نیست بلکه هدف یکدست‌سازی حروفی است که معنی یکسان داشته اما با کد (اسکی[4] یا یونیکد[5]) متفاوتی در متن ظاهر شده‌اند. به منظور کاهش فضای حالت، نرمال‌سازی حروف ضروری است. نمونه‌هایی از حروف با معنای مشابه اما با کد متفاوت در شکل ۴ نمایش داده شده‌اند. در نرمال‌سازی هر سه این حرف‌ها باید به یک حرف نگاشت شوند.

شکل ٤. نمونه‌ای از حروف با معنای یکسان و کد متفاوت.

ی - Ë - ي | ۶ - 6 -

مأخذ: یافته‌های پژوهش

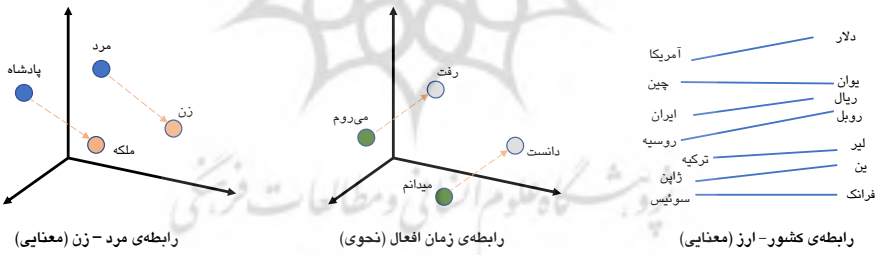
1. Embedding
2. Tokenization
3. Font
4. ASCII
5. Unicode

مرحله بعد، فرآیند توکن‌سازی است که در آن متن ورودی نرمال شده به دنباله‌ای از کلمات شکسته می‌شود. در ساده‌ترین حالت، بعد از تمیزسازی و نرمال‌سازی متن، اگر علائم نگارشی را با فاصله جایگزین کنیم، هر آنچه که با فاصله از یکدیگر جدا شده، یک توکن خواهد بود. البته روش‌های پیشرفته‌تری نیز برای توکن‌سازی وجود دارد (Dridan & Oepen, 2012) که برای کاربرد این مقاله همین ایده کفایت می‌کند. همچنین به دلیل استفاده از جاسازی‌های پیشرفته و مدل‌های زبانی بزرگ در مدل‌های نهایی، نیازی به داشتن مراحل ریشه‌یابی و حذف کلمات توقف نیست.

۳-۳. بازنمایی داده‌های متنی

به منظور بازنمایی داده‌های متنی به بردارهای عددی، از انواع جاسازی‌ها می‌توان استفاده کرد (Li & Yang, 2018). از روش‌های اولیه می‌توان به خورجین کلمات و تی‌اف-آی‌دی‌اف[1] اشاره نمود. در طول زمان به منظور مدل‌سازی و انعکاس روابط معنایی[2] و نحوی[3] زبان‌ها، بازنمایی‌ها و جاسازی‌های متنی پیشرفته‌تری تولید شدند. نمونه‌هایی از روابط معنایی و نحوی توکن‌ها در شکل ۵ نمایش داده شده است.

شکل ۵. نمونه‌ای از روابط معنایی و نحوی در فارسی.



مأخذ: بازطراحی شده توسط نویسندگان مقاله براساس (Mikolov, et al., 2013).

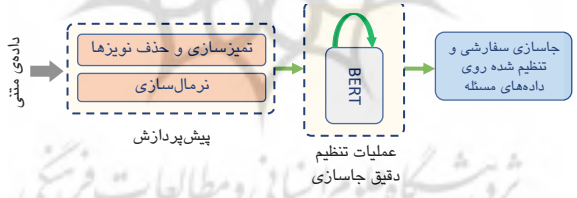
در این مقاله از جاسازی برت[4] فارسی (Farahani, et al., 2021) که توسعه‌ای بر مدل برت اصلی است، برای بازنمایی داده‌های متنی استفاده شده است. جاسازی برت که مخفف

1. TF-IDF
2. Semantic
3. Syntactic
4. BERT

بازنمایی‌های کدگذاری دوطرفه از ترنسفورمرها[1] است، در سال ۲۰۱۸ تحت مالکیت معنوی شرکت گوگل (Devlin, et al., 2018) و با هدف مدل‌سازی کلمات در متن با سرعت بالاتر ارائه گردید.

به دلیل آنکه جاسازی‌های بزرگ روی حجم عظیم داده‌ها آموزش داده شده و برای مدل‌سازی زبان به‌صورت عمومی به‌کار می‌روند، در کاربردهایی که روی بخش خاصی از ادبیات زبانی تمرکز دارند، عموماً کیفیت متوسطی دارند. در این شرایط باید این جاسازی‌ها را با داده‌های مرتبط با پژوهش، تنظیم دقیق[2] یا سفارشی‌سازی کنیم (شکل ۶). این فرآیند در واقع بخشی از فرآیند آموزش و تولید جاسازی است که با تمرکز بر داده‌های خاص و سفارشی شده انجام می‌شود.[3] با این عملیات کیفیت بازنمایی متون مورد استفاده در پژوهش توسط جاسازی، بالاتر خواهد رفت. خروجی جاسازی تنظیم شده به ازای هر سند ورودی، یک بردار با ابعاد ۷۶۸ است. در این مقاله، عملیات تنظیم دقیق به اندازه ۸ دور آموزش مدل[4] انجام شده است. این مقادیر با در نظر گرفتن جلوگیری از بیش‌برازش جاسازی برت لحاظ شده‌اند. همچنین نرخ یادگیری[5] در بازه $[1e-5, 5e-5]$ بوده و اندازه دسته[6] براساس پردازنده گرافیکی[7] در دسترس ۱۶ یا ۳۲ انتخاب شده است.

شکل ٦. تنظیم دقیق جاسازی روی داده‌های متنی حوزه پژوهش



مأخذ: یافته‌های پژوهش

1. Bi-directional Encoder Representation from Transformers
2. Fine-Tuning
۳. جزئیات این فرآیند در Sousa, et al., 2019 موجود است.
4. Epoch
5. Learning Rate
6. Batch Size
7. GPU

در سال‌های اخیر تلاش‌های بسیاری در جهت ارائه مدل‌های زبانی بزرگ انجام شده که شاخص‌ترین آنها مدل‌های زبانی جی‌پی‌تی[1] ارائه شده توسط شرکت اوپن‌ای‌آی[2] هستند (Brown, et al., 2020). تا زمان تهیه این سند ۷ نسخه رسمی از مدل زبانی جی‌پی‌تی ارائه شده که آخرین نسخه آن، نسخه GPT4-O است که در قالب چت‌جی‌پی‌تی[3] به‌شدت دنیا را تحت تأثیر قرار داده است (Mohamadi, et al., 2023). در این پژوهش، از مدل‌های زبانی بزرگ به‌عنوان یکی از مدل‌های ارزیابی استفاده شده است.

٤-٣. مهندسی ویژگی‌ها

مهندسی ویژگی‌ها فرآیندی است که در آن سودمندی یا زائد بودن ویژگی‌های موجود بررسی شده و یا ویژگی‌های جدیدی براساس نیاز مسئله تعریف می‌شود (Nargesian, et al., 2017).

در این پژوهش نیز ویژگی‌های مرتبط با نرخ ارز قبل از استفاده از آنها در آموزش مدل ارزیابی شده تا ویژگی‌های زائد مشخص شوند. یک ویژگی زائد، ویژگی است که در فرآیند کلاس‌بندی یا رگرسیون، نقش خاصی ایفا نمی‌کند. این ویژگی‌ها عموماً یا همبستگی زیادی با مسئله ندارند و یا توسط ویژگی‌های دیگر پوشش داده می‌شوند.

