Iranian Electric Industry Journal
Quality & Productivity
P.P 47-59

Vol. 13/ No. 2/ Summer 2024
ISSN 2322-2344
EISSN 2717-1639

Electricity Consumption Forecasting Using a Hybrid Approach Based on Transformer Model and LSTM Neural Network

Mohammad reza Ahmadipour, Esmat Rashedi, Maryam Amoozegar

Kerman - University of Postgraduate Industrial Education and Advanced Technology

Abstract:
In recent years, electricity consumption forecasting has gained significant importance by relying on new technologies and utilizing big data processing methods. Various methods including classical statistical time series analysis techniques, support vector machines, and recurrent neural networks with long short-term memory have been studied in this field. Statistical methods may not be suitable for forecasting and modeling some complex phenomena due to their inability to account for sudden changes. On the other hand, the support vector machine method operates based on increasing the data dimensions. Therefore, in cases where the data has high dimensions, this leads to increasing the complexity of the problem space. Various types of neural networks also face limitations such as vanishing gradients and the inability to account for temporal relationships. To achieve more accurate electricity consumption forecasting, this paper proposes a hybrid approach using the Transformer model and long short-term memory neural network. By addressing the gradient problem and learning complex patterns, this approach offers higher accuracy compared to other methods. Additionally, the Transformer model, with its attention mechanism, has the ability to focus on important components of the data, creating a more interpretable model with high resistance to noise. The proposed method has been evaluated on a standard dataset and compared with existing methods. The results show that this method achieves higher accuracy and lower error in metrics such as mean squared error and mean absolute percentageerror.

Keywords: Predicting time series , Long Short-Term Memory Neural Networks, Transformer Model, Attention Mechanism.

Submit date: 2023/12/31
Accepted date: 2024/10/06

Corresponding author's name: Maryam Amoozegar
Corresponding author's address: Kerman - University of Postgraduate Industrial Education and Advanced Technology

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]

Iranian Electric Industry Journal
Quality & Productivity
ص.ص ۵۹-۴۷

Vol. 13 / No. 2/ Summer 2024
ISSN 2322-2344
EISSN 2717-1639

پیش‌بینی مصرف برق با استفاده از رویکرد ترکیبی مبتنی بر مدل انتقالی و شبکه عصبی با حافظه طولانی کوتاه‌مدت

نوع مطالعه: پژوهشی

محمدرضا احمدی پور[۱]، عصمت راشدی[۱]، مریم آموزگار[۲]*

[۱]دانشکده مهندسی برق و کامپیوتر، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته

[۲]پژوهشکده کامپیوتر و فناوری اطلاعات، پژوهشگاه علوم و نکنولوژی پیشرفته و علوم محیطی، دانشگاه دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته

چکیده: در سال‌های اخیر پیش‌بینی مصرف برق، با تکیه بر فنّاوری‌های جدید و بهره‌گیری از روش‌های پردازش داده‌های بزرگ اهمیت زیادی پیداکرده است. روش‌های زیادی ازجمله روش‌های آماری و کلاسیک مختلف تحلیل سری زمانی، ماشین بردار پشتیبان، و شبکه‌های عصبی بازگشتی و حافظه طولانی کوتاه‌مدت در این حوزه موردبررسی قرارگرفته‌اند. روش‌های آماری به دلیل عدم در نظر گرفتن تغییرات ناگهانی، ممکن است برای پیش‌بینی و مدل‌سازی برخی پدیده‌های پیچیده چندان مناسب نباشند. از طرف دیگر، روش ماشین بردار پشتیبان بر مبنای افزایش ابعاد داده عمل می‌کند. بنابراین، در مواردی که داده‌ها دارای ابعاد بالایی باشند، این امر منجر به پیچیده‌تر شدن فضای مسأله می‌گردد. انواع بسیاری از شبکه‌های عصبی نیز با محدودیت‌هایی مثل محوشدگی گرادیان و عدم توجه به روابط زمانی مواجه هستند. برای پیش‌بینی دقیق‌تر مصرف برق، این مقاله یک رویکرد ترکیبی با استفاده از مدل انتقالی و شبکه عصبی با حافظه طولانی کوتاه‌مدت را پیشنهاد می‌دهد. این رویکرد، با حل مشکل گرادیان و یادگیری الگوهای پیچیده، دقت بالاتری نسبت به روش‌های دیگر ارائه می‌دهد. همچنین مدل انتقالی با استفاده از مکانیزم توجه، توانایی تمرکز بر اجزای مهم داده را داراست و مدلی با تفسیرپذیری بیشتر و مقاومت بالا در مقابل نویز ایجاد می‌کند. روش پیشنهادی روی مجموعه داده استاندارد ارزیابی و با روش‌های موجود مقایسه شده است. نتایج نشان می‌دهد این روش دقت بالاتر و خطای کمتری در معیارهایی مانند میانگین مربعات خطا و میانگین درصد خطای مطلق دارد.

واژه‌های کلیدی: پیش‌بینی سری زمانی -شبکه‌های عصبی حافظه طولانی کوتاه‌مدت- مدل انتقالی-مکانیزم توجه.

تاریخ ارسال: ۱۴۰۲/۱۰/۱۰
تاریخ پذیرش: ۱۴۰۳/۰۷/۱۵
نام نویسنده مسئول: امین سامان فر
نشانی نویسنده مسئول: گروه برق- دانشگاه آزاد اسلامی واحد خرم آباد- ایران

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]

۱_ مقدمه:

پیش‌بینی مصرف برق به معنای تخمین زدن حجم برق مورد نیاز برای یک منطقه در زمان‌های آینده است و از اهمیت بالایی برخوردار است. با داشتن مقدار مصرفی برق آینده می‌توان تولید و توزیع مصرف برق را برنامه‌ریزی و از هدر رفت برق جلوگیری کرد و هزینه‌های اضافی تولید و توزیع برق را مدیریت کرد. همچنین، شرکت‌های توزیع برق می‌توانند شبکه را بهتر مدیریت کرده و مشکلات عدم تطابق عرضه و تقاضا را کاهش دهند. پیش‌بینی مصرف برق همچنین به برنامه‌ریزی بهینه‌تر درزمینه تولید و انتقال برق کمک می‌کند و استفاده از منابع انرژی پاک را ترویج می‌دهد. درنتیجه، بهره‌برداری پایدارتر و حفظ محیط‌زیست را تضمین می‌کند.

روش‌های آماری، توانایی تحلیل داده‌ها و مدل‌سازی آن‌ها را دارند مانند مرجع (توان et al., 1396) که با الگوهای رگرسیون خطی فازی[1] (FLR) میانگین متحرک خودهمبسته یکپارچه [2](ARIMA) و میانگین متحرک خودهمبسته یکپارچه فازی[3] (FARIMA) مصرف برق بخش عمومی را پیش‌بینی می‌کند. سپس با استفاده از معیارهای مختلف انتخاب شده، مصرف برق سایر بخش‌ها را با الگوی دقیق‌تری پیش‌بینی می‌کند. چالش های اساسی روش‌های آماری عبارتند از نیاز به پارامترهای بسیار زیاد و پیچیده و نیز نیاز به انتخاب پارامترهای خاص برای شرایط متفاوت است. روش‌های آماری به دلیل فرضیات ساده سازی شده و محدودیت در مدلسازی رفتارهای پیچیده و غیرخطی داده ها، نمی توانند تغییرات ناگهانی و روندهای پویای موجود در داده های واقعی را به خوبی لحاظ کنند. این روش‌ها معمولا برای داده های ایستا و خطی مناسب هستند اما در مواجهه با داده های پویا، غیرخطی و دارای نویز که در دنیای واقعی رایج هستند، قابلیت پیش بینی دقیق را ندارند. به همین دلیل، نتایج مطلوبی را در پیش بینی برخی پدیده های پیچیده مانند مصرف برق تولید نمی کنند.

روش‌های یادگیری عمیق توانایی یادگیری داده‌های بزرگ و مدل‌سازی با پیچیدگی‌های بالا و ارتباطات غیرخطی را دارند. مانند مرجع (Tarmanini et al., 2023) که از شبکه عصبی مصنوعی[4](ANN) استفاده کرده و دلیل استفاده از شبکه عصبی مصنوعی، دقت بالای آن در مدل‌سازی داده‌های غیرخطی است. این مرجع به روش‌های مبتنی بر یادگیری ماشین برای پیش‌بینی بار با استفاده از دو روش پیش‌بینی میانگین متحرک یکپارچه بازگشتی خودکار و شبکه عصبی مصنوعی پرداخته که شبکه عصبی مصنوعی نتایج بهتری ارائه داده است. شبکه عصبی مصنوعی در صورت عدم تأمین داده‌های کافی، ممکن است به‌درستی عمل نکند. در مطالعه دیگری در (A. Xu) یک ماشین بولتزمن محدود شده[5](RBM) عمیق برای مدل‌سازی و پیش‌بینی مصرف انرژی و مبتنی بر الگوریتم کاهش تخمین گرادیان[6] برای تنظیم پارامترها ارائه شده است. در این مطالعه علاوه بر متغیرهای اقلیمی مانند دما و رطوبت، تأثیر سایر متغیرها مانند عوامل اقتصادی نیز مورد بررسی قرار گرفته است. ماشین بولتزمن محدود شده، در مسائل ساده و کوچک‌تر به‌سادگی آموزش می‌بیند و از الگوریتم کاهش تخمین گرادیان بهره می‌گیرد. اما این روش ممکن است به جای رسیدن به جواب بهینه، به یک نقطه تعادلی ناامیدکننده برسد.

