

A Novel Multi-Step Ahead Demand Forecasting Model Based on Deep Learning Techniques and Time Series Augmentation

Hossein Abbasimehr*, Reza Paki**

*Faculty of Information Technology and Computer Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran

**Graduated from Shahid Madani University of Azerbaijan

Abstract

In a business environment where there is fierce competition between companies, accurate demand forecasting is vital. If we collect customer demand data at discrete points in time, we obtain a demand time series. As a result, the demand forecasting problem can be formulated as a time series forecasting task. In the context of time series forecasting, deep learning methods have demonstrated good accuracy in predicting complex time series. However, the excellent performance of these methods is dependent on the amount of data available. For this purpose, in this study, we propose to use time series augmentation techniques to improve the performance of deep learning methods. In this study, three new methods have been used to test the effectiveness of the proposed approach, which are: 1) Long short-term memory, 2) Convolutional network 3) Multihead self-attention mechanism. This study also uses a multi-step forecasting approach that makes it possible to predict several future points in a forecasting operation. The proposed method is applied to the actual demand data of a furniture company. The experimental results show that the proposed approach improves the forecasting accuracy of the methods used in most different prediction scenarios. Also, the results on three public time series datasets confirms the superiority of the proposed method.

Keywords: Time Series, Deep Learning, Long short-term memory, Convolutional network, Multihead self-attention mechanism.

مدل جدید پیش بینی چند گامی تقاضا با استفاده از روش های یادگیری عمیق و تکنیک های داده افزایی سری زمانی

حسین عباسی مهر*، رضا پاکی**

*دانشکده فناوری اطلاعات و مهندسی کامپیوتر، استادیار، دانشگاه شهید مدنی آذربایجان، تبریز، ایران

**دانش آموخته، دانشگاه شهید مدنی آذربایجان

تاریخ دریافت: ۱۴۰۰/۰۶/۱۹ تاریخ پذیرش: ۱۴۰۰/۱۲/۰۷

نوع مقاله: پژوهشی

چکیده

در یک محیط تجاری که رقابت سختی بین شرکت ها وجود دارد، پیش بینی دقیق تقاضا یک امر مهمی است. اگر داده های مربوط به تقاضای مشتری را در نقاط گسسته ای از زمان جمع آوری کنیم، یک سری زمانی تقاضا به دست می آید. در نتیجه، مسئله پیش بینی تقاضا به عنوان یک مسئله پیش بینی سری های زمانی فرموله می شود. در زمینه پیش بینی سری های زمانی، روش های یادگیری عمیق دقت مناسبی در پیش بینی سری های زمانی پیچیده داشته اند. با این وجود عملکرد خوب این روش ها به میزان داده های در دسترس وابسته است. بدین منظور در این مطالعه استفاده از روش های داده افزایی سری زمانی در کنار روش های یادگیری عمیق پیشنهاد می شود. در این مطالعه سه روش نوین جهت تست کارایی رویکرد پیشنهادی به کار گرفته شده است که عبارت اند از: (۱) حافظه کوتاه مدت طولانی، (۲) شبکه کانولوشنی (۳) مکانیزم خودتوجه چند سر. همچنین در این مطالعه رویکرد پیش بینی چندگامی به کار گرفته می شود که امکان پیش بینی چند نقطه آینده را در یک عمل پیش بینی به وجود می آورد. روش پیشنهادی بر روی داده واقعی تقاضای یک شرکت مبلمان اعمال شده است. نتایج آزمایش ها نشان می دهد که رویکرد پیشنهادی باعث بهبود دقت پیش بینی روش های به کار گرفته شده در اکثر حالات مختلف پیش بینی می شود. همچنین بکارگیری روش پیشنهادی روی مجموعه داده های محک استاندارد، حاکی از عملکرد موفق مدل های بدست آمده نسبت به مدل های پایه می باشد.

واژگان کلیدی: سری زمانی، یادگیری عمیق، حافظه طولانی کوتاه-مدت، شبکه کانولوشنی، مکانیزم خودتوجه چندسر.

۱. مقدمه

تقاضا را به عنوان یک مسئله پیش بینی سری های زمانی فرموله کرد [۳]. سری های زمانی در حوزه های کاربردی مختلفی از جمله، پیش بینی مصرف بار الکتریکی (به عنوان مثال [۴، ۵])، پیش بینی تماس های ورودی مرکز تماس [۶]، تقاضای گردشگری [۷]، پیش بینی تقاضای پول نقد دستگاه خودپرداز [۸]، پیش بینی حجم تولید نفتی [۹]، پیش بینی هوا [۱۰]، پیش بینی رفتار مشتریان بانکی [۱۱] پیش بینی تعداد مبتلایان بیماری کووید-۱۹ [۱۲] و غیره به کار گرفته شده است.

در زمینه پیش بینی سری های زمانی، روش های یادگیری عمیق دقت مناسبی در پیش بینی سری های زمانی پیچیده داشته اند [۲، ۱۳،

پیش بینی تقاضا پایه تمام فعالیت های برنامه ریزی یک شرکت است [۱]. رقابت سخت میان شرکت ها در تمام زمینه ها، پیش بینی دقیق تقاضاهای مشتری با استفاده از شیوه های رایج تقاضا را برای کسب و کارها دشوار ساخته است؛ بنابراین، شرکت ها از روش های پیشرفته علم داده به منظور پیش بینی تقاضای مشتریان خود استفاده می کنند [۲]. اگر داده های مربوط به تقاضای مشتری را در نقاط گسسته ای از زمان (برای مثال روزانه، هفتگی، ماهانه) جمع آوری کنیم، یک سری زمانی تقاضا به دست می آید. بنابراین، می توان مسئله پیش بینی

بخش ۴ به توصیف داده ها، مراحل پیاده سازی پرداخته و نتایج مدل ها را مورد مقایسه و تحلیل قرار می دهد. در بخش ۵، روش پیشنهادی روی سه مجموعه داده محک استفاده شده و نتایج آن گزارش می شود. نتیجه گیری و ارائه پیشنهاد کارهای آتی در بخش ۶ ارائه می شود.

۲. تحقیقات مرتبط و مفاهیم پایه

در این بخش ابتدا به مرور تحقیقات مرتبط با پیش بینی سری های زمانی با استفاده از روش های یادگیری عمیق پرداخته و سپس روش های مورد استفاده در این مطالعه را توصیف می کنیم.

۲.۱ تحقیقات مرتبط در زمینه پیش بینی سری زمانی با روش های یادگیری عمیق

در سال های اخیر از روش های یادگیری عمیق در زمینه پیش بینی سری زمانی به طور فزاینده ای استفاده شده است.

جدول ۱ برخی از تحقیقات در زمینه پیش بینی سری زمانی با استفاده از آن روش ها را نمایش می دهد. همانطور که در جدول مشاهده می شود، روش LSTM پرکاربردترین روش از دسته روش های یادگیری عمیق در حوزه پیش بینی سری زمانی است. همچنین روش های مبتنی بر شبکه کانولوشنی، مکانیزم خودتوجه چندسره و روش های ترکیبی متشکل از این روش ها در تحقیقات قبلی مورد استفاده قرار گرفته است.