روش‌های مختلفی برای یافتن ویژگی‌های زائد و انتخاب ویژگی‌های مفید وجود دارد (Heaton, 2016). در ساده‌ترین حالت می‌توان همبستگی هر یک از ویژگی‌ها با یکدیگر و همچنین با برچسب خروجی را محاسبه کرده و ویژگی‌هایی که با خروجی همبستگی پائینی داشته یا توسط ویژگی دیگری توصیف می‌شوند، را حذف کرد. اگر برداری از ویژگی‌های به شکل $X = [X_1, \dots, X_n]$ داشته باشیم، همبستگی آن به صورت زیر است:

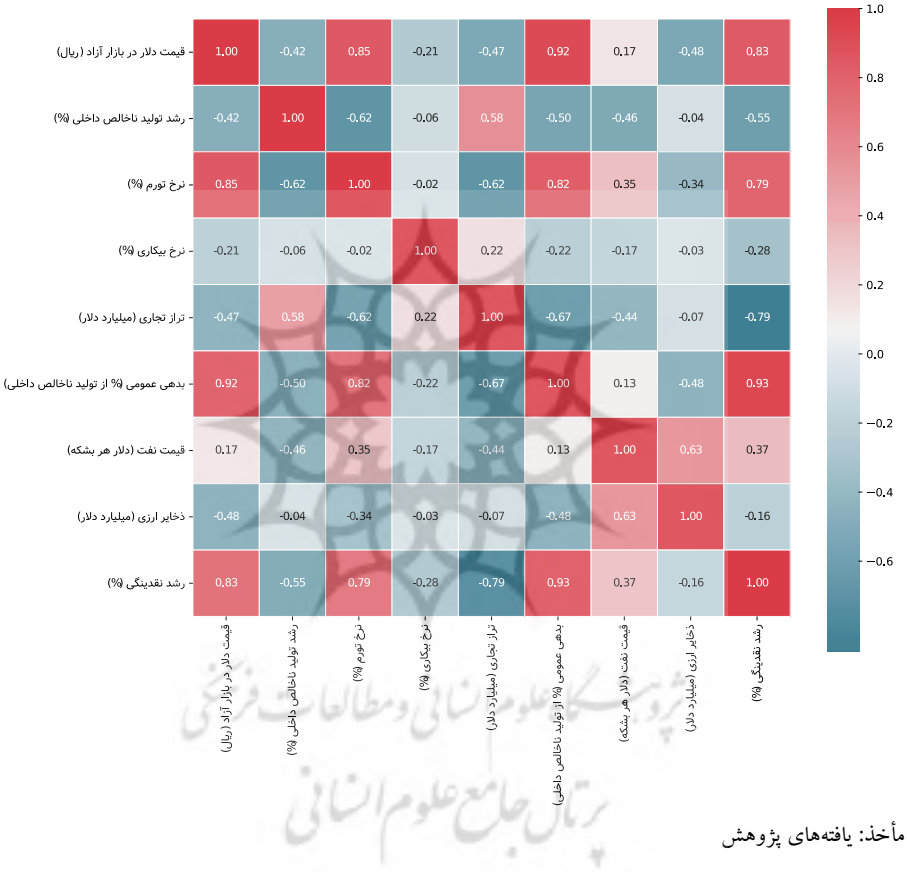
$$Corr(X) = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1n} \\ \vdots & \ddots & \vdots \\ \rho_{n1} & \cdots & \rho_{nn} \end{bmatrix},\ \rho_{nn} = \frac{COV(X_i, X_j)}{X_{\sigma_i} X_{\sigma_j}} \qquad (۲)$$

در ماتریس همبستگی X، اندازه بزرگ‌تر مقادیر ρ_{ij}، نشان‌دهنده هم‌خطی[4] بودن جدی در دو متغیر i و j است (Nargesian, et al., 2017). براساس شکل ۷، به نظر می‌رسد که

1. GPT
2. OpenAI
3. ChatGPT
4. Co-Linearity

نرخ تورم، میزان بدهی عمومی و رشد نقدینگی بیشترین همبستگی را با نرخ ارز بازار آزاد دارند. همچنین سایر شاخص‌های اقتصادی دیگر نیز با نرخ ارز دارای همبستگی هستند به‌طوری‌که هیچ ویژگی کم‌اهمیتی (زائد) در این مجموعه دیده نمی‌شود. این نتیجه مؤید صحت داده‌های جمع‌آوری شده و مطابقت آن با تئوری‌ها و مشاهدات اقتصادی است.

شکل ۳. نمودار همبستگی شاخص‌های اقتصادی ایران (۲۰۰۰ – ۲۰۲۳)



مأخذ: یافته‌های پژوهش

مسئله دیگر، بررسی متغیرهایی است که با یکدیگر (و نه با نرخ ارز) همبستگی بالایی دارند. براساس شکل ۷، متغیرهای «بدهی عمومی» و «رشد نقدینگی» بسیار به یکدیگر مرتبط‌اند. بنابراین، این متغیرها اطلاعات جدیدی به مسئله اضافه نکرده و داشتن یکی از آنها برای آموزش و عملکرد مدل کافی است. لذا از نظر تئوری و به منظور کاهش فضای حالت

مسئله، می‌توان یکی از این دو ویژگی را حذف کرد. بین این دو گزینه، رشد نقدینگی کاندیدای حذف است زیرا بدهی عمومی همبستگی بالاتری با نرخ ارز دارد.

۵-۳. مدل پیشنهادی

برای انتخاب مدل مناسب، مدل‌های مختلفی را پیاده‌سازی و آزموده‌ایم. مدل‌های مختلف برای دو حالت از مسئله آموزش داده شده‌اند؛ در حالت اول مسئله به‌عنوان کلاس‌بندی در نظر گرفته شده و در حالت دوم، مسئله به‌عنوان یک مسئله رگرسیون تحلیل شده است. همچنین برای هریک از حالت‌ها، مدل‌ها با دو نوع داده آموزش داده شده‌اند. نوع اول صرفاً شامل داده‌های عددی و نوع دوم شامل داده‌های عددی به‌علاوه داده‌های متنی بوده است. دلیل این رویکرد آن است که در بخش تحلیل بتوانیم تأثیر حقیقی استفاده از اخبار را بررسی کنیم.

نظر به این‌که مدل‌های کلاسیک یادگیری ماشین مانند ماشین بردار پشتیبان[1]، جنگل تصادفی[2] و امثالهم در مسائل پیچیده و بزرگ کارایی مناسبی ندارند (Khan, et al., 2023) لذا از مدل‌های یادگیری عمیق استفاده شده است. کارایی هر یک از این مدل‌ها به داده ورودی، معماری و ظرفیت مدل و همچنین کیفیت آموزش مدل بستگی دارد. از میان مدل‌های یادگیری ماشین که از ابتدا برای کار با داده‌های زمانی و یا سری‌ها و دنباله‌ها طراحی شده‌اند، روابط موجود در توالی داده‌ها در سری‌های زمانی را نیز در نظر می‌گیرند. ازجمله معروف‌ترین این مدل‌ها شبکه‌های عصبی بازگشتی[3]، حافظه کوتاه‌مدت بلند[4] و واحدهای بازگشتی دروازه‌دار[5] هستند. تمامی این مدل‌ها در زمره شبکه‌های عصبی جای می‌گیرند.

به جز مدل‌های ذکر شده در فوق، مدل‌های جدیدتر و با ظرفیت بیشتری نیز وجود دارند که ازجمله مهم‌ترین آنها می‌توان به مدل‌های مبتنی بر معماری ترنسفورمر که ابزارهای جذابی چون چت‌جی‌پی‌تی بر مبنای آنها بنا نهاده شده‌اند، اشاره کرد. از آخرین نسخه مدل

1. Support Vector Machine (SVM)
2. Random Forest (RF)
3. Recurrent Neural Network (RNN)
4. Long Short-Term Memory (LSTM)
5. Gated Recurrent Unit (GRU)

جی‌پی‌تی (در زمان نگارش این مقاله) به‌صورت دسترسی با ای‌پی‌آی[1] و با رویکرد تنظیم دقیق استفاده شده که نتایج و نکات آن در ادامه ذکر شده است.