در مرجع (Amalou et al., 2022) تحلیل مقایسه‌ای از چندین مدل یادگیری عمیق، به‌ویژه معماری‌های شبکه عصبی بازگشتی[7] مانند شبکه عصبی بازگشتی ساده، شبکه عصبی حافظه طولانی کوتاه‌مدت (LSTM[8]) و واحد بازگشتی گیتی (GRU[9]) ارائه شده است. همان‌طور که از معماری‌های RNN قابل انتظار است، با کمترین مقدار خطا از نوع مجذور مربعات خطا (RMSE[10]) و بیشترین مقدار امتیاز R2 بین سه معماری، GRU عملکرد بهتری نسبت به RNN ساده و LSTM داشت. این نتیجه می‌تواند به چندین دلیل توجیه شود که مهم‌ترین آن قابلیت غلبه بر مشکل محوشدگی گرادیان[11] توسط شبکه‌های بازگشتی است. GRU نیز ممکن است مشکل حافظه بلندمدت داشته باشد، به این معنا که اطلاعات مربوط به فاصله زمانی بزرگ‌تر را به‌درستی نگه نمی‌دارد. این مشکل می‌تواند باعث عدم یادگیری الگوهای طولانی‌مدت شود. بر اساس شبکه عصبی بازگشتی، LSTM دروازه‌هایی دارد که حالت‌های این شبکه در واحدهای مختلف از آن انتقال داده می‌شوند و هر واحد می‌تواند عملیاتی را برای افزایش یا کاهش اطلاعات انجام دهد. کنترل منطقی واحد دروازه تعیین می‌کند که آیا داده به‌روز شود یا دور انداخته شود. این ساختار مشکلات وزن RNN، محوشدگی و انفجار گرادیان را

[1] - Fuzzy Linear Regression
[2] - Autoregressive integrated moving average
[3] - fractional autoregressive integrated moving-average
[4] - Artificial neural network
[5] - Restricted Boltzmann machine
[6] - Contrastive divergence
[7] - Recurrent neural network
[8] - Long Short-Term Memory
[9] - Gated recurrent unit
[10] - Root mean square error
[11] -Gradient Vanishing

[Downloaded from ieijqp.ir on 2026-06-13]
[DOI: 10.61186/ieijqp.13.2.2]

برطرف می‌کند و باعث می‌شود شبکه بهتر و سریع‌تر همگرا شود. از این‌رو از مدل روش LSTM برای دریافت نتایج دقیق‌تر در پیش‌بینی مصرف برق استفاده می‌شود. مقالات مبتنی بر مدل LSTM، دقت پیش‌بینی بهتری را ارائه کرده‌اند(Cui et al., 2020).

باوجود این، ممکن است حافظه طولانی کوتاه‌مدت با مشکل بیش‌برازش به داده‌های آموزشی مواجه شود. از این‌رو در (Mahjoub et al., 2022) یک مدل پویا بر اساس تحلیل سری زمانی و LSTM معرفی‌شده است تا مصرف برق را پیش‌بینی، نقاط بیشینه مصرف را شناسایی و سپس بار را کاهش دهد. در مدل استفاده‌شده این مرجع، ماژول GRU برای مدل‌سازی تغییرات پویایی در داده‌های مصرف برق تاریخی استفاده می‌شود تا ویژگی‌های بالقوه یادگیری در داده‌های دنباله زمانی بهبود یابد. به طریق مشابه، روش LSTM به دلیل توانایی حفظ و آموزش ویژگی‌های داده‌های ارائه‌شده برای مدت زمان طولانی انتخاب می‌شود. نتایج تجربی در چندین افق زمانی نشان می‌دهد که مدل LSTM نتیجه بهتری نسبت به روش‌های پیش‌بینی GRU و Drop-GRU تولید کرده، خطاهای پیش‌بینی کمتر و دقت بهتری دارد. بنابراین، پیش‌بینی‌های مبتنی بر LSTM اجازه می‌دهد تا تصمیمات به‌صورت پیشگیرانه اتخاذ شده و در مواقعی که مصرف بیش از آستانه مجاز رخ می‌دهد، کاهش بار فعال شود. این امر تأثیر قابل‌توجهی در برنامه‌ریزی کیفیت برق و نگهداری تجهیزات برق خواهد داشت. با این اوصاف، LSTM دارای تعداد زیادی پارامتر قابل تنظیم است که باید به‌درستی تنظیم شوند. تنظیم نادرست پارامترهای LSTM می‌تواند منجر به عملکرد ضعیف یا عدم آموزش مطلوب شبکه شود.

در مرجع (H. Wang et al., 2023)، از یک مدل خودرگرسیونی عمیق جدید با ویژگی شبکه دوسویه LSTM تقویت‌شده-DAFA) (BiLSTMبه‌عنوان یکBiLSTM عمیق جدید استفاده‌شده است. ماژول تبدیل برداری (VA[12]) برای نمایش خصوصیات خطی و غیرخطی با تأخیر زمانی سیگنال‌های ورودی به روشی بدون نظارت به کار گرفته‌شده است. نتایج تجربی و تحلیل‌های آماری این مرجع نشان می‌دهد DAFA-BiLSTM دارای عملکرد تطبیقی خوب و نیز استحکام در محیط‌های پر نویز است. این راهکار ممکن است به شرایط محیطی خاصی که در آن ویژگی‌های زمانی تغییر می‌کنند، حساس باشد. اگر شرایط محیطی تغییرپذیر و پیچیده باشند، مدل ممکن است دچار عملکرد ضعیف‌تری شود و نتواند به‌طور دقیق پیش‌بینی کند.

در تحقیق (D. Wang et al., 2023) بر اساس ایده ادغام اطلاعات چندوجهی، یک مدل نوآورانه جامع برای پیش‌بینی تقاضای کلی انرژی با شبکه عصبی پیچشی و حافظه طولانی کوتاه‌مدت[13](CNN-LSTM) در یک محیط داده‌های چندشکلی ارائه‌شده است. نتایج تجربی نشان می‌دهد که مدل استفاده‌شده مؤثر است و ثابت می‌کند که ترکیب داده‌های سری زمانی و داده‌های متنی می‌تواند به بهبود عملکرد پیش‌بینی کمک کند. شبکه‌های عصبی پیچشی-حافظه طولانی کوتاه‌مدت مشکل کندی در آموزش در دنباله‌های بلند را داراست.

مدل انتقالی در پردازش زبان طبیعی[14](NLP)، به‌عنوان روشی مهم مطرح شده است و عملکرد بهتری نسبت به الگوریتم‌های RNN و Seq2Seq دارد. در (L’Heureux et al., 2022) معماری مبتنی بر مدل انتقالی برای پیش‌بینی بار با افزودن تبدیل N-فضا و طراحی یک روش نوآورانه برای کنترل ویژگی‌های زمینه پیشنهاد شده است. نتایج نشان می‌دهد که رویکرد پیشنهادی این مطالعه با موفقیت بر روی داده‌های سری زمانی با داده‌های زمینه کار می‌کند و از مدل‌های Seq2Seq مدرن عملکرد بهتری دارد. بااین‌حال مدل انتقالی نیاز به حجم بالای داده‌های آموزشی و تنظیم پارامترهای مدل برای هر مسئله خاص دارد.