۲.۲ روش حافظه طولانی کوتاه مدت

مدل LSTM یک نوع شبکه عصبی بازگشتی^۵ بهبودیافته است که توسط هوچریتز و اشمیت^۷ [۲۱] توسعه داده شده است. این مدل قابلیت قوی در پیش بینی سری های زمانی دارد LSTM می تواند اطلاعات وابستگی های طولانی مدت در داده های سری زمانی را استخراج کرده و به طور مناسب بین داده ورودی و خروجی نگاشت ایجاد کند [۲۲]. همانطور که در شکل ۱ نشان داده شده است LSTM از چهار بخش سلول حافظه C ، گیت ورودی i ، گیت فراموشی f و گیت خروجی o تشکیل شده است [۲۲]. در ادامه با در نظر گرفتن پارامترهای زیر به توصیف عملیاتی که در مدل LSTM انجام می شود می پردازیم

۱۴. مهم ترین خصوصیت روش های یادگیری عمیق توانایی آن ها در یادگیری بازنمایی از داده های ورودی خام است. در این مطالعه سه روش نوین جهت پیش بینی تقاضا بکارگرفته شده است که عبارت اند از: (۱) حافظه طولانی کوتاه مدت^۲، (۲) شبکه کانولوشنی^۳ مکانیزم توجه چند سر^۴.

علی رغم عملکرد خوب این روش ها در پیش بینی دقیق یک سری زمانی، قدرت پیش بینی آن ها به میزان داده های در دسترس وابسته است. معمولا سری های زمانی واقعی دارای طول کوتاه هستند و در نتیجه تعداد نمونه داده های استخراجی از آن ها جهت ساخت مدل پیش بینی ممکن است کافی نباشد [۱۴]. بدین منظور در این مطالعه استفاده از روش های داده افزایی سری زمانی^۵ [۱۵] در کنار روش های یادگیری عمیق پیشنهاد می شود. سه روش یادگیری عمیق شامل روش LSTM، روش مبتنی بر شبکه کانولوشنی و همچنین روش مبتنی بر توجه چندسره برای بررسی کارایی روش پیشنهادی پیاده سازی می شوند. در این مطالعه رویکرد پیش بینی چندگامی به کار گرفته می شود، رویکرد پیش بینی چندگامی، امکان پیش بینی چند نقطه آینده را در یک عمل پیش بینی به وجود می آورد بر خلاف استراتژی پیش بینی تک گامی که در هر پیش بینی فقط یک نقطه آینده را پیش بینی می کند. نوآوری های این مطالعه عبارتند:

(۱) ترکیب روش های داده افزایی سری زمانی با روش های یادگیری عمیق برای حل مشکل کمبود داده در سری های زمانی
(۲) استفاده از مدل مکانیزم خود توجه چندسره برای پیش بینی سری زمانی
(۳) طراحی مدل های یادگیری عمیق جهت انجام پیش بینی داده ها به صورت چندگامی
(۴) بکارگیری روش پیشنهادی روی سری های زمانی محک و تایید عملکرد موفق آن

روش پیشنهادی بر روی داده واقعی تقاضای یک شرکت مبلمان اعمال شده است. نتایج نشان می دهد که رویکرد پیشنهادی باعث بهبود دقت پیش بینی روش های به کار گرفته شده در اکثر حالات مختلف پیش بینی می شود. همچنین بکارگیری روش پیشنهادی روی مجموعه داده های محک استاندارد، حاکی از عملکرد موفق مدل های بدست آمده نسبت به مدل های پایه می باشد.

ساختار ادامه این مقاله بدین گونه است. بخش ۲ مرور ادبیاتی جامع از پیش بینی سری زمانی با استفاده از روش های یادگیری عمیق را ارائه می دهد. در بخش ۳، رویکرد پیشنهادی را توصیف می کنیم.

⁵ Time series augmentation

⁶ Recurrent Neural Network (RNN)

⁷ Hochreiter and Schmidhuber

² Long short-term memory (LSTM)

³ Convolutional Neural Network (CNN)

⁴ Multi head self-attention

جدول ۱. برخی از تحقیقات در زمینه پیش‌بینی سری زمانی با استفاده از روش‌های یادگیری عمیق

مرجع	نوآوری	نتایج	روش یادگیری عمیق
[۹]	ارائه یک معماری LSTM چند لایه برای پیش‌بینی سری زمانی	نتایج این مطالعه نشان‌دهنده این است که روش LSTM چند لایه ارائه شده، نسبت به روش آماری ARIMA و روش‌های استفاده شده در تحقیقات دیگر عملکرد موفق‌تری از خود نشان می‌دهد.	LSTM چند لایه
[۱۶]	ارائه یک مدل LSTM دو لایه جهت پیش‌بینی سری زمانی دمای پوسته ژئروسکوپ	در مقایسه با روش‌هایی شامل SVM و شبکه عصبی، روش LSTM عملکرد بهتری از خود نشان می‌دهد.	LSTM
[۱۷]	به‌کارگیری LSTM برای انجام پیش‌بینی شاخص‌های مالی	نتایج روی مجموعه داده مالی نشان می‌دهد که روش LSTM قادر به استخراج ویژگی‌های از سری‌های زمانی دارای نویز می‌باشد. همچنین LSTM نسبت به روش‌هایی نظیر شبکه عصبی استاندارد، رگرسیون لجستیک و روش جنگل تصادفی، عملکرد بهتری از خود نشان می‌دهد.	LSTM
[۱۴]	ارائه چارچوبی برای خوشه‌بندی سری‌های زمانی و سپس ایجاد نمونه داده از هر خوشه جهت ایجاد مدل پیش‌بینی	نتایج این مطالعه نشان می‌دهد که در اکثر سناریوها، ترکیب روش‌های خوشه‌بندی با LSTM بهتر از روش LSTM پایه عمل می‌کند.	LSTM
[۲]	ارائه یک مدل LSTM بهینه چند لایه جهت پیش‌بینی سری زمانی تقاضا	معماری چند لایه LSTM نتایج بهتری نسبت به سایر روش‌های مورد استفاده نظیر روش RNN، شبکه عصبی چندلایه پرسپترون، SVM، KNN و روش‌های آماری دارد	LSTM
[۱۸]	توسعه یک روش ترکیبی مبتنی بر شبکه کانولوشنی و شبکه LSTM جهت پیش‌بینی قیمت طلا	نتایج نشان می‌دهد که مدل LSTM-CNN نتایج بهتری نسبت به مدل‌های پایه دارد.	LSTM شبکه کانولوشنی
[۱۹]	پیش‌بینی شاخص سهام با استفاده از شبکه‌های کانولوشنی	نتایج این تحقیق نشان‌دهنده قدرت مدل‌های CNN در پیش‌بینی قیمت سهام با استفاده از مجموعه داده انتخاب شده متعلق به یک دوره چهار ساله بورس ملی هند است.	شبکه کانولوشنی
[۲۰]	پیش‌بینی تعداد مبتلایان کووید-۱۹ با روش‌های یادگیری عمیق	نتایج این تحقیق حاکی از آن است که روش LSTM دو طرفه بهترین عملکرد را در بین دیگر روش‌ها از جمله LSTM، روش SVR و روش ARIMA بدست آورده است.	LSTM Bidirectional LSTM شبکه عصبی بازگشتی ^۱ مبتنی بر گیت
[۱۲]	توسعه مدل‌های مبتنی بر یادگیری عمیق جهت پیش‌بینی سری زمانی مبتلایان کووید-۱۹	روش‌های مبتنی بر یادگیری عمیق عملکرد موفق‌تری در زمینه پیش‌بینی سری زمانی مبتلایان کووید-۱۹ در ۱۰ کشور با بالاترین تعداد ابتلا داشتند. روش‌های مورد استفاده در این تحقیق نسبت به مدل ارائه شده در دیگر تحقیقات عملکرد بهتری داشته‌اند.	LSTM، روش مبتنی بر مکانیزم توجه چندسره، روش مبتنی بر شبکه کانولوشنی

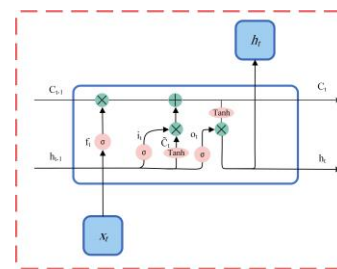
c_t و c_{t-1} حالات سلول در نقطه زمان‌های t و $t-1$

$b = \{b_a, b_f, b_c, b_o\}$ بردارهای سوگیری گیت ورودی، گیت فراموشی، سلول حافظه و گیت خروجی.