در ادامه، بروزترین مدل‌های متناسب با مسئله پیاده‌سازی، آموزش و تنظیم دقیق شده‌اند. در میان مدل‌های موجود مدل زبانی بزرگ تنظیم دقیق شده و مدل واحدهای بازگشتی دروازه‌دار دوطرفه بهترین نتیجه را در سناریوهای مختلف داشته‌اند. ازآنجاکه مدل زبانی بزرگ صرفاً از طریق ای‌پی‌آی در دسترس بوده و مانند یک جعبه سیاه به آن دسترسی داریم، در ادامه مدل واحدهای بازگشتی دروازه‌دار دوطرفه به‌عنوان بهترین مدل تشریح شده است.

۱-۵-۳. مدل واحدهای بازگشتی دروازه‌دار دوطرفه

این مدل در زمره شبکه‌های عصبی تکرارپذیر[2] قرار می‌گیرد. مزیت این مدل نسبت به مدل حافظه کوتاه‌مدت بلند دوطرفه[3] که در اکثر پژوهش‌های دیگر استفاده شده، سرعت آموزش بالاتر آن است (Seabe, et al., 2023). همچنین مدل پیشنهادی به دلیل ترکیب و استفاده از انواع داده در فرآیند آموزش و پیش‌بینی، ذیل تکنیک‌های تلفیق داده‌ها[4] قابل طبقه‌بندی است (Liu, et al., 2020).

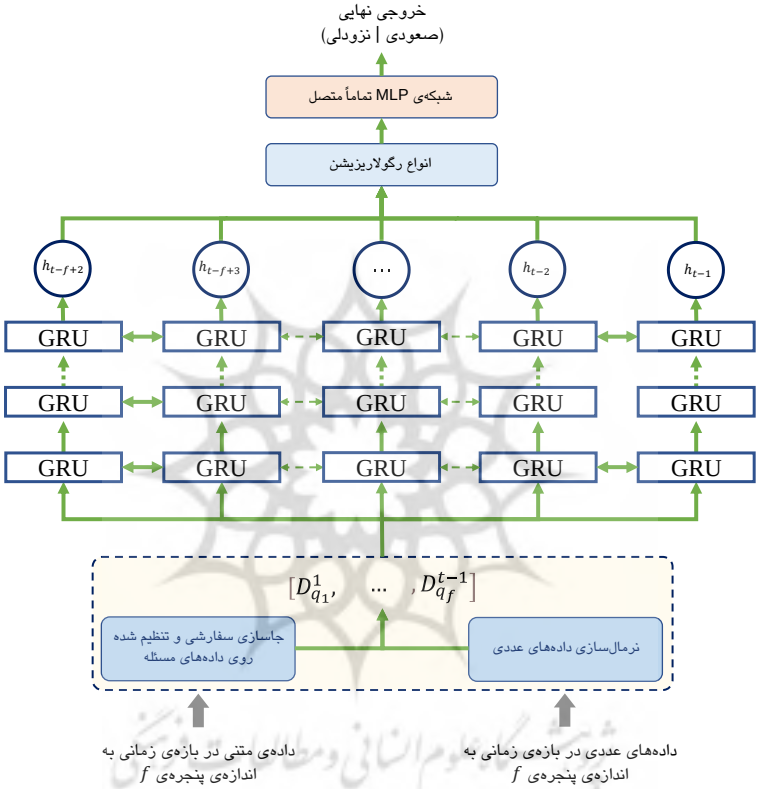
ساختار کلی مدل استفاده شده در شکل ۸ نمایش داده شده است. این مدل دارای تعدادی مرحله یا واحد است که در هر مرحله داده‌های مربوط به f روز گذشته را دریافت کرده و براساس آن وزن‌ها یا پارامترهای مدل را به‌روز می‌کند. لذا برای هر پیش‌بینی در زمان t، مقادیر ورودی در بازه $[t-f+1, \dots, t-1]$ را به مدل می‌دهیم. در هر دوره آموزش، وزن‌ها براساس میزان خطای مدل، به‌روز می‌شوند. به‌روزرسانی وزن‌ها تا جایی ادامه می‌یابد که مقدار خطای مدل به ازای چند دوره آموزش متوالی مدل کاهش نیافته و یا به آستانه مورد قبول تعریف شده برسد. تابع خطا در رابطه ۴ نمایش داده شده است. این تابع به تابع آنتروپی متقابل[5] معروف بوده و به دلیل ویژگی‌های محاسباتی مناسب در فرآیند آموزش، در مسائل کلاس‌بندی بسیار مرسوم است. در این رابطه $\hat{y}^t$ پیش‌بینی مدل و y^t برچسب

1. Access Point Interface (API)
2. Recurrent Neural Networks
3. Bi-LSTM
4. Data Fusion
5. Cross-Entropy

حقیقی داده در زمان (روز) t است. این تابع توسط بهینه‌ساز آدام[1] در فرآیند آموزش بهینه می‌شود.

$$Loss(t) = \frac{1}{T}\sum_{t=1}^{T}(y^t \log(\hat{y}^t) + (1-y^t)\log(1-\hat{y}^t)) \quad (۴)$$

شکل ۸. ساختار مدل واحدهای بازگشتی دروازه‌دار دوطرفه



مأخذ: یافته‌های پژوهش

به منظور جلوگیری از بیش‌برازش[2] مدل از تکنیک‌های تعدیل[3] استفاده می‌شود. در اینجا از دو تکنیک متداول تعدیل یعنی دوراندازی[4] و نرمال‌سازی دسته‌ای[5] است. به منظور تجمیع

1. ADAM
2. Overfitting
3. Regularization
4. Dropout
5. Batch Normalization

خروجی مراحل مختلف، در لایه آخر شبکه از یک لایه شبکه عصبی پرسپترون چندلایه[1] به‌صورت تماماً متصل[2] استفاده شده است. در حالت کلاس‌بندی، تابع فعال‌سازی این شبکه تابع سافت‌مکس[3] بوده و خروجی نهایی به‌صورت دودویی[4] مشخص می‌شود. در حالت رگرسیون تابع فعال‌سازی لایه آخر، یک تابع خطی همانی ساده است.

با قراردادن بردار حاصل از جاسازی متنی، به ازای هر سند یا خبر در روز t برداری به‌صورت $[e_1^t, \ \ldots \ , e_K^t \,|v_1^t, \ \ldots \ , e_R^t]$ خواهیم داشت که در آن e_i^t مقدار هر درایه از بردار جاسازی به طول K و v_i بیانگر مقدار نرمال شده یک متغیر عددی از مجموعه R متغیر عددی است. این بردار را به‌صورت D_p^t نمایش داده و آن را به‌صورت بردار داده برای خبر p در روز t تفسیر می‌کنیم. با این کدگذاری از مسئله، برای هر روز، به تعداد اخبار موجود از منابع مختلف، نمونه خواهیم داشت. برای ساخت داده‌های مربوط به روز $t-i$، متغیرهای عددی را عیناً استفاده کرده و برای بخش بردار جاسازی، میانگین جاسازی‌های تمامی اخبار منتشره در $t-i$ را قرار می‌دهیم.