در یک جمع‌بندی باید ذکر کرد که شبکه عصبی حافظه طولانی کوتاه‌مدت قادر به حفظ اطلاعات بلندمدت و شناسایی الگوها و وابستگی‌های طولانی است، بدین ترتیب به‌خوبی این توانایی را دارد که سری زمانی مصرف برق را پیش‌بینی کند و ضمن حل مشکل محوشدگی گرادیان قابلیت پردازش سری‌های زمانی با طول‌های متغیر و یادگیری الگوهای پیچیده را دارد. از طرف دیگر مدل انتقالی با لایه‌های توجه، به ارتباطات سراسری توجه دقیق‌تری می‌کند و الگوهای پیچیده‌تر را شناسایی می‌کند. لذا به‌واسطه ویژگی‌هایش می‌تواند به‌خوبی از عهده مسئله پیش‌بینی مصرف برق که تحت تأثیر عوامل مختلفی مانند فصل، سال، ایام تعطیل، شرایط جوی، رویدادهای خاص و غیره است، برآید. در واقع مکانیزم توجه[15] مدل انتقالی را قادر می‌سازد تا روی بخش‌های مختلفی از داده‌های ورودی، توجه بیشتری گذاشته و دقت پیش‌بینی را بهبود ببخشد. در این مقاله راهکاری ارائه می‌شود که با ترکیب و تقویت شبکه عصبی حافظه طولانی کوتاه‌مدت با مدل انتقالی پیش‌بینی با دقت بالاتری نسبت به هردو روش به دست می‌آورد.

[12] - vector autoregression
[13] - Both Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM)
[14] - Natural language processing
[15] - Attention Mechanism

[Downloaded from ieijqp.ir on 2026-06-13]
[DOI: 10.61186/ieijqp.13.2.2]

در ادامه در بخش دوم مفاهیم اولیه توضیح داده‌شده است. بخش سوم به تشریح روش پیشنهادی می‌پردازد. در بخش چهارم ارزیابی راهکار پیشنهادی موردبررسی قرار می‌گیرد. بخش انتهایی نیز به بیان نتیجه‌گیری می‌پردازد.

۲_ شبکه عصبی حافظه طولانی کوتاه‌مدت

حافظه طولانی کوتاه‌مدت یک نوع شبکه عصبی بازگشتی است که برای مدل‌سازی داده‌های دنباله‌ای مانند سری‌های زمانی استفاده می‌شود. این شبکه عصبی قادر است الگوهای طولانی‌مدت را در داده‌های دنباله‌ای شناسایی و حفظ و در پیش‌بینی آینده استفاده کند. ساختار این شبکه شامل چندین واحد است که هر واحد شامل یک سلول حافظه و سه گیت (گیت فراموشی، گیت ورودی و گیت خروجی) هستند. سلول حافظه به‌عنوان واحد اصلی این شبکه عمل می‌کند و توانایی ذخیره و حذف اطلاعات را دارد. گیت فراموشی تصمیم می‌گیرد کدام اطلاعات در سلول حافظه حفظ شوند و کدام اطلاعات حذف شوند. گیت ورودی تعیین می‌کند کدام اطلاعات جدید به سلول حافظه اضافه شوند و گیت خروجی تصمیم می‌گیرد کدام اطلاعات از سلول حافظه استخراج‌شده و به لایه خروجی ارسال شوند.

از طرف دیگر، شبکه‌های عصبی بازگشتی در درک ساختارهای بلندمدت ضعیف عمل می‌کنند. اطلاعات باید از طریق نرون‌های متعددی عبور کند تا به نرون در حال پردازش برسد. اما در مرحله انتشار به عقب، این اطلاعات بلندمدت به دلیل ضرب شدن در مقادیر کوچک، به سمت صفر میل می‌کنند و از بین می‌روند. این مسئله به‌عنوان مشکل محوشدگی گرادیان شناخته می‌شود (Ozcan et al., 2021). حالت سلولی در LSTMباعث می‌شود مشکل محوشدگی گرادیان در RNN حل شود. بنابراین LSTM نسبت به RNN عملکرد بهتری در گرفتن اطلاعات تاریخچه‌ای و پیش‌بینی تحولات آینده برای مسائل مربوط به وابستگی‌های طولانی‌مدت در داده‌های توالی دارد. ساختار LSTM با در نظر گرفتن وابستگی بلندمدت در داده‌ها، می‌تواند الگوهای پیچیده‌تری را در مقایسه با RNN ساده یاد بگیرد (Song et al., 2020).

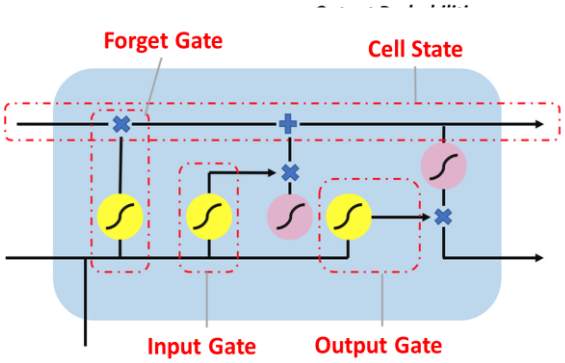
۲-۱- مدل انتقالی

مدل انتقالی یک معماری شبکه عصبی است که برای مدل‌سازی داده‌های دنباله‌ای مانند سری‌های زمانی استفاده می‌شود. این معماری، به‌خصوص برای مسائل ترجمه ماشینی، پرسش و پاسخ و تولید متن توجه بسیاری را به مکان‌های مختلف داده‌های ورودی اعمال می‌کند.

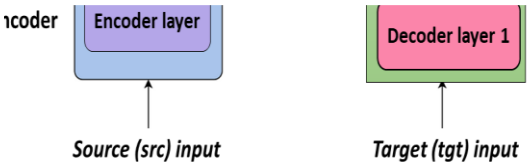
یک سری زمانی ثابت، نباید به زمان مشاهده وابسته باشد. به‌عبارت‌دیگر، در یک سری زمانی ثابت، الگوهای قابل‌تشخیصی برای پیش‌بینی آینده توسط نگاه به داده‌ها وجود ندارند. اگر داده‌ها ثابت باشند، فرآیند پردازش و پیش‌بینی بار آینده به راحتی انجام می‌شود. اغلب داده‌های مربوط به سری زمانی بار، دارای تمام ویژگی‌های روندها، فصلی و نویز به‌طور هم‌زمان هستند (Farsi et al., 2021). معماری مدل انتقالی به شبکه این امکان را می‌دهد تا الگوهای طولانی‌مدت را در داده‌های دنباله‌ای شناسایی کند. مدل انتقالی ویژگی‌ها و مولفه های مهمی مانند لایه توجه دارد. مکانیزم توجه روشی است که ویژگی توجه انسان را تقلید می‌کند. وقتی انسان به‌طور سطحی به یک نقاشی نگاه می‌کند، می‌تواند کل تصویر را مشاهده کند، اما وقتی با دقت به نقاشی نگاه می‌کند، توجه‌اش محدود به برخی نواحی نقاشی می‌شود. به‌عبارت‌دیگر، توجه مغز انسان به‌کل نقاشی متوازن نیست، بلکه تمرکز مشخصی دارد. مکانیزم توجه استفاده از پارامترهای وزنی متفاوت برای هر عنصر ورودی است، به‌گونه‌ای که درحالی‌که به اطلاعات موردنیاز توجه خودکار می‌کند، اطلاعات بی‌فایده دیگر را سرکوب می‌کند (Wang et al., 2022). این لایه به مدل این امکان را می‌دهد تا به تمام ویژگی‌های ورودی به‌طور هم‌زمان توجه کند و از آن‌ها استفاده کند. با توجه به شکل ۲ معماری رمزگذار-رمزگشا از دیگر ویژگی‌های مهم مدل انتقالی است که از دو بخش اصلی تشکیل‌شده است. رمزگذار برای تبدیل دنباله ورودی به نمایش معنایی و ایجاد نمایش انتقالی به‌عنوان خروجی استفاده می‌شود. رمزگشا برای تولید دنباله خروجی بر اساس نمایش انتقالی ایجادشده توسط رمزگذار به کار می‌رود.

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]



شکل۱- معماری شبکه عصبی حافظه طولانی کوتاه‌مدت



شکل ۲- مدل انتقالی

۳_ روش پیشنهادی

در مسئله پیش‌بینی مصرف برق که یک موضوع سری‌زمانی است، چالش‌های مختلفی وجود دارد که روی دقت و کارایی مدل‌ها تأثیرگذار هستند. برخی از این چالش‌ها عبارت‌اند از وابستگی زمانی، نویز، عدم قطعیت، گرادیان محوشونده و حجم زیاد داده‌ها. سری‌زمانی دارای وابستگی‌های زمانی است که باید در مدل‌ها مدنظر قرار بگیرد. معمولاً رویدادهای گذشته در پیش‌بینی آینده تأثیرگذار هستند و مدل‌ها باید بتوانند این وابستگی‌ها را درک کنند. همچنین عوامل ناپایدار، نویز یا انحرافات ناشی از عوامل خارجی هستند که باعث ایجاد قطعیت ناهمگون و پیچیدگی در پیش‌بینی می‌شود. مدل‌ها باید بتوانند این عدم قطعیت را مدیریت کنند و به دقتی که به تعداد زیادی از حالت‌ها پاسخ دهند برسند. در برخی موارد، سری‌زمانی ممکن است با گذشت زمان تغییر کند. به عنوان مثال، الگوها و رفتارها ممکن است در طول زمان تغییر کنند و این تغییرات بر پیش‌بینی دقیق تأثیر می‌گذارند. مدل‌ها باید بتوانند این تغییرات را تشخیص داده و با آن‌ها سازگار شوند. برخی سری‌های زمانی ممکن است در فواصل زمانی نامنظم نمونه‌برداری شوند. این مسئله ممکن است باعث مشکلاتی در مدل‌ها شود، زیرا زمان‌بندی یکنواخت داده‌ها وجود ندارد و مدل باید بتواند با این نامنظمی‌ها سازگاری داشته باشد. سری‌زمانی معمولاً دارای حجم بالایی از داده است که می‌تواند به مشکلات محاسباتی و مدیریتی منجر شود. مدل‌ها باید قدرت محاسباتی مناسبی داشته باشند و روش‌های مؤثری برای مقابله با حجم بالای داده‌ها را ارائه دهند.