$W = \{W_a, W_f, W_c, W_o\}$ ماتریس‌های وزنی گیت ورودی، گیت فراموشی، سلول حافظه و گیت خروجی.

$U = \{U_a, U_f, U_c, U_o\}$ وزن‌های بازگشتی گیت ورودی، گیت فراموشی، سلول حافظه و گیت خروجی

h_t ، خروجی واحد LSTM از طریق رابطه (۱) محاسبه می‌شود:



شکل ۱. ساختار شبکه [LSTM 22]

x_t : مقدار ورودی در نقطه زمانی t

h_t و h_{t-1} مقدار خروجی در نقطه زمان‌های t و $t-1$

^۱ Gated Recurrent Unit (GRU)

نسبت به داده های خام ورودی داشته و باعث بهبود دقت پیش بینی می شود.

۴.۲ روش مبتنی بر مکانیزم خودتوجه چندسر

در سال های اخیر بکارگیری مکانیزم توجه در حوزه پردازش زبان طبیعی موفقیت آمیز بوده است [۲۵، ۲۶]. مطالعه وسوانی^۱ و همکاران [۲۷] کارایی مکانیزم توجه را در پردازش داده های ترتیبی به خوبی نشان داده است. یک تابع توجه یک کوئری Q به همراه یک مجموعه از کلید-مقدارها $\langle K, V \rangle$ را دریافت می کند تا خروجی O را محاسبه کند. ورودی لایه خودتوجه سه پارامتر با نام های key-value-query هستند. این سه ورودی می توانند منبع یکسانی داشته باشند و یا از لایه های قبلی گرفته شوند. در پیش بینی سری های زمانی از یک ورودی استفاده می کنیم.

۳. رویکرد پیشنهادی و معماری مدل های مورد استفاده

در این مطالعه ابتدا سه روش یادگیری عمیق شامل روش LSTM، روش شبکه کانولوشنی و همچنین روش خودتوجه چندسر پیاده سازی می شود؛ سپس یک رویکرد مبتنی بر داده افزایی سری زمانی برای بهبود عملکرد این روش ها ارائه می شود.

۳.۱ رویکرد پیشنهادی

همانطور که قبلا اشاره شد، یک سری زمانی دنباله ای از مشاهدات بر روی یک متغیر مورد نظر است که به صورت منظم در نقاط گسسته ای از زمان گردآوری می شوند. در یک مسئله پیش بینی سری زمانی با استفاده از روش های یادگیری ماشین، سری به بخش آموزش و بخش تست تقسیم می شود. بخشی از داده های آموزش برای ارزیابی و انتخاب مدل بهینه انتخاب می شوند که به آن ها داده های اعتبارسنجی گفته می شود. در یک سری زمانی بخش تست از نقاط انتهایی سری انتخاب می شود و داده های نزدیک به بخش تست تاثیر بیشتری در دقت مدل دارند. داده های از نوع سری زمانی دارای خصوصیات متفاوتی نسبت به داده های غیرزمان دار دارند. در یک مسئله یادگیری ماشین مرتبط با داده های غیرزمانی هر چقدر داده ها بیشتر باشد معمولا دقت مدل بهتر می شود اما این موضوع در زمینه داده های سری زمانی وابسته به ویژگی های سری است. همچنین به دلیل اینکه طول بعضی از سری های زمانی کوتاه است، مدل های یادگیری ماشین/عمیق با مشکل بیش برآزش مواجه می شوند. برای

$$h_t = o_t \tanh(c_t) \quad (1)$$

جایی که o_t اطلاعاتی خروجی واحد LSTM را تنظیم می کند و c_t حافظه واحد LSTM در زمان t است. o_t با رابطه (۲) محاسبه می شود

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (2)$$

که در آن σ به عنوان تابع سیگموئید است.

سلول حافظه، c_t با فراموش کردن بخشی از اطلاعات حافظه کنونی، c_{t-1} و اضافه کردن محتوای جدید حافظه، \tilde{c}_t آپدیت می شود (رابطه (۳)).

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \quad (3)$$

که محتوای جدید حافظه، \tilde{c}_t با استفاده از رابطه (۴) زیر بدست می آید:

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

اینکه تا چه حدی اطلاعات حافظه فعلی فراموش شود توسط گیت فراموشی، f_t تعیین می شود؛ و همچنین اینکه تا چه میزان محتوای حافظه جدید به سلول حافظه اضافه شود توسط گیت ورودی، i_t مشخص می شود. گیت های خروجی و ورودی به ترتیب توسط روابط (۵) و (۶) محاسبه می شود

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (5)$$

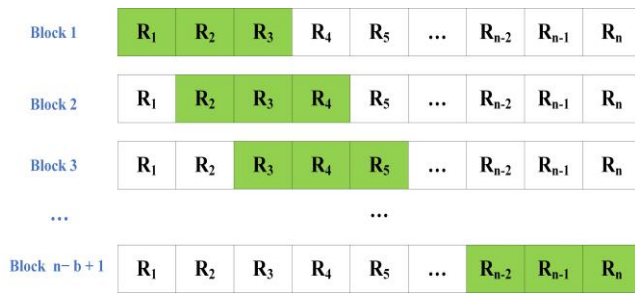
$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (6)$$

که در آن σ به عنوان تابع سیگموئید است [۲۳].

۳.۲ روش مبتنی بر شبکه کانولوشنی

CNN کاربردهای موفقی در حوزه مختلف مخصوصا در زمینه بینایی ماشین داشته است [۲۳]. یک لایه کانولوشنی داده های ورودی خام را دریافت کرده و با انجام عملیات کانولوشن با استفاده از کرنل های کانولوشن ویژگی های جدیدی استخراج می کند. هر لایه کانولوشن شامل یک کرنل (یک پنجره کوچک) است که روی داده های حرکت کرده و از طریق انجام عملیات کانولوشنی ویژگی های جدید را محاسبه می کند [۲۴]. ویژگی های جدید قابلیت متمایزسازی بالایی

¹ Vaswani



شکل ۱. فرآیند ایجاد بلوک‌های همپوشان از سری زمانی [۳۰]

سپس برای ساخت سری نمونه برداری شده، به طور تصادفی $k = n / b$ بلوک (با جایگزینی) انتخاب شده و تلفیق می‌شوند.

سری بدست آمده از طریق داده‌افزایی همانطور که در شکل ۲ نشان داده شده است جهت ایجاد مجموعه داده اعتبار سنجی در مدل‌سازی‌ها به کار گرفته می‌شود.