۲-۵-۳. پیکربندی مدل واحدهای بازگشتی دروازه‌دار دوطرفه

منظور از پیکربندی[5]، جزئیات مربوط به معماری و ابرپارامترهای مدل است. مدل مورد استفاده از خانواده‌ی مدل‌های واحدهای بازگشتی دروازه‌دار دوطرفه است که برای روابط داده‌های متوالی مناسب است. در حالت اول، اندازه بردار ورودی صرفاً شامل تمامی داده‌های عددی موجود از متغیرهای انتخاب شده است. تعداد این متغیرها ۲۱ عدد می‌باشد که ۴ مورد از آنها مربوط به داده‌های روزانه نرخ ارز، ۷ مورد مربوط به داده‌های منتج شده و نگاشت شده از شاخص‌های اقتصادی، و ده مورد بیانگر مرجع خبر است.

در حالت دوم، به ازای هر داده متنی یک بازنمایی برداری با ابعاد ۷۶۸ خواهیم داشت. این بردار در کنار ۲۱ متغیر دیگر، بردار ورودی به طول ۷۸۹ مؤلفه را ایجاد می‌کند.

1. Multi Layer Perceptron (MLP)
2. Fully Connected
3. SoftMax
4. Binary
5. Configuration

در معماری مدل ۳ لایه از مدل واحدهای بازگشتی دروازه‌دار دوطرفه به‌صورت چیده شده روی هم در نظر گرفته شده است. همچنین در هر لایه، ۳۲ واحد پردازشی از این مدل در نظر گرفته شده است.

به منظور جلوگیری از بیش‌براش مدل، از تکنیک دوراندازی استفاده شده و پارامتر دوراندازی به اندازه ۰/۲ تا ۰/۵ وزن‌ها در نظر گرفته شده است.

به دنبال پشته‌ای از لایه‌های محاسباتی بازگشتی دروازه‌دار دوطرفه، ۲ لایه تماماً متصل قرار داده شده که در هر لایه ۶۴ الی ۲۵۶ نورون در نظر گرفته شده است. همچنین تابع فعال‌سازی برای لایه‌های پنهان تابع رلو[1]، برای لایه خروجی در حالت رگرسیون تابع خطی و برای لایه خروجی در حالت کلاس‌بندی تابع سیگموید[2] لحاظ شده است.

به منظور آموزش مدل، اندازه دسته[3] متناسب با پردازنده گرافیکی و محدودیت‌های سرعت و حافظه، ۳۲ یا ۶۴ انتخاب شده و نرخ یادگیری در بازه $[1e-3,1e-4]$ تنظیم شده است. برای آموزش کامل مدل حداقل ۵۰ و حداکثر ۱۰۰ دوره آموزش لازم است. عدد دقیق براساس وضعیت همگرایی مدل مشخص می‌شود. برای بهینه‌سازی تابع هزینه در مدل، از بهینه‌ساز آدام با پارامترهای $\beta_1 = ۰/۹$ و $\beta_2 = ۰/۹۹۹$ استفاده شده است.

تابع خطا برای حالت رگرسیون تابع میانگین مربعات خطا و برای حالت کلاس‌بندی، تابع آنتروپی متقابل محاسبه شده است. به منظور جلوگیری از بیش‌برازش مدل، علاوه بر تکنیک دوراندازی ذکر شده، از تعدیل ال۲[4] (کاهش وزن‌ها) در لایه‌های تماماً متصل خروجی استفاده شده است.

۳-۵-۳. کدگذاری مسئله کلاس‌بندی

منظور از کدگذاری در اینجا تبدیل مقادیر پیوسته نرخ ارز در خروجی، به مقادیر گسسته است. هفت کلاس در جهت مثبت و منفی در نظر گرفته‌ایم:

- یک کلاس، معادل با بدون تغییر، برای مقدار تغییرات کمتر از ۰/۵٪
- دو کلاس برای تغییرات بین ۰/۵٪ تا ۱٪

1. Rectifier Linear Unite (ReLU)
2. Sigmoid
3. Batch Size
4. L2 Regularization

- دو کلاس برای تغییرات بین ۱٪ تا ۳٪
- دو کلاس برای تغییرات بیش از ۳٪

لذا به ازای هر نمونه ورودی، یکی از ۷ کلاس فوق به نمونه نسبت داده می‌شود. برای نمایش برچسب هر نمونه از کدگذاری وان-هات استفاده شده است (شکل ۳). لذا هفت ستون به ازای هر نمونه داریم که مقدار ستون متناظر با کلاسی که نمونه به آن تعلق دارد برابر با ۱ بوده و مابقی مقادیر صفر هستند.

٤-٥-٣. مدل زبانی بزرگ

در حالتی دیگر از یک مدل زبانی بزرگ که با استفاده از داده آموزش تنظیم دقیق شده است، استفاده کرده‌ایم. مدل مورد استفاده در اینجا جی‌پی‌تی-چهاراو[1] از شرکت اپن‌ای‌آی است. به منظور انجام تنظیم دقیق مدل روی داده‌های آموزشی، داده شکل‌داده شده[2] را از طریق واسط کاربری موجود به سرورهای شرکت اپن‌ای‌ای ارسال کرده و درخواست تنظیم دقیق مدل را ارسال کرده‌ایم. در این روش، همانطور که پیش‌تر نیز اشاره شد، دسترسی مستقیم به مدل نداریم و صرفاً از طریق واسط‌های طراحی شده می‌توانیم با آن تعامل کنیم. پس از انجام عملیات تنظیم دقیق، نیاز به یک فرمان داریم تا براساس داده ورودی، مدل خروجی مدنظر را به ما ارائه دهد.

٤. آزمون‌ها و نتایج

به منظور ارزیابی درست و استاندارد، از مجموعه داده آزمون[3] استفاده شده است. این داده‌ها پیش‌تر توسط هیچ یک از مدل‌ها عیناً دیده نشده‌اند. البته در فرآیند آموزش، داده‌های مشابه با آنها در آموزش مدل‌ها استفاده شده است. در این پژوهش مجموعه داده آزمون به اندازه ۱۰٪ از کل داده‌ها و به‌صورت یک توالی زمانی از ۲۵٪ انتهایی کل مجموعه داده انتخاب شده است. این داده‌ها به‌صورت تصادفی و زمان-آگاه[4] انتخاب شده‌اند.

1. GPT-4o
2. Formatted
3. Test Set
4. Time-Aware

متأسفانه مقایسه مدل‌های پیشنهادی با سایر روش‌های پیشنهاد شده در مقالات، به دلایل در دسترس نبودن مجموعه داده‌ها، مکانیزم‌های نرمال‌سازی و پالایش، الگوریتم‌های پیاده‌سازی، پارامترها و ابرپارامترهای مربوط به هر مدل، میسر نیست.

لذا در این پژوهش تمامی مدل‌های مدرن و مرتبط پیاده‌سازی شده و نتایج آنها گزارش شده است.

به منظور جلوگیری از سوگیری مدل روی بخش خاصی از داده، از روش اعتبارسنجی متقابل[1] استفاده کرده‌ایم. نتیجه گزارش‌شده نهایی در این روش، میانگین تمام نتایج در همه حالت‌هاست.

۱-٤. معیارهای ارزیابی

در صورتی که پیش‌بینی نرخ ارز به‌عنوان یک مسئله رگرسیون مطرح باشد، از میانگین قدرمطلق خطاها[2] (رابطه ۵) و همچنین میانگین مربعات خطا[3] (رابطه ۶) استفاده شده است.

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \quad (۵)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (۶)$$

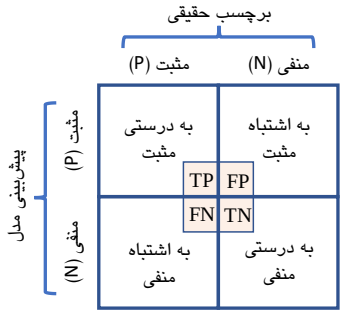
در صورتی که مسئله را به‌عنوان یک مسئله کلاس‌بندی تفسیر کنیم، عموماً از معیار ارزیابی صحت[4] برای ارزیابی کارایی مدل‌های پیش‌بینی نرخ ارز استفاده شده است (رابطه ۷).