مشکل گرادیان محوشونده یکی از چالش‌های اصلی در آموزش شبکه‌های عصبی عمیق است که می‌تواند از یادگیری الگوهای پیچیده و وابستگی‌های طولانی مدت در داده‌ها جلوگیری کند. در راهکار اول، شبکه LSTM برای مقابله با این مشکل استفاده شد. در این شبکه از تکنیک‌هایی مانند گیت‌های ورودی، خروجی و فراموشی و لایه‌های نرمال‌سازی و تابع فعال‌ساز ReLU که پیاده‌سازی شده بودند، استفاده گردید. این تکنیک‌ها نقش مهمی در حفظ جریان گرادیان در طول لایه‌های عمیق، انتقال مستقیم اطلاعات از ورودی به خروجی، پایدارسازی توابع فعال‌ساز و در نهایت جلوگیری از محو شدن گرادیان ایفا می‌کنند. شبکه LSTM با داشتن حافظه سلولی و گیت فراموشی، قابلیت حفظ اطلاعات برای دوره‌های طولانی و تصمیم‌گیری در مورد اطلاعات قابل فراموشی یا حفظ از گذشته را دارد که به آن کمک می‌کند تا الگوها و روابط زمانی طولانی مدت مهم را به درستی شناسایی و حفظ کند.

برای پیش‌بینی مصرف برق، دو چالش اساسی نویز و عدم قطعیت وجود دارد. راهکار دوم که برای مقابله با این چالش‌ها استفاده شد، مدل انتقالی با مکانیزم توجه بود. مکانیزم توجه در این مدل می‌تواند به شناسایی و تمرکز بر روی ویژگی‌های مهم و مرتبط با الگوهای مصرف برق، مانند دما، رطوبت، تاریخ و ساعت، کمک کند. این مکانیزم با محاسبه وزن‌های متغیر برای هر ویژگی ورودی بر اساس اهمیت آن در پیش‌بینی مصرف برق عمل می‌کند. بنابراین، مدل می‌تواند بر ویژگی‌های با وزن توجه بالاتر که بیشتر با الگوهای مصرف برق مرتبط هستند، تمرکز کند و از ویژگی‌های نامرتبط یا نویزی چشم‌پوشی کند. این راهکار نه تنها می‌تواند تفسیرپذیری را از طریق بررسی توزیع وزن‌های توجه افزایش دهد که درک عملکرد مدل را در پیش‌بینی مصرف برق آسان‌تر می‌کند، بلکه مقاومت در برابر نویز را با نادیده گرفتن ویژگی‌های نویزی به ارمغان می‌آورد. علاوه بر این، از آنجا که بر ویژگی‌های کلیدی مرتبط با مصرف برق متمرکز می‌شود، ظرفیت مدلسازی روابط طولانی‌مدت را که برای پیش‌بینی مصرف برق بسیار مهم است، افزایش می‌دهد. مدل انتقالی توجهی همچنین می‌تواند در برطرف کردن چالش عدم قطعیت در داده‌ها نیز کمک کننده باشد، زیرا توانایی مدلسازی روابط پیچیده و طولانی مدت را دارد و می‌تواند با داده‌های نویزی به شکل انعطاف‌پذیری برخورد کند.(Niu et al., 2021)

به منظور ارائه راهکار جامعی برای رفع کلیه چالش های پیش بینی مورد بحث، روش پیشنهادی در این تحقیق ترکیبی از شبکه عصبی بازگشتی حافظه طولانی کوتاه مدت و مدل انتقالی را پیشنهاد می‌کند. این مدل برای پیش‌بینی دقیق آینده در یک دنباله، طراحی‌شده است. برای

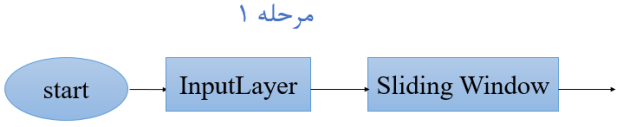
[Downloaded from ieijqp.ir on 2026-06-13]
[DOI: 10.61186/ieijqp.13.2.2]

برطرف نمودن برخی چالش‌های پیش‌بینی مصرف برق، از مدل انتقالی در روش ترکیبی استفاده شد که از لایه‌های مدل انتقالی برای استخراج ویژگی‌ها و مدل‌سازی روابط غیرخطی در داده‌ها استفاده می‌کند. این لایه‌ها توانایی ارتباط بلندمدت در داده‌ها را دارند و بهبود قابل‌توجهی در پیش‌بینی دقیق مرحله آینده ایجاد می‌کنند. روش پیشنهادی از لایه‌های شبکه عصبی حافظه طولانی کوتاه‌مدت که قابلیت حفظ اطلاعات بلندمدت و شناسایی الگوهای پیچیده و وابستگی‌های زمانی را در داده‌ها دارد، استفاده می‌کند. روش پیشنهادی امکان استفاده از قابلیت‌های هر دو مدل را فراهم می‌کند و به‌دقت و کارایی پیش‌بینی کمک می‌کند. روش پیشنهادی از لایه‌های شبکه عصبی چندلایه[16] با تعداد واحدهای قابل تنظیم استفاده می‌کند. این لایه‌ها توانایی نمایش و تخمین توابع پیچیده را دارند و به مدل انعطاف‌پذیری بیشتری می‌دهند. این نوآوری‌ها باعث بهبود دقت و عملکرد مدل در پیش‌بینی می‌شوند. جزئیات روش در ادامه آورده شده است.

ورودی مدل یک سری زمانی چند یا تک متغیره با ابعاد (n_steps, n_features) است که تعداد گام‌های زمانی با (n_steps) و تعداد ویژگی‌ها با (n_features) نشان داده می‌شود. لایه رمزگذار انتقالی شامل یک بلوک انتقالی است که وظیفه پردازش و تبدیل ورودی‌ها را دارد. همان‌طور که در شکل ۴ آمده است ابتدا ورودی‌ها با استفاده از لایه‌های نرمال‌سازی و توجه چندسری، تبدیل می‌شوند. سپس با استفاده از لایه‌های Dropout، اطلاعات تبدیل‌شده نگه‌داشته می‌شوند. درنهایت، خروجی تبدیل‌شده با ورودی اصلی جمع و از لایه نرمال‌سازی عبور می‌کند. در ادامه یک‌لایه Conv1D با استفاده از یک مجموعه از فیلترهای کرنل[17] به طول ورودی، یک عمل پیچش را اعمال می‌کند. هر فیلتر کرنل، الگوهای خاصی را در داده‌های ورودی تشخیص می‌دهد. با اعمال عمل پیچش با این فیلترها، ویژگی‌های مختلف در داده‌های ورودی استخراج می‌شوند. سپس از یک‌لایه LSTM با تعداد واحدهای مشخص‌شده، استفاده می‌شود. این لایه اطلاعات را به طول توالی ورودی بازگردانده و به لایه بعدی منتقل می‌کند. پس از لایه LSTM، از لایه تجمع میانگین کلی استفاده می‌شود تا اطلاعات سری زمانی به یک بردار با ابعاد (mlp_units) تبدیل شوند. این لایه با استفاده از میانگین‌گیری در طول توالی، اطلاعات مهم را استخراج می‌کنند. در ادامه، چندین لایه Dense با تعداد واحدهای مشخص‌شده (mlp_units) استفاده می‌شود. این لایه‌ها با استفاده از تابع فعال‌سازی ReLU، اطلاعات را تبدیل و منتقل می‌کنند. در پایان، یک‌لایه Dropout برای پیشگیری از برازش بیش‌ازحد و یک‌لایه Dense با تعداد واحدهای (f_step) و فعال‌سازی سیگموید برای تولید خروجی‌ها استفاده می‌شود. شکل ۳ معماری روش پیشنهادی را نشان می‌دهد.