الگوریتم ۱. شبه کد روش پیشنهادی

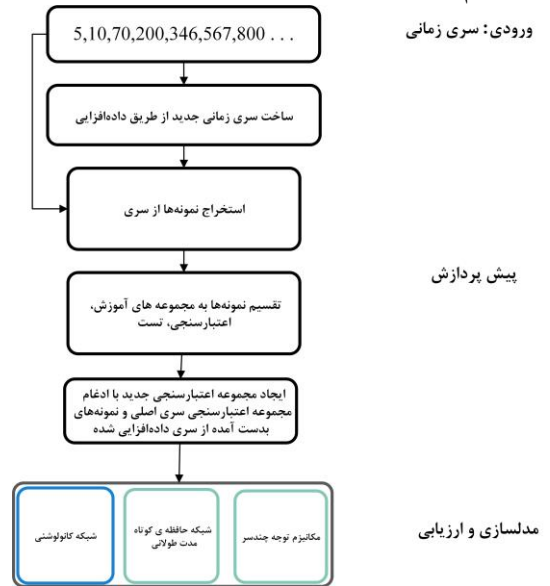
Algorithm 1:

Input: Time series (TS), n_Boot, Lag, Output_Window

Output: Bootstrapped Validation Set

- 1- BC_TS=**Box-Cox_transform** (TS)
- 2- [Trend, Seasonal, Remainder] = **Decompose** (BC_TS)
- 3- **For** i in 1 to n_Boot:
- 4- Boot reminder[i]=**MBB** (Reminder)
- 5- Series_BC[i]=Trend + Seasonal + Boot reminder [i]
- 6- New_Series[i]=**Invert_Box-Cox_transform** (Series_BC[i])
- 7- **End**
- 8- Augmented_series=**Mean** (New_Series)
- 9- [train,val, test]=**GenerateSamples** (ts, Lag, Output_Window)
- 10- [aug_train,aug_val, aug_test]=**GenerateSamples** (Augmented_series, Lag, Output_Window)
- 11- Val_new= **Concat** (val, val_aug)
- 12- **Return** Val_new

حل این مشکل در این مطالعه استفاده از روش‌های داده‌افزایی سری زمانی و استفاده از آن‌ها به عنوان داده اعتبارسنجی پیشنهاد می‌شود. روال کلی ایده پیشنهادی در شکل ۲ و شبه کد آن در الگوریتم ۱ نمایش داده شده است.



شکل ۲. روال کلی روش پیشنهادی

در رویکرد پیشنهادی برای انجام داده‌افزایی سری زمانی از الگوریتم داده‌افزایی توسعه داده شده در [۱۵] استفاده می‌کنیم. این الگوریتم ابتدا یک تبدیل Box-Cox [۲۸] روی سری انجام داده و سپس سری را به اجزای آن شامل تغییرات فصلی^۱، روند^۲ و باقی مانده^۳ با استفاده از روش‌های تجزیه سری STL^۴ یا loess [۲۹] بر اساس نوع سری تجزیه می‌کند. برای سری‌های دارای تغییرات فصلی از روش STL جهت تجزیه سری به اجزا آن و همچنین برای سری‌های غیرفصلی از روش loess جهت تجزیه سری به اجزا روند و باقی مانده استفاده می‌شود. سپس نمونه برداری از روی سری باقی مانده را با استفاده از روش نمونه برداری بلوک متحرک (MBB)^۵ [۳۰] انجام داده و مولفه تغییرات فصلی و روند را به سری نمونه‌برداری شده اضافه می‌کند و در آخر تبدیل معکوس Box-Cox روی سری بدست آمده انجام می‌دهد. در روش نمونه‌برداری بلوک متحرک، ابتدا با در نظر گرفتن اندازه بلوک، b ، برای یک سری زمانی به طول n ، $n - b - 1$ بلوک همپوشان ساخته می‌شود. شکل ۱ بلوک‌های همپوشان را برای حالتی که $b = 3$ است، نشان می‌دهد. نقاط $R_1 : R_b$ بلوک اول را نشان می‌دهد، نقاط $R_2 : R_{b+1}$ بلوک دوم را نشان می‌دهد و به همین ترتیب، بلوک آخر شامل نقاط $R_{n-b+1} : R_n$ است.

⁴ Seasonal and Trend decomposition using Loess

⁵ Moving Block Bootstrapping

¹ Seasonality

² Trend

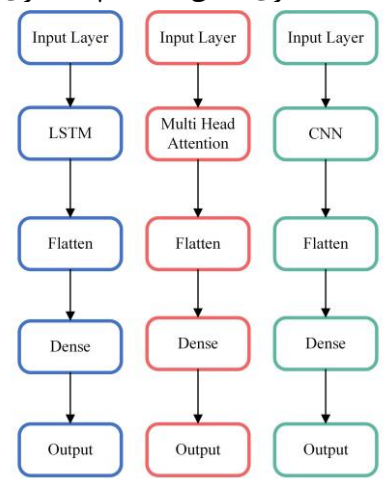
³ Residual

۲.۳ معماری مدل های مورد استفاده جهت تست رویکرد پیشنهادی

همانطور که در ابتدای این بخش بیان شد، در این مطالعه سه روش یادگیری عمیق شامل روش LSTM، روش شبکه کانولوشنی و همچنین روش خودتوجه چندسره جهت سنجش کارایی رویکرد پیشنهادی پیاده سازی می شود. رویکرد پیشنهادی در این پژوهش به عنوان یک روش پیش پردازش محسوب می شود و مجموعه داده مناسب برای اعتبارسنجی مدل ها را فراهم می کند. معماری روش های پیاده سازی شده در شکل توصیف شده است. همانطور که در شکل قابل مشاهده است، هر کدام از مدل ها دارای یک لایه ورودی، یک لایه اصلی، لایه مسطح ساز^۱، لایه کاملاً متصل^۲ و لایه خروجی هستند. لایه ورودی نمونه داده های استخراج شده از سری زمانی را به عنوان ورودی دریافت کرده و هر کدام از لایه های اصلی کار یادگیری بازنمایی از داده های ورودی را انجام می دهند و ویژگی های جدیدی را استخراج می کنند. ویژگی های بدست آمده توسط لایه مسطح ساز در یک برداری مسطح شده و به لایه کاملاً متصل خوراند می شود. گره های این لایه به لایه خروجی متصل شده و کار پیش بینی را انجام می دهند. تعداد گره های لایه خروجی نشان دهنده تعداد گام های پیش بینی مدل است.

۴. آزمایش های عملی و نتایج

در این بخش ابتدا به توصیف مجموعه داده مورد استفاده جهت انجام آزمایش ها پرداخته و سپس نتایج عملکرد مدل ها بیان شده و مورد تحلیل قرار می گیرد. الگوریتم های مورد استفاده در این تحقیق با استفاده از کتابخانه یادگیری عمیق Keras^۳ پیاده سازی شده اند.



شکل ۳. معماری مدل های پیاده سازی شده

۱.۴ مجموعه داده مطالعه موردی

داده های مورد استفاده در این مطالعه، داده های تقاضای یک محصول پرفروش یک شرکت مبلمان است که در طی ۱۱ سال به طور روزانه جمع آوری شده است. این داده ها به صورت ماهانه تجمیع شده و در قالب یک سری زمانی ماهانه مورد استفاده قرار گرفتند. تعداد نقاط زمانی این سری ۱۳۲ ماه است. در این مطالعه ۱۲ نقطه انتهایی سری به عنوان بخش تست در نظر گرفته شده و مابقی داده ها برای آموزش و ساخت مدل ها بکار گرفته شده است.

۲.۴ تنظیم پارامترها

در مسئله پیش بینی سری های زمانی با استفاده از روش های یادگیری عمیق ابرپارامترهای متعددی وجود دارد که انتخاب بهینه آن ها تاثیر زیادی در افزایش دقت مدل ساخته ایجاد می کند. به جهت اینکه مدل های یادگیری عمیق شامل پارامترهای زیادی هستند، آموزش آن ها زمان بر بوده و امکان بررسی تمامی ترکیب های مختلف ابرپارامترها وجود ندارد؛ بنابراین در این مطالعه از یک الگوریتم بهینه سازی بیزین برای انتخاب مدل ها استفاده کردیم. در

جدول ۲ ابرپارامترهای مورد استفاده در این مطالعه برای هر روش و محدوده هر کدام از آن ها توصیف شده است. برخی از ابرپارامترها در بین تمامی مدل ها مشترک است. ابرپارامتر اندازه پنجره در ساخت نمونه ها از سری زمانی به کار می رود و انتخاب بهینه آن نقش مهمی در دقت مدل های ساخته شده دارد.