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (۷)$$

تفسیر پارامترهای فوق براساس ماتریس درهمی[5] است که در شکل ۹ نمایش داده شده که البته قابل تعمیم به مسائل چندکلاسه است.

1. Cross-Validation
2. Mean Absolute Error (MAE)
3. Mean Squared Error (MSE)
4. Accuracy
5. Confusion Matrix

شکل ۹. ماتریس درهمی



مأخذ: یافته‌های پژوهش

علاوه‌براین معیار، معیار استاندارد دیگر مورد استفاده، امتیاز F1[1] است که دو معیار دقت[2] و یادآوری[3] را با یکدیگر ترکیب می‌کند. معیارهای دقت و یادآوری براساس جدول درهمی به صورت زیر تعریف می‌شوند:

$$Precision = \frac{TP}{TP+FP} \quad (۸)$$

$$Recall = \frac{TP}{TP+FN} \quad (۹)$$

معیار دقت بیان می‌کند که از بین تمامی نمونه‌هایی که مدل به‌عنوان خروجی مثبت پیش‌بینی کرده است، چه تعدادی از آنها واقعاً مثبت بوده‌اند؟ معیار یادآوری بیان می‌کند که از میان تمامی نمونه‌های مثبت، چه تعدادی از آنها را مدل به‌درستی شناسایی کرده است؟

دو معیار دقت و یادآوری با یکدیگر رابطه پایاپایی[4] دارند و بهبود یکی به کاهش دیگری می‌انجامد. معیار امتیاز F1 به منظور ایجاد ترازی[5] بین این دو معیار معرفی شده تا بتوان با استفاده از یک معیار، وضعیت یک مدل را مشخص نمود. این معیار درواقع نوعی میانگین هندسی است و به‌صورت زیر محاسبه می‌شود:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (۱۰)$$

باید توجه نمود که برای گزارش عملکرد یک مدل، نیاز به هر دو معیار صحت و امتیاز F1 داریم. این مسئله به توزیع کلاس‌ها در مجموعه‌های داده بازمی‌گردد. در مجموعه‌هایی

1. F1 Score
2. Precision
3. Recall
4. Trade-Off
5. Balance

که در آنها ناترازی[1] در کلاس‌ها وجود دارد، تنها استفاده از معیار صحت گمراه‌کننده است. مثلاً برای تشخیص بیماران، عموماً اکثرا افراد سالم هستند و تعداد کمی بیمار هستند. اگر مدل ما همه نمونه‌ها را سالم پیش‌پینی کند، عملاً نرخ صحت بالایی (مثلاً ۹۷٪) خواهد داشت! لذا لازم است که نرخ F1 که ناتراز بودن توزیع کلاس‌ها را نیز در نظر می‌گیرد، گزارش کنیم.

۲-٤. ارزیابی مدل‌ها به‌عنوان رگرسور

به منظور بررسی کارایی مدل، ابتدا تمامی مدل‌های مختلف را پیاده‌سازی و بررسی کرده‌ایم. نتایج این بررسی در جدول ۲، نمایش داده شده است.

جدول ۲. نتایج ارزیابی مدل‌ها به‌عنوان رگرسور در پیش‌بینی نرخ ارز

	بدون استفاده از داده متنی		با استفاده از داده متنی	
	میانگین مربعات خطا	میانگین قدرمطلق خطاها	میانگین مربعات خطا	میانگین قدرمطلق خطاها
واحد بازگشتی دروازه‌دار	۰/۰۳۵	۰/۱۲۵	۰/۰۱۵	۰/۰۸۵
حافظه کوتاه‌مدت بلند	۰/۰۳۳	۰/۱۱۵	۰/۰۱۵	۰/۰۸۰
واحد بازگشتی دروازه‌دار دوطرفه	۰/۰۳۲	۰/۱۱۲	۰/۰۱٤	۰/۰۷۸
حافظه کوتاه‌مدت بلند دوطرفه	۰/۰۳۳	۰/۱۱٤	۰/۰۱۵	۰/۰۸۰
شبکه عصبی پیچشی- حافظه کوتاه‌مدت بلند[2]	۰/۰۳۷	۰/۱۲۰	۰/۰۱۷	۰/۰۸۸
شبکه عصبی پیچشی زمانی[3]	۰/۰٤۰	۰/۱۲۵	۰/۰۱٦	۰/۰۹۰
میانگین متحرک یکپارچه خودرگرسیونی	۰/۰۵۰	۰/۱۳۵	---	---
پرافت[4]	۰/۰٤۵	۰/۱۳۵	---	---

1. Imbalance
2. Convolutional Neural Network – Long Short-Term Memory
3. Temporal Convolutional Network
4. Prophet

ادامه جدول ۲. نتایج ارزیابی مدل‌ها به‌عنوان رگرسور در پیش‌بینی نرخ ارز

	با استفاده از داده متنی		بدون استفاده از داده متنی	
	میانگین قدرمطلق خطاها	میانگین مربعات خطا	میانگین قدرمطلق خطاها	میانگین مربعات خطا
تقویت گرادیان شدید[1]	۰/۰۹۰	۰/۰۱۷	۰/۱۲۵	۰/۰۴۰
ماشین تقویت گرادیان سبک[2]	۰/۰۹۲	۰/۰۱۸	۰/۱۳۵	۰/۰۴۳
مدل زبانی تنظیم دقیق شده	۰/۰۷۰	۰/۰۱۲	۰/۱۰۴	۰/۰۲۷

مأخذ: یافته‌های پژوهش

با توجه به جدول ۲، مدل مبتنی بر مدل زبان بزرگ از مدل‌های دیگر عملکرد بهتری داشته است. در مقاوم دوم، مدل واحد بازگشتی دروازه‌دار دوطرفه قرار دارد. همچنین استفاده از داده متنی در تمامی سناریوها منجر به بهبود کارایی مدل و کاهش خطا شده است. نکته جالب توجه در این جدول، فاصله نسبتاً کم روش‌های حافظه کوتاه‌مدت بلند و واحد بازگشتی دروازه‌دار نسبت به یکدیگر است. باید دقت کرد که با وجود این فاصله کم، روش واحد بازگشتی دروازه‌دار در آموزش و استنتاج سریع‌تر و کاراتر از روش‌های حافظه کوتاه‌مدت بلند است.

۳-٤. ارزیابی مدل‌ها به‌عنوان کلاس‌بند

نتایج عملکرد مدل به‌عنوان کلاس‌بند در جدول ۳ نمایش داده شده است. در اینجا نیز مجدداً روش مبتنی بر مدل زبانی بزرگ، بر سایر روش‌ها برتری دارد. در رتبه دوم، واحد بازگشتی دروازه‌دار دوطرفه قرار دارد. نکته حائز اهمیت در اینجا در رابطه با روش‌های آماری میانگین متحرک یکپارچه خودرگرسیونی و پرافت است. این روش‌ها ذاتاً برای کلاس‌بندی طراحی نشده‌اند. برای داشتن مقایسه‌ای از کارایی این روش‌ها، نتایج در آن‌ها را مطابق با کدگذاری برچسب که پیش‌تر توضیح داده شد، تبدیل به خروجی کلاس‌بند شده‌اند.