۳-۱- مرحله آماده‌سازی داده

در مرحله اول که پیش‌پردازش است با توجه به شکل۴ ابتدا داده‌ها از ورودی خوانده‌شده و پیش‌پردازش می‌شوند. در اینجا، داده‌هایی که شامل داده‌های غیرعددی هستند حذف و داده‌ها به نوع عددی تبدیل می‌شوند. در ادامه، پنجره لغزان[18] که یک روش برای تقسیم دنباله داده‌ها به نمونه‌های قابل‌استفاده در مدل‌های پیش‌بینی یا آموزش است، اعمال می‌شود. در این روش پنجره‌ای از روی داده‌ها عبور کرده و در هر مرحله یک نمونه جدید برداشت می‌شود. با استفاده از پنجره لغزان، مدل می‌تواند الگوها و روابط زمانی بین نمونه‌ها را در داده‌ها یاد بگیرد و پیش‌بینی‌هایی برای نمونه‌های آینده ارائه دهد. این روش برای مسائل پیش‌بینی توالی مفید است. روند تقسیم داده در پنجره لغزان، به این صورت است که ابتدا یک پنجره با طول n_steps از ابتدای داده انتخاب می‌شود. بخش ورودی این پنجره به‌عنوان ورودی شبکه عصبی استفاده می‌شود. به تعداد f_step گام زمانی از توالی داده به‌عنوان خروجی شبکه عصبی در نظر گرفته می‌شود. این پنجره به‌طور مرحله‌به‌مرحله به سمت راست حرکت کرده و فرآیند تکرار می‌شود تا به انتهای داده برسد. درنهایت، ماتریس x شامل بخش‌های ورودی و ماتریس y شامل مقادیر خروجی متناظر با هر بخش ورودی ایجاد می‌شوند و برای استفاده در مدل آماده می‌شوند. سپس داده‌ها نرمال‌سازی می‌شوند.



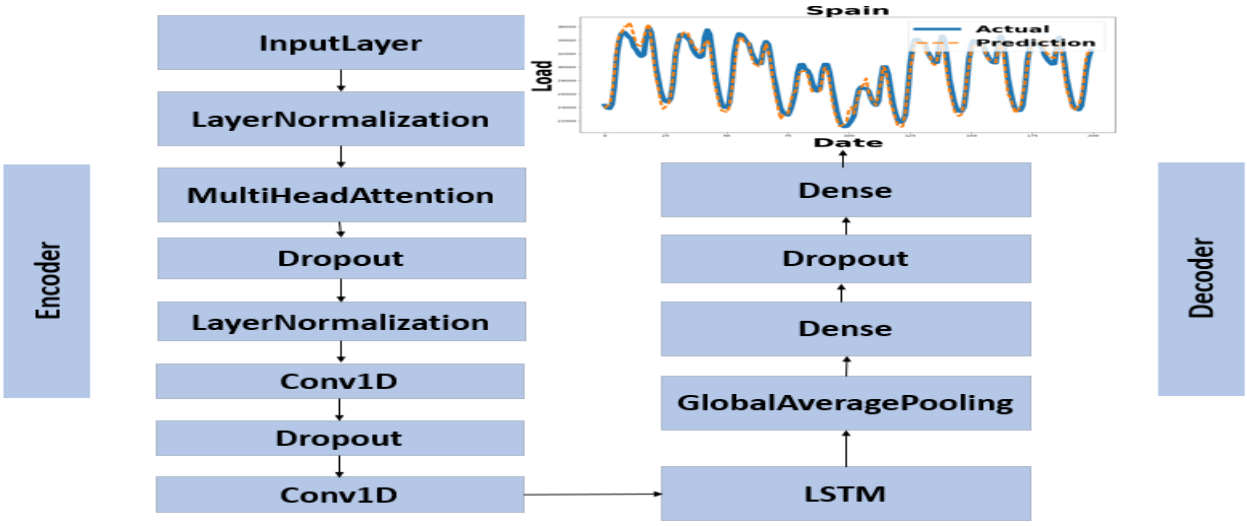
شکل۴- مرحله آماده سازی داده ها در روش پیشنهادی.

۳-۲- مرحله توجه چندسر

با توجه به شکل۵ که در زیر نمایش داده‌شده ورودی به لایه چند سر توجه وارد می‌شود. این لایه برای محاسبه وزن‌های، از مکانیزم توجه بین ورودی و خروجی استفاده می‌کند. با استفاده از این وزن‌ها، امکان تمرکز بر بخش‌های مهم ورودی و جلوگیری از اطلاعات غیرضروری فراهم می‌شود. سپس برای جلوگیری از بیش‌برازش از لایه Dropout استفاده می‌شود. این لایه به‌طور تصادفی، برخی از واحدهای خروجی را به صفر

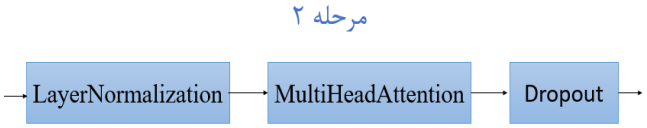
[16] - Multilayer perceptron(mlp)
[17] - kernel filters
[18] - Sliding window

[Downloaded from ieijqp.ir on 2026-06-13]
[DOI: 10.61186/ieijqp.13.2.2]



شکل ۳- چارچوب و معماری روش پیشنهادی

تنظیم می‌کند که منجر به حذف برخی ارتباطات بین واحدها می‌شود. این کار باعث افزایش تنوع و انعطاف‌پذیری مدل می‌شود.

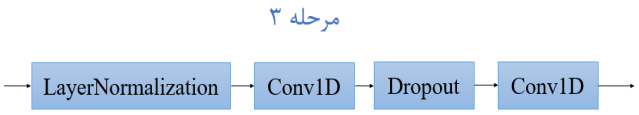


شکل۵- مرحله توجه در روش پیشنهادی.

۳-۳- مرحله استخراج ویژگی

با توجه به شکل۶ دوباره فرآیند نرمال‌سازی انجام می‌گیرد. سپس برای استخراج ویژگی‌ها از لایه Conv1D بر روی دنباله‌های زمانی استفاده می‌شود. CNN یک نوع شبکه عصبی پیش‌خور است که عملکرد خوبی در پردازش تصویر و پردازش زبان طبیعی (NLP) دارد. CNN می‌تواند به‌طور مؤثری در پیش‌بینی سری‌های زمانی به کار گرفته شود. ادراک محلی و به اشتراک‌گذاری وزن در CNN می‌تواند تعداد پارامترها را به‌طور قابل‌توجهی کاهش دهد و بنابراین کارایی مدل‌های یادگیری را بهبود بخشد (Lu et al., 2021).

Conv1D (کانولوشن یک‌بعدی) یک عملگر پردازش است که در شبکه‌های عصبی عمیق به‌ویژه در حوزه تشخیص الگو و پردازش سیگنال استفاده می‌شود. این عملگر برای پردازش داده‌های یک‌بعدی مانند سری‌های زمانی، سیگنال‌های صوتی و سیگنال‌های زمان-فرکانس استفاده می‌شود. برخلاف Conv2D که برای تصاویر دوبعدی استفاده می‌شود، Conv1D تنها یک بعد داده را در نظر می‌گیرد. این لایه با استفاده از فیلترهای پیچشی و عملیات پیچشی یک‌بعدی، اطلاعات مهم را از داده‌ها استخراج می‌کند. سپس دوباره لایه Dropout و لایه پیچشی یک‌بعدی تکرار می‌شود.



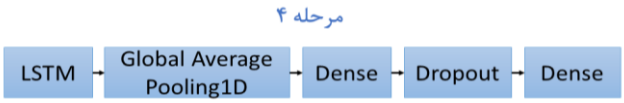
شکل۶- مرحله استخراج ویژگی در روش پیشنهادی.

۳-۴- مرحله شبکه حافظه طولانی کوتاه‌مدت

در شکل ۷ که گام آخر مدل پیشنهادی است، از لایه LSTM که یک نوع خاص از شبکه‌های بازگشتی و برای پردازش دنباله‌های زمانی است استفاده می‌شود. این لایه قادر است به‌طور خاص بازدهی بهتری در مواردی که وابستگی‌های طولانی‌مدت در دنباله وجود دارد، ارائه دهد. این لایه توانایی حفظ و فراموشی اطلاعات را دارد و به شبکه عصبی اجازه می‌دهد از اطلاعات گذشته برای پیش‌بینی از برخی ویژگی‌ها استفاده کند. در مرحله بعد از لایه تجمع میانگین کلی یک‌بعدی استفاده می‌شود که این لایه به‌منظور کاهش ابعاد خروجی استفاده می‌شود. با اعمال این لایه، میانگین مقادیر در تمام بعدهای داده محاسبه می‌شود و به‌عنوان خروجی تولید می‌شود. این کار باعث کاهش تعداد پارامترها و ابعاد داده‌ها می‌شود و درنتیجه مدل را سبک‌تر می‌کند. در آخر از لایه Dense استفاده می‌شود. این لایه معمولاً در پایان مدل برای تبدیل بردارهای ویژگی به خروجی نهایی استفاده می‌شود.