جدول ۲. ابرپارامترهای بهینه شده توسط بهینه ساز بیزین

محدوده ابرپارامتر	ابروپارامتر
<ul style="list-style-type: none"> اندازه پنجره: [۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵] تعداد تکرار (Epoch): حداکثر ۵۰۰ - بر اساس تغییرات خطای مدل روی داده های اعتبارسنجی تنظیم می شود. نرخ یادگیری: [۰.۰۰۰۱، ۰.۰۰۰۵، ۰.۰۰۱، ۰.۰۰۰۵، ۰.۰۰۱] بهینه ساز: Adam [۲۳] تابع فعال سازی لایه های کاملاً متصل و خروجی: [ReLU, Linear] 	ابروپارامترهای مشترک بین تمامی مدل ها
<ul style="list-style-type: none"> تعداد واحدها: [۴، ۸، ۱۶، ۳۲، ۶۴، ۱۲۸] تابع فعال سازی: [ReLU, Linear] نرخ dropout: [۰.۰، ۰.۱، ۰.۲، ۰.۳، ۰.۴، ۰.۵] 	مدل های مبتنی بر LSTM
<ul style="list-style-type: none"> اندازه کرنل: [۲، ۳، ۴] تعداد فیلترها: [۳۲، ۶۴، ۱۲۸، ۲۵۶] 	مدل های مبتنی بر CNN
<ul style="list-style-type: none"> تابع فعال سازی: [ReLU, Linear] 	مدل های مبتنی بر مکانیزم توجه چند سر

³ <https://keras.io/>

¹ Flatten

² Fully connected (Dense)

$$SMAPE = \frac{1}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{\frac{(|A_t| + |F_t|)}{2}} \times 100 \quad (۷)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \times 100 \quad (۸)$$

که در این رابطه A_t و F_t ، به ترتیب مقادیر واقعی و پیش‌بینی در نقطه زمانی t هستند.

۵.۴ نتایج

همانطور که قبلاً بیان شد در مطالعه از رویکرد پیش‌بینی چندگام آینده استفاده می‌کنیم. بنابراین، برای سنجش کارایی روش پیشنهادی، در این مطالعه مدل‌ها بر اساس تعداد خروجی‌های مختلف مورد ارزیابی قرار گرفته‌اند. پرکاربردترین روش در مطالعات قبلی استفاده از طول گام ۱ است. برای انجام پیش‌بینی به اندازه افق h ، طول گام انتخابی بایستی مضربی از h باشد. با توجه به اینکه طول بخش تست در این مطالعه برابر ۱۲ است، بنابراین طول گام‌های مورد آزمایش برابر ۱، ۲، ۳، ۴، ۶ انتخاب شدند. در این مطالعه علاوه بر سه مدل پایه شامل مدل مبتنی بر حافظه طولانی کوتاه‌مدت (LSTM) مدل مبتنی بر شبکه کانولوشنی (CNN) و مدل مبتنی بر مکانیزم خودتوجه چند سر (ATT)، سه مدل جدید با ترکیب رویکرد پیشنهادی و مدل‌های پایه شامل مدل LSTM_AUG (LSTM + داده افزایشی)، مدل CNN_AUG (CNN + داده افزایشی) و مدل ATT_AUG (ATT + داده افزایشی) پیاده‌سازی شدند.

نتایج آزمایش‌ها در

جدول ۳ بر مبنای معیار SMAPE نمایش داده شده است. همانطور که مشاهده می‌شود، روش پیشنهادی بیشترین بهبود را روی ATT گذاشته است. روش ATT_AUG در هر ۵ حالت بهتر از روش توجه عمل کرده است که این بیان‌گر تاثیر روش پیشنهادی روی این مدل است. برای روش LSTM همانطور که مشاهده می‌کنید در سه حالت خروجی ۱، ۲ و ۳ روش LSTM_AUG باعث بهبود نتایج شده است. همچنین برای روش CNN در ۴ حالت خروجی ۱، ۲، ۳ و ۴ روش CNN_AUG منجر به افزایش دقت مدل گشته است. به طور کلی از بین ۱۵ ترکیب از مدل‌ها، در ۱۲ مورد مدل‌های حاصله از داده‌افزایی بهترین نتیجه را نسبت به مدل‌های پایه بدست آورده‌اند. همچنین مقایسه تمامی مدل‌ها نشان می‌دهد که برای پیش‌بینی با

۳.۴ اقدامات مربوط به جلوگیری از بیش‌برازش مدل‌ها

مدل‌های پیچیده‌ای نظیر شبکه عصبی بدون تنظیم مناسب با خطر بیش‌برازش مواجه هستند. بیش‌برازش زمانی اتفاق می‌افتد که هنگام آموزش مدل که در یک تکرار مشخصی، خطا روی داده‌های آموزشی کاهش می‌یابد اما خطا روی داده‌های اعتبارسنجی (که به آن خطای تعمیم^۱ می‌گویند) شروع به افزایش می‌کند [۳۱]. بنابراین انجام اقداماتی جهت تنظیم شبکه مخصوصاً زمانی که تعداد داده‌ها کم باشد، اهمیت زیادی دارد [۱۴]. در این مطالعه برای جلوگیری از خطر بیش‌برازش مدل‌ها، از روش توقف زودهنگام^۲ [۳۱] استفاده می‌شود. روش توقف زودهنگام عملکرد بهتری نسبت به سایر روش‌های تنظیم آموزش شبکه دارد [۳۱]. مراحل این روش به صورت زیر است:

- ۱- تقسیم داده‌های آموزشی به مجموعه‌های داده آموزش و اعتبارسنجی
- ۲- آموزش مدل روی داده‌های آموزشی و محاسبه خطا روی مجموعه داده اعتبارسنجی
- ۳- متوقف کردن آموزش مدل به محض اینکه خطا روی مجموعه اعتبارسنجی در تعداد تکرار مشخصی (مثلاً ۱۰ تکرار بعدی)، نسبت به خطای اندازه‌گیری شده تکرارهای قبلی بهبود پیدا نکند.
- ۴- در نظر گرفتن مدل بدست آمده در تکرار قبلی به عنوان مدل نهایی

همچنین در این مطالعه، برای جلوگیری از بیش‌برازش، لایه حذف تصادفی^۳ در پیکره‌بندی تمامی مدل‌های استفاده شده گنجانده شده است.

۴.۴ معیار ارزیابی

در این مطالعه از معیار میانگین درصد قدرمطلق خطا متقارن ($SMAPE^4$) و میانگین درصد قدرمطلق خطا ($MAPE^5$) که معیارهای رایج در زمینه اندازه‌گیری دقت مدل‌ها در پیش‌بینی سری‌های زمانی هستند استفاده می‌کنیم. این معیارها با رابطه‌های ۷ و ۸ محاسبه می‌شود:

⁴ Symmetric Mean Absolute Percentage Error

⁵ Mean Absolute Percentage Error

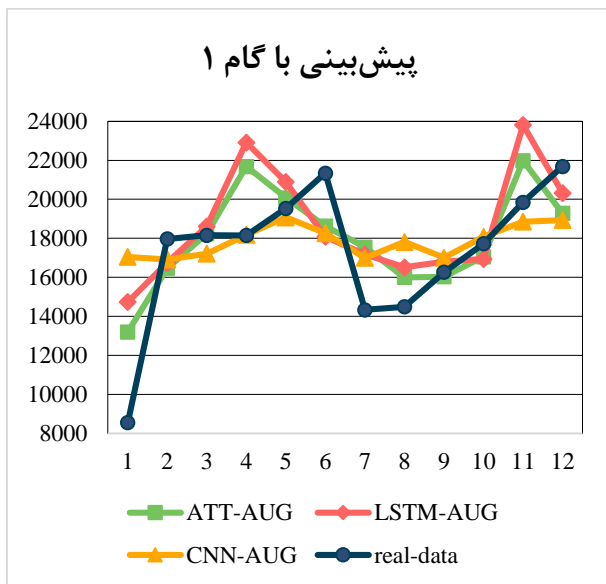
⁶ Forecast horizon

¹ Generalization error

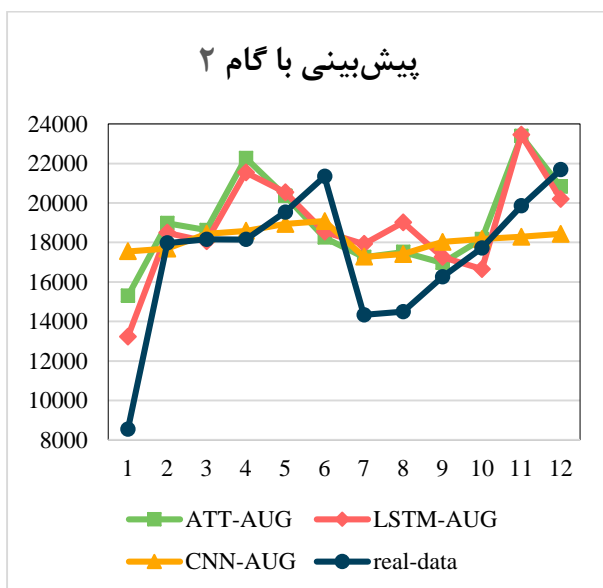
² Early stopping

³ Dropout

گام ۱، روش ATT_AUG کمترین خطا (MAPE = ۱۱,۹۵) را بدست آورده است. برای خروجی ۲ و ۳ روش CNN_AUG بهترین مدل ها را تولید کرده است. برای خروجی ۴ و ۶ به ترتیب روش LSTM و روش CNN بهترین نتایج را کسب کرده اند. ترکیب این نتایج نشان دهنده مفید بودن روش پیشنهادی است. همانطور که قبلا بیان شد، برای جلوگیری از بیش برآزش مدل های مبتنی بر یادگیری عمیق، نیاز است تا داده اعتبارسنجی مورد استفاد متفاوت از داده های سری انتخاب شود. همچنین جدول ۴ نتایج عملکرد مدل ها را بر مبنای معیار MAPE نشان می دهد. ATT_AUG و CNN_AUG در هر ۵ حالت نسبت به روش های پایه خود یعنی روش ATT و CNN عملکرد بهتری را داشته اند. بررسی نتایج روش LSTM_AUG نشان می دهد که این روش برای گام های ۱، ۲ و ۳ دارای بیشترین بهبود نسبت روش پایه، یعنی LSTM داشته است. همچنین برای گام ۴، این روش اندکی موفق تر از روش پایه ظاهر شده است.



شکل ۴. نمودار داده های واقعی و پیش بینی با گام ۱



شکل ۵. نمودار داده های واقعی و پیش بینی با گام ۲

جدول ۳ نتایج عملکرد مدل ها را بر مبنای معیار MAPE نشان می دهد. ATT_AUG و CNN_AUG در هر ۵ حالت نسبت به روش های پایه خود یعنی روش ATT و CNN عملکرد بهتری را داشته اند. بررسی نتایج روش LSTM_AUG نشان می دهد که این روش برای گام های ۱، ۲ و ۳ دارای بیشترین بهبود نسبت روش پایه، یعنی LSTM داشته است. همچنین برای گام ۴، این روش اندکی موفق تر از روش پایه ظاهر شده است.

جدول ۳. نتایج مدل ها

روش	تعداد خروجی				
	۱	۲	۳	۴	۶
ATT	۱۴,۷۰	۱۴,۷۷	۱۴,۶۹	۱۴,۷۱	۱۵,۹۹
ATT_AUG	۱۱,۹۵	۱۴,۱۲	۱۴,۱۰	۱۳,۸۵	۱۵,۶۰
LSTM	۱۴,۵۵	۱۴,۷۴	۱۵,۱۱	۱۳,۴۷	۱۴,۱۸
LSTM_AUG	۱۴,۴۰	۱۴,۵۸	۱۴,۲۳	۱۴,۰۲	۱۴,۷۹
CNN	۱۳,۵۵	۱۴,۵۸	۱۳,۴۱	۱۵,۳۵	۱۳,۹۹
CNN_AUG	۱۳,۲۲	۱۳,۵۹	۱۲,۶۷	۱۵,۱۲	۱۴,۲۷

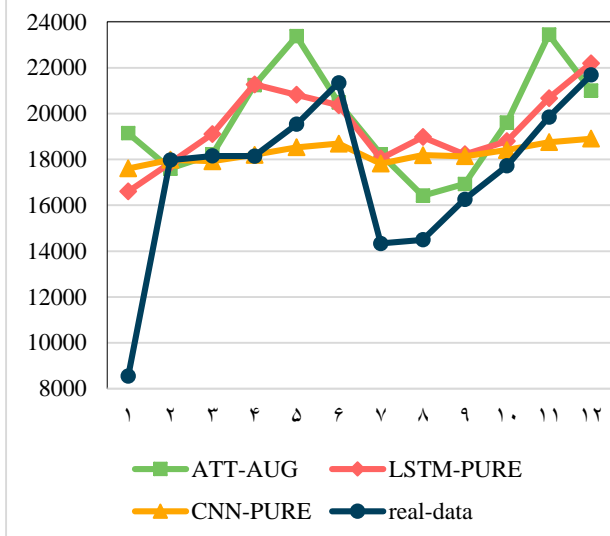
جدول ۴. نتایج مدل ها بر اساس معیار MAPE

روش	تعداد خروجی				
	۱	۲	۳	۴	۶
ATT	۱۷,۱۰	۱۷,۸۴	۱۷,۹۰	۱۷,۸۹	۲۰,۴۹
ATT_AUG	۱۳,۱۹	۱۶,۰۶	۱۶,۲۹	۱۶,۶۵	۲۰,۳۲
LSTM	۱۷,۰۳	۱۹,۰۹	۱۸,۳۰	۱۷,۱۵	۱۷,۰۱
LSTM_AUG	۱۶,۰۱	۱۵,۷۳	۱۶,۴۰	۱۷,۱۷	۱۷,۴۷
CNN	۱۶,۰۱	۱۷,۴۹	۱۵,۸۵	۱۸,۲۲	۱۷,۴۱
CNN_AUG	۱۶,۱۱	۱۶,۸۱	۱۵,۳۳	۱۸,۳۰	۱۷,۳۸

به منظور تحلیل بیشتر نتایج، برای هر طول گام، پیش بینی حاصل از بهترین مدل ها و داده های واقعی در شکل های ۴، ۵، ۶، ۷ و ۸ ترسیم شده است. همانطور که برای گام ۱ مشخص روش ATT_AUG پیش بینی نزدیک تری نسبت به داده واقعی (real-)

دارند. این مجموعه داده‌ها به صورت گسترده در تحقیقات مربوط به سری زمانی مورد استفاده قرار گرفته‌اند. در این خصوص نتایج بدست آمده در تحقیق [۳۲] جهت مقایسه عملکرد مدل‌ها مورد استفاده قرار می‌گیرد. برای ایجاد شرایط یکسان در آزمایشات، تعداد نقاط آموزش و تست در سری زمانی مطابق تحقیق [۳۲] انتخاب شدند. همچنین مشابه تحقیق [۳۲] استراتژی پیش‌بینی یک گامی در نظر گرفته می‌شود. در جدول ۵ توصیف‌های مربوط به هر مجموعه داده به همراه تعداد نقاط آموزش و اعتبارسنجی و آزمون ارائه شده است.