1. Extreme Gradient Boosting
2. Light Gradient Boosting Machine

جدول ۳. نتایج ارزیابی مدل‌ها به‌عنوان کلاس‌بند در پیش‌بینی نرخ ارز

	بدون استفاده از داده متنی		با استفاده از داده متنی	
	صحت	امتیاز F1	صحت	امتیاز F1
واحد بازگشتی دروازه‌دار	۰/۷۵	۰/۷۳	۰/۸۳	۰/۸۱
حافظه کوتاه‌مدت بلند	۰/۷۶	۰/۷۵	۰/۸٤	۰/۸۲
واحد بازگشتی دروازه‌دار دوطرفه	۰/۷۸	۰/۷۶	۰/۸٦	۰/۸٤
حافظه کوتاه‌مدت بلند دوطرفه	۰/۷۷	۰/۷۵	۰/۸٤	۰/۸۳
شبکه عصبی پیچشی- حافظه کوتاه‌مدت بلند	۰/۷۳	۰/۷۱	۰/۸۱	۰/۷۹
شبکه عصبی پیچشی زمانی	۰/۷۳	۰/۷۰	۰/۸۰	۰/۷۷
میانگین متحرک یکپارچه خودرگرسیونی	۰/٦۳	۰/٦۱	---	---
پرافت	۰/٦٦	۰/٦۳	---	---
تقویت گرادیان شدید	۰/۷۳	۰/۷۱	۰/۷۸	۰/۷۵
ماشین تقویت گرادیان سبک	۰/۷۱	۰/٦۸	۰/۷٦	۰/۷۳
مدل زبانی تنظیم دقیق شده	۰/۸۲	۰/۸۰	۰/۸۸	۰/۸۷

مأخذ: یافته‌های پژوهش

ملاحظه می‌شود که در نظر گرفتن اخبار، تأثیر قابل ملاحظه‌ای در بهبود عملکرد پیش‌بینی نرخ ارز دارد. نکته جالب دیگر در این نتایج، خروجی مدل زبانی بزرگ است. این مدل بهترین نتیجه را در بین مدل‌های موجود دارد اما مشکلاتی نیز دارد. در ساده‌ترین حالت عملکرد این مدل مشابه یک جعبه سیاه است که ما از داخل آن و نحوه عملکرد آن، اطلاع دقیقی نداریم. این مسئله تفسیرپذیری خروجی مدل را کم می‌کند. به عبارت دیگر، وقتی پیش‌بینی‌ها با کاهش دقت مواجه شوند، نمی‌توان به‌طور دقیق مدل را عیب‌یابی کرده و دلایل افت عملکرد را متوجه شد. این مسئله در مدل‌های دیگر و به‌طور کلی در تمامی مدل‌هایی که پیاده‌سازی و کد منبع آنها در دسترس است، وجود ندارد. مسئله دیگر، حفاظت و صیانت از داده‌هاست. در بسیاری از کاربردها، داده‌های مورد استفاده دارای ارزش سازمانی و ارگانی بوده و ارسال آنها به یک مرجع خارجی، خلاف سیاست‌های آن سازمان است. لذا با وجود

عملکرد بهتر مدل زبانی بزرگ، در استفاده از آن باید ملاحظات سازمانی و اهداف بلندمدت را لحاظ کرد.

۵. بحث و نتیجه‌گیری

در این پژوهش تلاش شده تا با بازنمایی مناسب و تلفیق داده، انواع داده‌های عددی، متنی و دسته‌ای در کنار یکدیگر و در یک مدل جامع به منظور پیش‌بینی نرخ ارز به‌کار گرفته شوند. رویکرد پیشنهادی شامل پیش‌پردازش داده‌ها، بازنمایی عددی انواع داده‌ها، مهندسی ویژگی و تلفیق داده و آموزش مدل یادگیری ماشین مبتنی بر یادگیری عمیق است. در این راستا، طیف گسترده‌ای از مدل‌های یادگیری ماشین ازجمله مدل‌های شبکه‌های عصبی تکرارپذیر، مدل واحد بازگشتی حافظه‌دار دوطرفه، حافظه کوتاه‌مدت بلند و همچنین مدل‌های زبانی بزرگ مانند جی‌پی‌تی-چهاراو پیاده‌سازی، آموزش و آزموده شده‌اند تا بهترین مدل برای پیش‌بینی دقیق‌تر نرخ ارز انتخاب شود.

نتایج نشان داد که مدل واحد بازگشتی حافظه‌دار دوطرفه به دلیل توانایی بالاتر در درک روابط پیچیده میان داده‌های متنی و سری‌های زمانی، عملکرد بهتری نسبت به مدل‌های سنتی مانند میانگین متحرک یکپارچه خودرگرسیونی و برخی مدل‌های یادگیری عمیق دیگر داشت. علاوه‌براین، استفاده از مدل‌های زبانی بزرگ مانند جی‌پی‌تی-چهاراو نیز نتایج مطلوبی در پیش‌بینی ارائه داد؛ هرچند محدودیت‌هایی در دسترسی به این مدل‌ها وجود داشت. علاوه بر مشکلات دسترسی، مسئله مالکیت داده‌های سازمانی نیز دغدغه مهمی بود که این مدل را به انتخاب اول این پژوهش تبدیل نکرد. ذکر این نکته نیز حائز اهمیت است که به دلیل عدم اطلاع از ساختار داخلی مدل‌های زبانی بزرگ که به‌صورت سرویس در دسترس هستند، تفسیر نتایج و عیب‌یابی آن‌ها دشوار است. این مسئله نیز محدودیت دیگری است که احتمالاً در آینده مرتفع گردد.

ارزیابی‌ها در هر دو حالت رگرسیون و کلاس‌بندی انجام شده و کارایی مدل‌ها براساس معیارهای خطای‌میانگین‌مطلق و خطای‌میانگین‌مربعات برای رگرسیون و نرخ صحت و امتیاز F1 برای کلاس‌بندی گزارش شده است.

این پژوهش سعی کرده است با ارائه یک مدل جامع که از داده‌های متنی و اخبار استفاده می‌کند، به دیگر حوزه‌های مدلسازی اقتصادی نیز تعمیم‌پذیر باشد. این مدل‌ها به تحلیل و

پیش‌بینی‌های دقیق‌تری کمک می‌کنند زیرا متون و اخبار به‌عنوان یک منبع اطلاعاتی لحظه‌ای و تأثیرگذار، معمولاً از شاخص‌های اقتصادی سنتی مانند نرخ تورم و رشد نقدینگی سریع‌تر تأثیر خود را بر بازار می‌گذارند. این پژوهش نشان داده که با بهره‌گیری از داده‌های متنی در کنار داده‌های عددی، می‌توان به ابزارهای پیش‌بینی بهتری دست یافت که برای کاربردهای مالی و اقتصادی بسیار مفید خواهد بود. به عبارت دیگر، به دلیل نوع‌آوری ایجاد شده در پشتیبانی از انواع داده‌ها، می‌توان اطمینان حاصل نمود که تأثیر انواع داده در بازه‌های زمانی مختلف، چه کوتاه‌مدت و چه بلندمدت، توسط مدل پیشنهادی قابل درک، تفسیر و پیش‌بینی است.

رویکرد اتخاذ شده در این مقاله مستقل از بازار بوده و می‌توان آن را در آینده برای سایر بازارها مانند بازارها سهام نیز گسترش داد. همچنین به دلیل ساختار قطعه‌قطعه‌ای در نظر گرفته شده، می‌توان هر یک از بخش‌ها در هر یک از چهار مرحله مذکور را بنا بر نیاز تغییر داده، به‌روز کرده و یا بهبود داد. به‌عنوان مثال می‌توان از جاسازی‌های مدرن‌تر استفاده کرده و یا در صورت وجود منابع پردازشی، مدل‌های پیچیده‌تری را نیز آزمایش کرد.