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]



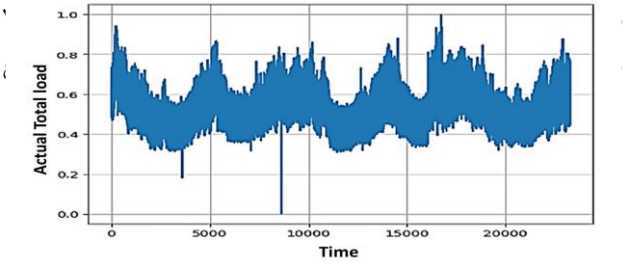
شکل ۷- مرحله لایه بازگشتی در روش پیشنهادی.

۴_ ارزیابی راهکار پیشنهادی

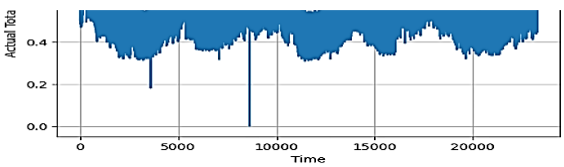
اغلب روش‌ها برای پیش‌بینی مصرف برق در شرایط ایده آل به‌خوبی عمل می‌کنند اما به دلیل برخی چالش‌های اساسی دقت بالایی ندارند. تک متغیره و چند متغیره بودن مدل از سناریوهایی هستند که می‌توانند روی دقت مدل تأثیر بگذارند. در نظر گرفتن عوامل محیطی تأثیرگذار بر روی پارامتر هدف می‌تواند پیش‌بینی مدل را بهبود بخشد. برای مسئله پیش‌بینی مصرف برق در نظر گرفتن عواملی همچون دما، فشار، تعطیلات، و شرایط اقتصادی، به پیش‌بینی با دقت بهتر مدل کمک می‌کند. همان‌طور که در مدل پیشنهادی بیان شد، وجود ویژگی‌هایی همچون دما و فشار در ورودی باعث عملکرد بهتر روش پیشنهادی می‌شود. در ادامه داده استفاده‌شده در این مطالعه و معیار اعتبار سنجی روش پیشنهادی و روش‌های مقایسه شده تشریح می‌شوند. در انتها سناریوهای طراحی‌شده در پیش‌بینی مصرف برق مورد تحلیل و ارزیابی قرار خواهند گرفت.

۴-۱- معرفی مجموعه داده انتخابی

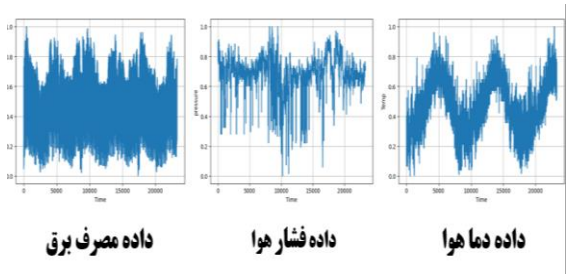
در این پژوهش از داده‌های بازار انرژی اسپانیا و مونته‌نگرو استفاده شده است. داده‌های اسپانیا شامل داده‌های ساعتی مصرف برق و داده‌های آب‌وهوایی از سال ۲۰۱۷ تا ۲۰۱۹ هستند. داده‌های مصرف برق از ENTSOE Transparency Platform و داده‌های آب‌وهوایی از OpenWeatherApi به دست آمده‌اند . در مجموع ۲۳۲۷۵ داده مصرف برق و داده‌های مربوط به دما و فشار جمع‌آوری شده‌اند. داده‌های



شکل ۸- مجموعه داده مونته نگرو

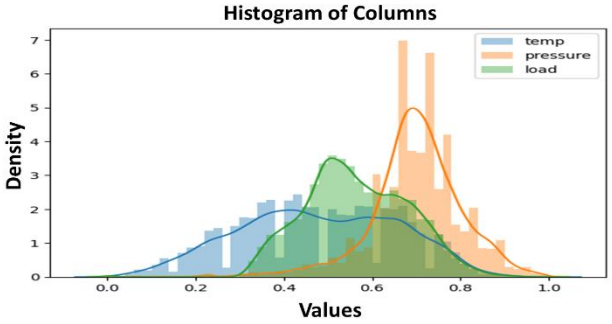


شکل ۸- مجموعه داده مونته نگرو



شکل ۹- مجموعه داده بازار انرژی اسپانیا.

داده‌های ورودی به دو بخش آموزش و آزمون تقسیم می‌شوند. بخش آموزش ۸۰٪ از داده‌ها را شامل می‌شود و بخش آزمون ۲۰٪ باقی‌مانده است. در شکل ۱۰ هیستوگرام اطلاعات توزیع یک مجموعه داده را به‌صورت گرافیکی مشاهده می‌کنید. هدف اصلی استفاده از هیستوگرام، شناخت توزیع یک مجموعه داده است. در این مثال، دما و فشار ویژگی‌ها یا متغیرهای مستقل هستند که بر روی محور x نمایش داده می‌شوند. مصرف برق به‌عنوان متغیر وابسته یا هدف در نظر گرفته‌شده است که بر روی محور y قرار می‌گیرد. هیستوگرام نشان می‌دهد که چگونه تغییر در دما و فشار ممکن است بر مصرف برق تأثیر بگذارد. به‌عنوان مثال، ممکن است هنگامی که دما بالاتر است، مصرف برق بیشتر باشد زیرا احتمالاً از تهویه مطبوع و تجهیزات خنک‌کننده بیشتر استفاده می‌شود.



شکل ۱۰- هیستوگرام مجموعه داده.

۴-۲- معیارهای اعتبار سنجی

معیارهای ارزیابی، دقت و عملکرد مدل را مورد بررسی و ارزیابی قرار می‌دهند. چهار معیار در این مطالعه برای اعتبارسنجی و ارزیابی روش پیشنهادی مورد استفاده قرار گرفته است. هر یک از این معیارها دارای مزایا و معایبی هستند. RMSE تابع زیان مرتبه دوم را ارائه می‌دهد که بر خطاهای بزرگتر تأکید بیشتری دارد. MAE یکنواخت و طبیعی بوده و می‌تواند میانگین خطا را اندازه‌گیری کند. MAPE به مقیاس بستگی ندارد و می‌تواند به راحتی برای داده پرحجم و کم‌حجم به کار رود (Al Mamun et al., 2020).

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]

ریشه میانگین مربعات خطا یک معیار شاخص است که برای اندازه‌گیری میزان تفاوت بین مقادیر پیش‌بینی شده و مقادیر واقعی استفاده می‌شود. این معیار معمولاً در مسائل پیش‌بینی و رگرسیون استفاده می‌شود و اندازه میانگین مربعات خطا را نشان می‌دهد.. رابطه محاسبه RMSE به‌صورت رابطه ۱ است:

(۱)

$$RMSE = \frac{\sqrt{\sum_{i=1}^{n} \|y(i) - \hat{y}(i)\|^2}}{\sqrt{n}}$$

که در آن N تعداد نقاط داده، y(i) i-مقدار اندازه‌گیری و ŷ(i) پیش‌بینی متناظر آن است. برای محاسبه RMSE، ابتدا تفاضل بین هر مقدار واقعی (x) و مقدار پیش‌بینی شده متناظر آن (y) در مجموعه داده آزمون را محاسبه می‌کنیم. سپس مربع این تفاضل‌ها را می‌گیریم. پس از آن، میانگین این مربعات را محاسبه می‌کنیم و در نهایت ریشه دوم این میانگین را می‌گیریم تا RMSE به دست آید.

معیار دوم معیار میانگین خطای مطلق یا MAE به صورت رابطه ۲ است:

(۲)

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

y_i = مقدار پیش‌بینی و x_i = مقدار واقعی و n = تعداد کل نقاط داده را نشان می‌دهد. برای محاسبه MAE، قدرمطلق تفاضل بین هر مقدار واقعی (x) و مقدار پیش‌بینی شده متناظر آن (y) در مجموعه داده آزمون را محاسبه می‌کنیم. سپس میانگین این قدرمطلق‌ها را می‌گیریم که همان MAE است.