پیش‌بینی با گام ۶

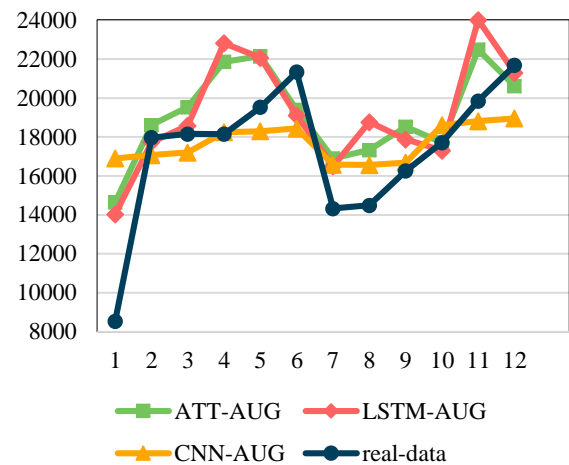


شکل ۸. نمودار داده‌های واقعی و پیش‌بینی با گام ۶

جدول ۵. توصیف مجموعه داده‌های محک

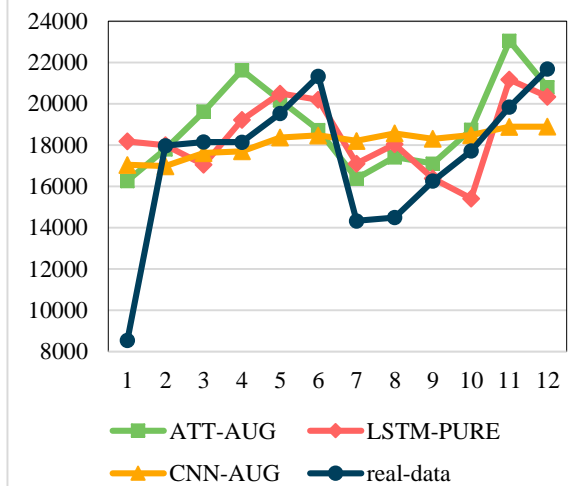
مجموعه داده	توصیف	تعداد نقاط مجموعه آموزش	تعداد نقاط اعتبارسنجی	تعداد نقاط آزمون
Sunspot	سری زمانی تعداد لکه‌های خورشیدی ۱۹۸۷-۱۷۰۰	۱۷۲	۵۸	۵۸
Lynx	تعداد سالانه سیاهگوش‌های به دام افتاده ^۱ ، رودخانه مکنزی، ۱۹۳۴-۱۸۲۱	۶۸	۲۳	۲۳
IBM	سری زمانی قیمت پایانی سهام شرکت IBM در بازه زمانی ۱۷/۵/۱۹۶۱ تا ۱۱/۲/۱۹۶۲	۲۲۱	۷۴	۷۴

پیش‌بینی با گام ۳



شکل ۶. نمودار داده‌های واقعی و پیش‌بینی با گام ۳

پیش‌بینی با گام ۴



شکل ۷. نمودار داده‌های واقعی و پیش‌بینی با گام ۴

۵. ارزیابی روش پیشنهادی روی مجموعه داده‌های محک

در بخش قبلی تاثیر روش پیشنهادی در بهبود دقت پیش‌بینی روش‌های یادگیری عمیق روی مجموعه داده مربوط به شرکت مبلمان مورد بررسی قرار گرفت. در این بخش، عملکرد روش پیشنهادی روی ۳ مجموعه داده محک استاندارد مورد بررسی قرار می‌گیرد. این ۳ مجموعه داده عبارت‌اند از: (۱) مجموعه داده Wolf's sunspot (۲) مجموعه داده Canadian lynx و (۳) مجموعه داده سری زمانی قیمت پایانی سهام شرکت IBM. این سری‌های زمانی مربوط به حوزه‌های متفاوتی هستند و ویژگی‌های آماری متفاوتی

¹ Lynx

همچنین هر سه مدل مبتنی بر روش پیشنهادی شامل LSTM_AUG، CNN_AUG و ATT_AUG دارای عملکرد بهتری نسبت به مدل ارائه شده در تحقیق [32] ETS_ANN هستند.

جدول ۷. مقایسه عملکرد روش پیشنهادی با نتایج مدل های تحقیقات فعلی روی

مجموعه داده Lynx

مدل	SMAPE
LSTM	۵۵,۶۸
LSTM_AUG	۴۹,۶۸
CNN	۵۶,۹۹
CNN_AUG	۴۲,۵۹
ATT	۵۵,۱۸
ATT_AUG	۴۲,۶۰
ETS_ANN[32]	۵۶,۶۷۴

۳.۵ مقایسه عملکرد روش ها روی مجموعه داده

IBM

جدول ۸ نتایج عملکرد مدل های بدست با استفاده از روش پیشنهادی را روی سری زمانی IBM نشان می دهد. همانطور که در جدول مشاهده می شود، هر سه روش مبتنی بر داده افزایی عملکرد بهتری نسبت به روش های پایه خود دارند. مدل های LSTM_AUG، CNN_AUG و ATT_AUG به ترتیب دارای بهبود ۲,۲۴٪، ۶,۲۶٪ و ۶,۷۶٪ نسبت به مدل های پایه خود دارند. همچنین، در مقایسه با بهترین روش موجود در تحقیقات فعلی، یعنی روش ترکیبی ETS-ANN، روش های LSTM_AUG، LSTM_AUG و ATT_AUG خطای کمتری دارند. مدل LSTM_AUG بهترین عملکرد را در بین تمامی مدل ها از خود نشان می دهد.

جدول ۸. مقایسه عملکرد روش پیشنهادی با نتایج مدل های تحقیقات فعلی روی

مجموعه داده IBM

مدل	SMAPE
LSTM	۱,۵۸۲۸
LSTM_AUG	۱,۵۴۷۳
CNN	۱,۹۲۵۴
CNN_AUG	۱,۸۰۴۹
ATT	۱,۶۸۰
ATT_AUG	۱,۵۶۶۴
ETS_ANN[32]	۱,۵۹۵۵

۶ نتیجه گیری

در این پژوهش یک رویکرد مبتنی بر روش های داده افزایی سری زمانی جهت افزایش دقت روش های یادگیری عمیق ارائه شد. رویکرد پیشنهادی با سه مدل نوین یادگیری عمیق شامل حافظه طولانی

در ادامه به ارائه نتایج عملکردی روش پیشنهادی روی هر کدام از مجموعه داده ها می پردازیم. شایان ذکر است که مطابق با تحقیق [۳۲]، در این بخش نتایج عملکرد مدل ها را با استفاده از معیار SMAPE گزارش می دهیم.

۱.۵ مقایسه عملکرد روش ها روی مجموعه داده

Sunspot

جدول ۶ میزان خطای مدل ها را بر اساس معیار SMAPE روی مجموعه داده Sunspot نشان می دهد. مدل ATT_AUG بهترین عملکرد را در بین روش ها دارد. همچنین مقایسه نتایج نشان می دهد که مدل های بدست آمده با روش پیشنهادی نسبت به روش های پایه دارای خطای کمتری هستند. مدل LSTM_AUG نسبت به مدل LSTM دارای بهبود ۲۷,۴۴٪ است؛ همچنین CNN_AUG دارای بهبود ۱۹,۲۹ درصدی نسبت مدل CNN است. علاوه بر این، مدل ATT_AUG دارای بهبود ۸,۱۳٪ نسبت به مدل ATT است. همچنین تمامی مدل های مبتنی بر روش پیشنهادی شامل LSTM_AUG، CNN_AUG و ATT_AUG از مدل ارائه شده در تحقیقات فعلی (مدل ETS_ANN[32]) عملکرد بهتری دارند. به طور کلی نتایج نشان می دهد که روش پیشنهادی روی این مجموعه داده باعث افزایش عملکرد مدل های پایه شده است.