ازآنجاکه ساختار طراحی شده امکان پذیرش هر نوع داده‌ای را میسر می‌سازد، با افزودن انواع دیگری از داده‌ها که از مراجع مختلف جمع‌آوری شده‌اند، احتمالاً کارایی مدل‌ها افزایش خواهد داشت. به‌عنوان مثال بسیاری از اخبار در شبکه‌های اجتماعی و کانال‌های پیام‌رسان‌ها منتشر می‌شوند. داده‌های به‌دست آمده از این منابع می‌تواند به غنای مجموعه داده بیفزاید.

در نهایت، این مقاله گامی در جهت توسعه مدل‌های هوشمندتر و جامع‌تر برای مدلسازی اقتصادی است که می‌تواند داده‌های متنی را نیز در تحلیل‌ها و پیش‌بینی‌های خود دخیل کند. پیشنهاد می‌شود که در تحقیقات آینده از داده‌های متنی گسترده‌تر و مدل‌های پیچیده‌تری استفاده شود تا به نتایج دقیق‌تر و کاربردی‌تری در حوزه مالی و اقتصادی دست یابیم.

تعارض منافع

تعارض منافعی وجود ندارد.

ORCID

Elmira Asle Roosta https://orcid.org/0009-0006-2926-1273
Alireza Erfani https://orcid.org/0000-0003-1493-216X
Abdolmohammad Kashian https://orcid.org/0000-0002-5352-1446

منابع

ابونوری، اسماعیل، خانعلی‌پور، امیر و عباسی، جعفر. (۱۳۸۸). اثر اخبار بر نوسانات نرخ ارز در ایران: کاربردی از خانواده ARCH. *پژوهشنامه بازرگانی*، ۵۰، ۱۰۱-۱۲۰.
dor: 20.1001.1.17350794.1388.13.50.4.8

امیری، فرهاد، درخشانی درآبی، کاوه و آسایش، حمید. (۱۳۹۹). بررسی آپار نوسانات نرخ ارز بر ارزش افزوده در زیربخش‌های اقتصاد ایران. *فصلنامه علمی مطالعات اقتصادی کابردی ایران*، ۱۰ (۳۹)، ۲۴۷-۲۶۷

بیات، ندا. (۱۳۹۷). پیش‌بینی نرخ ارز با استفاده از نقشه‌های خودسازمان‌ده بازگشتی. *اقتصاد و تجارت نوین*، ۱۳، ۵۵-۸۴

تقوی، مهدی و خدام، محمود. (۱۳۹۰). بررسی تطبیقی کارآمدی نظریه‌های ارزی در پیش‌بینی تغییرات نرخ ارز در بازار تبادلات بین‌المللی ارز. *دانش مالی تحلیل اوراق بهادار (مطالعات مالی)*، ۹، ۱۴۷-۱۹۲.

خاشعی، مهدی، بیجاری، مهدی و مخاطب رفیعی، فریماه. (۱۳۹۲). پیش‌بینی نرخ ارز با به‌کارگیری مدل‌های ترکیبی پرسپترون‌های چندلایه (MLPs) و طبقه‌بندی‌کننده‌های عصبی احتمالی (PNNs). *فصلنامه روش‌های عددی در مهندسی*، ۳۲(۱)، ۱۴-۱

خداویسی، حسن و ملابهرامی، احمد. (۱۳۹۱). مدل‌سازی و پیش‌بینی نرخ ارز براساس معادلات دیفرانسیل تصادفی. *تحقیقات اقتصادی*، ۱۰۰ (۴۷)، ۱۲۹-۱۴۴.
doi: 10.22059/jte.2012.29257

رحیمی‌بروجردی، علیرضا. (۱۳۷۹). نظام مطلوب ارزی و تنظیم و پیش‌بینی نرخ ارز برای اقتصاد ایران. *پژوهش‌های اقتصادی ایران*، ۵، ۴۰-۴۵.

زراءنژاد، منصور، فقه‌مجیدی، علی و رضایی، روح‌الله. (۱۳۸۷). پیش‌بینی نرخ ارز با استفاده از شبکه‌های عصبی مصنوعی و مدل ARIMA. *اقتصاد مقداری (بررسی‌های اقتصادی)*، ۱۹، ۱۰۷-۱۳۰.

شریف‌مقدم، شفق و هاشمی، سیدذبیح‌اله. (۱۳۹۷). پیش‌بینی نرخ ارز یورو به دلار با تکنیک شبکه عصبی مصنوعی. *مهندسی مالی و مدیریت اوراق بهادار*، ۹ (۳۷)، ۳۹۹-۴۱۳.

شیرازی، همایون و نصرالهی، خدیجه. (۱۳۹۲). مدل‌های پولی و پیش‌بینی نرخ ارز در ایران: از تئوری تا شواهد تجربی. *فصلنامه سیاست‌های مالی و اقتصادی*، *۱* (۴)، ۲۴-۵

طیبی، سیدکمیل، موحدنیا، ناصر و کاظمینی، معصومه. (۱۳۸۷). به‌کارگیری شبکه‌های عصبی مصنوعی در پیش‌بینی متغیرهای اقتصادی و مقایسه آن با روش‌های اقتصادسنجی: پیش‌بینی روند نرخ ارز در ایران. *مهندسی صنایع و مدیریت (ویژه علوم مهندسی شریف)*، ۴۳، ۱۰۴-۱۹۹.

فتاحی، شهرام، احمدی، آرش و اکرم‌میرزایی، علی. (۱۳۹۲). مقایسه دقت روش الگوریتم ژنتیک با روش‌های دیگر پیش‌بینی‌های نرخ ارز. *مطالعات و سیاست‌های اقتصادی*، ۹۶، ۱۱۳-۱۳۶. doi: 10.22096/esp.2013.26156

گودرزی فراهانی، یزدان، عادلی، امیدعلی و قربانی، عاطفه. (۱۳۹۹). تأثیر نااطمینانی سیاست‌های اقتصادی بر نوسانات نرخ ارز با استفاده از رویکرد مدل خودرگرسیون با وقفه‌های توزیعی غیرخطی (NARDL). *مدل‌سازی اقتصادسنجی*، ۵(۴)، ۱۴۷-۱۷۱.

doi: 10.22075/jem.2021.22243.1547

مرزبان، حسین، جواهری، بهنام و اکبریان، رضا. (۱۳۸۴). یک مقایسه بین مدل‌های اقتصادسنجی ساختاری، سری زمانی و شبکه عصبی برای پیش‌بینی نرخ ارز. *مجله تحقیقات اقتصادی*، ۴۰ (۲).

dor: 20.1001.1.00398969.1384.40.2.8.6

منصوری گرگری، حامد و خداویسی، حسن. (۱۳۹۸). پیش‌بینی نرخ ارز: مقایسه الگوهای رشد لجستیک با الگوهای رقیب. *اقتصاد و الگوسازی*، ۱۰، ۱۴۱-۱۷۹. doi: 10.48308/ecoj.10.3.157

هاشمی‌دیزج، عبدالرحیم، حاضری‌نیری، هاتف و پوروحدانی، رسول. (۱۳۹۹). مقایسه عملکرد مدل‌های شبکه‌های عصبی مصنوعی برای پیش‌بینی نرخ ارز در ایران. *دوفصلنامه علمی مطالعات و سیاست‌های اقتصادی*، *۷* (۲)، ۵۳-۸۰ doi: 10.22096/esp.2020.43397

یارمحمدی، مسعود و محمودوند، رحیم. (۱۳۹۴). پیش‌بینی نرخ ارز با استفاده از روش تحلیل مجموعه مقادیر تکین. *مجله اقتصاد*، ۱۸، ۱۳۷-۱۴۶. doi: 10.22084/aes.2016.1497

References

Abounouri, A., Khanalipour, A. & Abbasi, J. (2009). The effect of news on exchange rate volatility in Iran: An application of the ARCH family. *Pajouheshnameh Bazargani*, 50, 101-120. [In Persian] dor:20.1001.1.17350794.1388.13.50.4.8

Amiri, Farhad, Derakhshani-Daraabi, Kaveh, & Asayesh, Hamid. (2020). Investigation the Effects of Exchange Rate Fluctuations on the Sub-Sectors Value Added in Iran. *Iranian Journal of Applied Economic Studies*, *10*(39), 247–267. [In Persian]

Bayat, N. (2018). Forecasting exchange rates using recurrent self-organizing maps. *Economics and Modern Trade*, 13, 55-84. [In Persian]

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901. doi:10.48550/arXiv.2005.14165

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv*: 1810.04805. doi:10.48550/arXiv.1810.04805

Dridan, R. & Oepen, S. (2012). Tokenization: Returning to a long solved problem—a survey, contrastive experiment, recommendations, and toolkit. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers).