معیار سوم معیار دقت است که این معیار با استفاده از رابطه ۳ محاسبه می‌شود:

$$Accuracy = \frac{True\ Prediction \times 100}{Total\ samples} \quad (۳)$$

True Predictions تعداد نمونه‌هایی که به‌درستی توسط مدل پیش‌بینی شده‌اند. Total Samples کل تعداد نمونه‌ها در مجموعه‌ی ارزیابی است. این معیار نشان می‌دهد که چه درصد از نمونه‌ها به‌درستی توسط مدل پیش‌بینی شده‌اند. برای محاسبه دقت، ابتدا تعداد نمونه‌هایی را که مدل در مجموعه داده آزمون به درستی پیش‌بینی کرده است (True Predictions) را شمارش می‌کنیم. سپس این تعداد را بر تعداد کل نمونه‌ها در مجموعه داده آزمون (Total Samples) تقسیم می‌کنیم تا دقت به دست آید.

معیار آخر میانگین درصد خطای مطلق (MAPE[19]) است که برای اندازه‌گیری میزان خطای مدل در پیش‌بینی مقادیر عددی استفاده می‌شود. فرمول محاسبه MAPE به‌صورت رابطه ۴ است:

$$MAPE = \frac{1 \times \sum_{t=1}^{n} \frac{|A_t - F_t|}{|A_t|}}{n} \quad (۴)$$

n = تعداد دفعاتی که تکرار اتفاق می‌افتد و At = ارزش واقعی و Ft = مقدار پیش‌بینی است. برای محاسبه MAPE، ابتدا قدرمطلق تفاضل بین هر مقدار واقعی (At) و مقدار پیش‌بینی شده متناظر آن (Ft) در مجموعه داده آزمون را محاسبه می‌کنیم. سپس این قدرمطلق را بر مقدار واقعی (At) تقسیم می‌کنیم تا درصد خطای مطلق برای آن نمونه به دست آید. این کار را برای همه نمونه‌ها انجام داده و میانگین این درصدهای خطا را می‌گیریم که همان MAPE است.

۴-۳-روش‌های قابل‌مقایسه و تنظیم پارامترها

به‌منظور مقایسه روش پیشنهادی، از داده‌های بازار انرژی اسپانیا و مونته‌نگرو استفاده شده است. روش پیشنهادی با دو روش شبکه عصبی حافظه طولانی کوتاه‌مدت و مدل انتقالی بر روی هر دو مجموعه داده مقایسه شده است . تمامی مدل‌ها از تنظیمات مشترکی بهره برده‌اند که شامل تعداد دوره‌های آموزش برابر ۳۰، اندازه دسته برابر ۱۰، تقسیم داده‌ها به نسبت ۸۰-۲۰ برای آموزش و آزمون، تعداد گام‌های زمانی برابر ۹۶، تعداد گام‌های آینده‌نگر برابر ۲ است. این تنظیمات مشترک اعمال شده‌اند تا مقایسه‌ای منصفانه بین مدل‌ها انجام شود. در مدل انتقالی تعداد سرها برابر ۴ و در روش پیشنهادی برابر ۲ است. همچنین نرخ یادگیری مدل انتقالی ۰.۰۰۰۱ و برای روش پیشنهادی ۰.۰۰۱ می‌باشد. با استفاده از این تنظیمات، نتایج و عملکرد هر مدل قابل مقایسه است. مدل با استفاده از الگوریتم بهینه‌سازی Adam و توابع هزینه ریشه میانگین مربعات خطا، میانگین خطای مطلق و میانگین درصد خطای مطلق آموزش داده می‌شود. همچنین برای پایش عملکرد مدل، از معیار دقت استفاده می‌شود. با به‌کارگیری الگوریتم بهینه‌سازی و توابع هزینه مختلف، آموزش و ارزیابی مدل به‌طور مؤثر انجام می‌شود.

۴-۴-نتایج و ارزیابی

در این بخش، پیش‌بینی مصرف برق در دو سناریوی مجزا مورد بررسی قرار گرفته است. در سناریوی اول، مسئله به صورت تک‌متغیره و با

[19] - Mean Absolute Percentage Error

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]

استفاده از دو مجموعه داده مصرفی مونته‌نگرو و اسپانیا و با در نظر گرفتن تنها میزان مصرف، مدل‌سازی شده است. در سناریوی دوم، از دو مجموعه داده اسپانیا استفاده شده و اثر متغیرهای دما و فشار نیز لحاظ گردیده است. در این سناریو مدل‌سازی چندمتغیره انجام شده است. سپس نتایج حاصل از هر دو سناریو ارائه و تحلیل‌های لازم بر روی آن‌ها صورت گرفته است. روش پیشنهادی با دو شبکه عصبی حافظه طولانی کوتاه‌مدت و مدل انتقالی مقایسه شده است. برای ارزیابی از معیارهای دقت، ریشه میانگین مربعات خطا، میانگین خطای مطلق و میانگین درصد خطای مطلق استفاده شده است.

۴-۴-۱- سناریو تک متغیره

در سناریوی تک متغیره فرض بر این است که فقط یک متغیر و آن هم مصرف برق در نظر گرفته می‌شود. این سری زمانی به‌عنوان ورودی به مدل برای پیش‌بینی داده می‌شود. شبکه با تحلیل الگوهای موجود در این داده‌های تاریخی، یک مدل پیش‌بینی ایجاد می‌کند و می‌تواند مقدار مصرف برق را برای آینده پیش‌بینی کند.

جدول ۲و۳ نتایج مقایسه‌ای چند مدل تک متغیره برای پیش‌بینی مصرف برق را نشان می‌دهد. با توجه به نتایج ، روش پیشنهادی یعنی مدل ترکیبی تک متغیره در هر ۴ معیار، بهترین عملکرد را داشته است. RMSE و MAE به مراتب کمتر از دو مدل دیگر است. صحت وMAPE نیز بهتر از بقیه مدل‌هاست. بنابراین می‌توان نتیجه گرفت که در این مقایسه، مدل Transformer-LSTM عملکرد بهتری نسبت به مدل‌های LSTM و انتقالی داشته است.

جدول۲- ارزیابی روش ها در سناریوی تک متغیره بر روی داده اسپانیا.

نام روش	RMSE	دقت	MSE	MAPE
مدل انتقالی	۲۷۲.۲۰	۹۸.۵۱	۴۰۸	۱.۴۸
شبکه عصبی حافظه طولانی کوتاه مدت	۵۱۲.۲۹	۹۸.۶۲	۳۷۰	۱.۳۷
روش ترکیبی پیشنهادی	۶۶.۳۴	۹۹.۱۲	۲۳۹	۰.۸۷

جدول۳- ارزیابی روش ها در سناریو تک متغیره بر روی داده مونته نگرو.

نام روش	RMSE	دقت	MSE	MAPE
مدل انتقالی	۱۲.۹۸	۹۷.۴۴	۱۰	۲.۵۵
شبکه عصبی حافظه طولانی کوتاه مدت	۱۲.۰۴	۹۷.۶۳	۹	۲.۳۶
روش ترکیبی پیشنهادی	۱۰.۲۸	۹۸.۰۹	۷	۱.۹

۴-۴-۲- سناریو چند متغیره

در سناریوی چند متغیره برای پیش‌بینی مصرف برق، از چندین متغیر ورودی استفاده می‌شود. علاوه بر سری زمانی تاریخی مصرف برق، متغیرهایی مانند تاریخ، ساعت روز، دما، رطوبت هوا و رویدادهای خاص مثل تعطیلات نیز می‌توانند به‌عنوان ورودی در نظر گرفته شوند. ازجمله مهم‌ترین این متغیرها، دما است که تأثیر زیادی بر میزان و الگوی مصرف برق دارد. به‌طور کلی در روزهای گرم‌تر مصرف برق بیشتر است. همچنین تغییرات فشار جو نیز می‌تواند بر مصرف تأثیرگذار باشد. افزودن این متغیرهای تأثیرگذار به‌عنوان ورودی، به مدل اجازه می‌دهد تا الگوهای فصلی و تأثیر دما و فشار جو را در پیش‌بینی‌های خود لحاظ کند. درنتیجه استفاده از سناریو چندمتغیره می‌تواند باعث افزایش دقت و بهبود عملکرد مدل شود.