جدول ۶. مقایسه عملکرد روش پیشنهادی با نتایج مدل های تحقیقات فعلی روی

مجموعه داده Sunspot

مدل	SMAPE
LSTM	۳۷,۱۰
LSTM_AUG	۲۶,۹۲
CNN	۳۲,۶۶
CNN_AUG	۲۶,۳۶
ATT	۲۸,۲۹
ATT_AUG	۲۵,۹۹
ETS_ANN[32]	۲۹,۴۵۳

۲.۵ مقایسه عملکرد روش ها روی مجموعه داده

Lynx

نتایج عملکرد مدل ها روی مجموعه داده Lynx براساس معیار SMAPE در جدول ۷ آمده است. برای این مجموعه داده نیز، مدل های مبتنی بر روش پیشنهادی دارای نتایج بهتری نسبت به مدل های پایه دارند. مدل LSTM_AUG با خطای ۴۹,۶۸ دارای ۹,۹ درصد بهبود نسبت به LSTM پایه است. همچنین مدل CNN_AUG با میزان خطای ۴۲,۵۹ بهترین عملکرد را در بین تمامی مدل ها بدست آورده است. علاوه بر این، این مدل نسبت به مدل پایه خود یعنی مدل CNN دارای ۲۵,۲۶ درصد بهبود است.

- [7] R. Law, G. Li, D. K. C. Fong, and X. Han, "Tourism demand forecasting: A deep learning approach," *Annals of Tourism Research*, vol. 75, pp. 410-423, 2019.
- [8] F. Martínez, M. P. Frías, M. D. Pérez, and A. J. Rivera, "A methodology for applying k-nearest neighbor to time series forecasting," *Artificial Intelligence Review*, vol. 52, pp. 2019-2037, 2019.
- [9] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks," *Neurocomputing*, vol. 323, pp. 203-213, 2019.
- [10] I. Maqsood, M. R. Khan, and A. Abraham, "An ensemble of neural networks for weather forecasting," *Neural Computing & Applications*, vol. 13, pp. 112-122, 2004.
- [11] H. Abbasimehr and M. Shabani, "A new framework for predicting customer behavior in terms of RFM by considering the temporal aspect based on time series techniques," *Journal of Ambient Intelligence and Humanized Computing*, 2020/05/05 2020.
- [12] H. Abbasimehr and R. Paki, "Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization," *Chaos, Solitons & Fractals*, vol. 142, p. 110511, 2021.
- [13] K. Bandara, H. Hewamalage, Y.-H. Liu, Y. Kang, and C. Bergmeir, "Improving the Accuracy of Global Forecasting Models using Time Series Data Augmentation," *arXiv preprint arXiv:2008.02663*, 2020.
- [14] K. Bandara, C. Bergmeir, and S. Smyl, "Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach," *Expert Systems with Applications*, vol. 140, p. 112896, 2020/02/01/ 2020.
- [15] C. Bergmeir, Rob J. Hyndman, and José M. Benítez, "Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation," *International Journal of Forecasting*, vol. 32, pp. 303-312, 2016/04/01/ 2016.
- [16] H. Shi, S. Hu, and J. Zhang, "LSTM based prediction algorithm and abnormal change detection for temperature in aerospace gyroscope shell," *International Journal of*
- کوتاه-مدت، شبکه کانولوشنی و روش خودتوجه چندسری پیاده‌سازی شد. روش‌های مورد استفاده در این تحقیق با خروجی چندگانه طراحی شده‌اند طوری که امکان پیش‌بینی چندگام به صورت یک‌جا دارند. همچنین ابرپارامترهای مدل‌های پیاده‌سازی شده با استفاده از بهینه‌ساز بیزین انتخاب شدند. نتایج روی داده‌ها نشان داد که در اکثر حالت‌ها (پیش‌بینی با گام‌های مختلف) رویکرد پیشنهادی باعث افزایش دقت پیش‌بینی مدل‌های پایه می‌شود. این رویکرد برای روش خودتوجه چندسری، شبکه کانولوشنی و حافظه طولانی کوتاه-مدت به ترتیب باعث افزایش دقت در ۵، ۴ و ۳ حالت از ۵ حالت مختلف پیش‌بینی می‌شود. همچنین نتایج روی مجموعه داده‌های محک حاکی از عملکرد موفق مدل‌های بدست آمده در مقایسه با مدل‌های پایه و همچنین مدل برتر موجود در این زمینه است. رویکرد پیشنهادی برای پیش‌بینی سری‌های زمانی در حوزه‌های دیگر نظیر پیش‌بینی بارالکتريکی مصرفی می‌تواند مورد استفاده قرار گیرد.

مراجع

- [1] H. Haberleitner, H. Meyr, and A. Taudes, "Implementation of a demand planning system using advance order information," *International Journal of Production Economics*, vol. 128, pp. 518-526, 2010.
- [2] H. Abbasimehr, M. Shabani, and M. Yousefi, "An optimized model using LSTM network for demand forecasting," *Computers & Industrial Engineering*, vol. 143, p. 106435, 2020/05/01/ 2020.
- [3] M. A. Villegas, D. J. Pedregal, and J. R. Trapero, "A support vector machine for model selection in demand forecasting applications," *Computers & Industrial Engineering*, vol. 121, pp. 1-7, 2018/07/01/ 2018.
- [4] N. J. Johannesen, M. Kolhe, and M. Goodwin, "Relative evaluation of regression tools for urban area electrical energy demand forecasting," *Journal of Cleaner Production*, vol. 218, pp. 555-564, 2019.
- [5] M. Q. Raza, M. Nadarajah, and C. Ekanayake, "Demand forecast of PV integrated bioclimatic buildings using ensemble framework," *Applied energy*, vol. 208, pp. 1626-1638, 2017.
- [6] J. W. Taylor, "A comparison of univariate time series methods for forecasting intraday arrivals at a call center," *Management Science*, vol. 54, pp. 253-265, 2008.

- deep learning," *Ieee Access*, vol. 7, pp. 82390-82412, 2019.
- [26] K. Sangeetha and D. Prabha, "Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-10, 2020.
- [27] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., "Attention is all you need," in *31st International Conference on Neural Information Processing Systems*, Long Beach, California, USA, 2017, pp. 6000–6010.
- [28] G. E. P. Box and D. R. Cox, "An Analysis of Transformations," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 26, pp. 211-243, 1964.
- [29] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, "STL: A seasonal-trend decomposition," *Journal of official statistics*, vol. 6, pp. 3-73, 1990.
- [30] S. N. Lahiri, *Resampling methods for dependent data*: Springer Science & Business Media, 2013.
- [31] L. Prechelt, "Early Stopping — But When?," in *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 53-67.
- [32] S. Panigrahi and H. S. Behera, "A hybrid ETS–ANN model for time series forecasting," *Engineering Applications of Artificial Intelligence*, vol. 66, pp. 49-59, 2017.
- [17] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, pp. 654-669, 2018.
- [18] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN–LSTM model for gold price time-series forecasting," *Neural computing and applications*, vol. 32, pp. 17351-17360, 2020.
- [19] S. Mehtab and J. Sen, "Stock price prediction using convolutional neural networks on a multivariate timeseries," *arXiv preprint arXiv:2001.09769*, 2020.
- [20] F. Shahid, A. Zameer, and M. Muneeb, "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM," *Chaos, Solitons & Fractals*, vol. 140, p. 110212, 2020/11/01/ 2020.
- [21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735-1780, 1997.
- [22] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, pp. 2222-2232, 2017.
- [23] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*: MIT press, 2016.
- [24] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural computation*, vol. 29, pp. 2352-2449, 2017.
- [25] A. Ntakaris, G. Mirone, J. Kannianen, M. Gabbouj, and A. Iosifidis, "Feature engineering for mid-price prediction with