Farahani, M., Gharachorloo, M., Farahani, M. & Manthouri, M. (2021). ParsBERT: Transformer-based model for Persian language understanding. *Neural Processing Letters*, 53, 3831-3847. doi:10.48550/arXiv.2005.12515

Fatahi, Sh., Ahmadi, A. & Akram-Mirzaei, A. (2013). Comparison of the accuracy of genetic algorithm methods with other exchange rate forecasting methods. *Economic Studies and Policies*, 96, 113-136. [In Persian]. doi:10.22096/esp.2013.26156

Goudarzi Farahani, G., Adeli, O.A. & Ghorbani, M. (2020). The impact of economic policy uncertainty on exchange rate volatility using the NARDL model approach. *Econometric Modeling*, *5*(4), 147-171. [In Persian]. doi:10.22075/jem.2021.22243.1547

Hashemi-Dizaj, A., Hazari-Neiri, H. & Pourvohedani, R. (2020). Comparing the performance of artificial neural network models for forecasting exchange rates in Iran. *Scientific Biannual Journal of Economic Studies and Policies*, *7*(2), 53-80. [In Persian]. doi:10.22096/esp.2020.43397

Khashaei, Mehdi, Bijari, Mehdi, & Mokhatab-Rafi'i, Farimah. (2013). Forecasting the Exchange Rate Using Hybrid Models of Multilayer Perceptrons (MLPs) and Probabilistic Neural Network Classifiers (PNNs). *Numerical Methods in Engineering Quarterly*, *32*(1), 1–14. [In Persian]

Khodaveisi, H. & Molabahrami, A. (2012). Modeling and forecasting exchange rates based on stochastic differential equations. *Economic Research*, *100*(47), 129-144. [In Persian]. doi:10.22059/jte.2012.29257

Heaton, J. (2016). An empirical analysis of feature engineering for predictive modeling. *SoutheastCon*, 1-6. doi:10.48550/arXiv.1701.07852

Josse, J. & Husson, F. (2012). Handling missing values in exploratory multivariate data analysis methods. *Journal de la Société Française de Statistique*, *153*(2), 79-99.

Khan, W., Daud, A., Khan, K., Muhammad, S. & Haq, R. (2023). Exploring the frontiers of deep learning and natural language processing: A

comprehensive overview of key challenges and emerging trends. *Natural Language Processing Journal*, 100026. doi:10.1016/j.nlp.2023.100026

Li, Y. & Yang, T. (2018). Word embedding for understanding natural language: A survey. *Guide to Big Data Applications*, 83-104. doi:10.1007/978-3-319-53817-4

Liu, J., Li, T., Xie, P., Du, S., Teng, F. & Yang, X. (2020). Urban big data fusion based on deep learning: An overview. *Information Fusion*, 53, 123-133. doi:10.1016/j.inffus.2019.06.016

Mansouri-Gorgori, H. & Khodavisi, H. (2019). Exchange rate forecasting: Comparison of logistic growth models with competing models. *Economics and Modeling*, 10, 141-179. [In Persian]. doi:10.48308/ecoj.10.3.157

Marzban, D.H., Javaheri, B.B. & Akbarian, A. (2005). A comparison between structural econometric models, time series, and neural networks for exchange rate forecasting. *Economic Research Journal*, *40*(2). [In Persian]. dor: 20.1001.1.00398969.1384.40.2.8.6

Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv*: 1301.3781. doi:10.48550/arXiv.1301.3781

Mohamadi, S., Mujtaba, G., Le, N., Doretto, G. & Adjeroh, D.A. (2023). ChatGPT in the Age of Generative AI and Large Language Models: A Concise Survey. *arXiv preprint arXiv: 2307.04251*. doi:10.48550/arXiv.2307.04251

Nargesian, F., Samulowitz, H., Khurana, U., Khalil, E.B. & Turaga, D.S. (2017). Learning feature engineering for classification. *IJCAI*, 2529-2535. doi:10.24963/ijcai.2017/352

Potdar, K., Pardawala, T.S. & Pai, C.D. (2017). A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications*, *175*(4), 7-9. doi:10.5120/IJCA2017915495

Rahimi-Boroujerdi, A. (2000). The optimal exchange system and forecasting exchange rates for Iran's economy. *Iranian Economic Research*, 5, 40-45. [In Persian].

Seabe, P.L., Moutsinga, C.R.B. & Pindza, E. (2023). Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: A deep learning approach. *Fractal and Fractional*, *7*(2), 203. doi:10.3390/fractalfract7020203

Sharif-Moghaddam, Sh. & Hashemi, S.A. (2018). Forecasting euro to dollar exchange rate using artificial neural networks. *Financial Engineering and Securities Management*, *9*(37), 399-413. [In Persian].

Shirazi, Homayoun, & Nasrollahi, Khadijeh. (2013). Monetary Models and Exchange Rate Forecasting in Iran: From Theory to Empirical Evidence. *Financial and Economic Policies Quarterly*, *1*(4), 5–24. [In Persian]

Singh, D. & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97, 105524. doi:10.1016/j.asoc.2019.105524

Sousa, M.G., Sakiyama, K., de Souza Rodrigues, L., Moraes, P.H., Fernandes, E.R. & Matsubara, E.T. (2019). BERT for stock market sentiment analysis. 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), 1597-1601. doi.org:10.1109/ICTAI.2019.00231

Taghavi, M. & Khodam, M. (2011). Comparative efficiency of exchange rate theories in predicting exchange rate changes in the international foreign exchange market. *Financial Knowledge of Securities Analysis (Financial Studies)*, 9, 147-192. [In Persian].

Tayebi, S., Moheddinia, N. & Kazemini, M. (2008). Utilizing artificial neural networks in forecasting economic variables and comparison with econometric methods: Forecasting exchange rate trends in Iran. *Industrial Engineering and Management (Sharif Special Edition for Engineering Sciences)*, 43, 104-199. [In Persian]

Villamil, L., Bausback, R., Salman, S., Liu, T.L., Horn, C. & Liu, X. (2023). Improved stock price movement classification using news articles based on embeddings and label smoothing. *arXiv preprint arXiv:2301.10458*. doi:10.48550/arXiv.2301.10458

Yarmohammadi, M. & Mahmoodvand, R. (2015). Forecasting exchange rates using singular value decomposition analysis. *Journal of Economics*, 18, 137-146. [In Persian].

Zaranejad, M., Fagh-Majidi, A. & Rezaei, R.A. (2008). Exchange rate forecasting using artificial neural networks and ARIMA models. *Quantitative Economics (Economic Studies)*, 19, 107-130. [In Persian]

استناد به این مقاله: اصل روستا، المیرا، عرفانی، علیرضا و کاشیان، عبدالمحید. (۱۴۰۴). پیش‌بینی نرخ ارز در ایران با استفاده از تلفیق داده‌ها و مدل جامع مبتنی بر یادگیری ماشین. *پژوهش‌های اقتصادی ایران*، ۳۰(۱۰۳)، ۱۰۱-۱۳۷.