استفاده از سناریو تک متغیره یا چند متغیره بستگی به مسئله خاص و داده‌های موجود دارد. درصورتی‌که تنها به پیش‌بینی مصرف برق بدون در نظر گرفتن عوامل دیگر نیاز دارید، سناریو تک متغیره مناسب است. اما اگر عوامل مختلفی بر مصرف برق تأثیر دارند و لازم رابطه بین آن‌ها لحاظ شود، سناریو چند متغیره مناسب‌تر خواهد بود. به‌طور کلی، هرچند روش‌های به‌کاربرده شده می‌تواند خطاهای پیش‌بینی را در مدل‌های تک متغیره کاهش دهد، اما خطاها همچنان بیشتر از مدل‌های چندمتغیره است. این نشان‌دهنده اهمیت همبستگی درون دوره‌ای بین متغیرهای مختلف برای پیش‌بینی دقیق‌تر است (Shah et al., 2022). برای بررسی این موضوع، نتایج جدول ۴ نشان می‌دهد که مدل پیشنهادی بر روی مجموعه داده اسپانیا که از چند متغیر ورودی استفاده می‌کنند، در مقایسه با مدل‌های تک‌متغیره‌ای که فقط از یک متغیر ورودی بهره می‌برند، عملکرد بهتری در بیشتر معیارهای ارزیابی داشته‌اند. به‌طور خاص، مدل پیشنهادی ترکیبی این مقاله در سناریو چندمتغیره در مقایسه با سایر روش‌های تک‌متغیره و چندمتغیره، مجموعه داده اسپانیا توانسته است معیارهایی مانند میانگین مربعات خطا، میانگین خطای مطلق، میانگین درصد خطای مطلق و دقت پیش‌بینی را به‌طور قابل‌توجهی بهبود بخشد. این نتایج نشان می‌دهد که استفاده از چندین متغیر مرتبط با مصرف برق مانند دما، فشار در کنار سری زمانی تاریخی، می‌تواند باعث افزایش قدرت پیش‌بینی مدل شود. زیرا مدل می‌تواند از ارتباط بین این متغیرهای مختلف برای پیش‌بینی دقیق‌تر مصرف برق استفاده کند.

ترکیب دو مدل LSTM و Transformer در روش پیشنهادی برای پیش‌بینی مصرف برق، چالش‌هایی مانند نیاز به منابع محاسباتی بیشتر، زمان آموزش طولانی‌تر، حجم زیاد داده‌های آموزشی را به همراه دارد.

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]

اما این چالش ها برای دستیابی به عملکرد بهتر قابل قبول هستند و با مدیریت صحیح منابع و داده‌ها قابل مدیریت می‌باشند. و می توان از مزایای ترکیب دو مدل قدرتمند بهره‌مند شد. در این پژوهش، روش ترکیبی پیشنهادی برای پیش‌بینی مصرف برق با استفاده از زبان پایتون و با منابع محاسباتی محدود شامل پردازنده دوهسته‌ای Xeon 2.2 گیگاهرتزی، ۱۳ گیگابایت RAM و ۱۲ گیگابایت فضای دیسک سخت، پیاده‌سازی شده است. محدودیت زمانی منابع، امکان افزایش طولانی دوره‌های آموزش را محدود کرده که می‌تواند بر دقت و همگرایی مدل تاثیرگذار باشد. . به منظور غلبه بر این محدودیت‌ها و افزایش عملکرد، تکنیک‌های بهینه‌سازی و پیش‌پردازش داده‌ها اعمال شده و آزمایش‌ها با تنظیمات مختلف چندین بار تکرار شده تا از اعتبار نتایج اطمینان حاصل شود.

جدول ۴- ارزیابی روش‌های پیشنهادی ردر سناریو چند متغی بر روی داده اسپانیا.

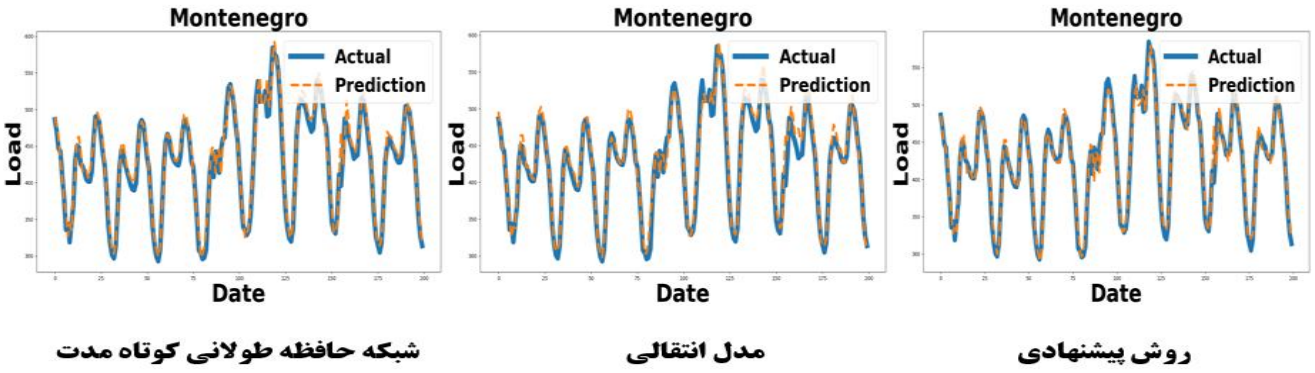
نام روش	RMSE	دقت	MAE	MAPE
مدل انتقالی	۸۴.۳۵	۹۸.۶۷	۳۶۰	۱.۳۲
حافظه طولانی کوتاه مدت	۵۲۵.۶	۹۸.۷۱	۳۶۷	۱.۲۸
ترکیبی پیشنهادی	۱۵.۹۳	۹۹.۲۱	۲۱۵	۰.۷۸

این مطالعه با ارائه روش ترکیبی جدید مبتنی بر ترکیب شبکه های عصبی LSTM و مدل Transformer، به پیش بینی قابل قبولی دست یافته است. این دقت باعث بهبود برنامه ریزی تولید انرژی، افزایش بهره وری شرکت های برق، کاهش هدر رفت انرژی و هزینه ها، ارتقای قابلیت اطمینان شبکه توزیع، بهینه سازی الگوی مصرف در سطح مشترکین و تسهیل گذار به انرژی های تجدیدپذیر می شود. در نهایت، منجر به بهبود کارایی و پایداری کل شبکه برق رسانی خواهد شد.

۵-نتیجه‌گیری

در زمینه برنامه‌ریزی تولید برق، پیش‌بینی مصرف برق به شرکت‌ها امکان می‌دهد تا تولید برق را با تقاضای واقعی هماهنگ کنند و از هدر رفت انرژی و افت کیفیت خدمات جلوگیری کنند. همچنین، در سیستم‌های توزیع برق، داشتن پیش‌بینی صحیح مصرف برق به شرکت‌های توزیع امکان می‌دهد تا ظرفیت و منابع خود را بهینه کنند و درنتیجه هزینه‌ها را کاهش دهند. طراحی یک شبکه عصبی باقابلیت پیش‌بینی صحیح بسیار حائز اهمیت است. این مقاله شبکه عصبی عمیق با حافظه طولانی کوتاه مدت را به جهت مدلسازی روابط زمانی در نظر گرفته است و آن‌را به واسطه مدل انتقالی با لایه‌های توجه تقویت کرده است. بدین ترتیب ضمن توجه به ارتباطات سراسری، الگوهای پیچیده‌تر را شناسایی می‌کند. این شبکه عصبی توانسته است عملکرد بهتری نسبت به شبکه‌های عصبی با حافظه طولانی کوتاه مدت بدست بیاورد.

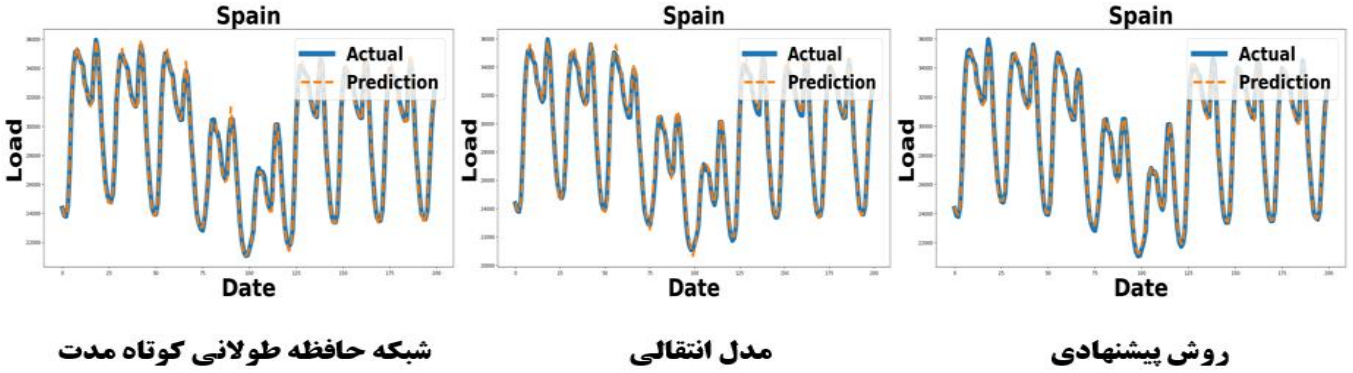
راهکار این مقاله قطعاً قابل تعمیم به داده‌های داخلی است، اما در این پژوهش از داده‌های بازار انرژی اسپانیا و مونته‌نگرو استفاده شده است زیرا این داده‌ها از کیفیت بالاتری برخوردار بوده و می‌توانند به ارزیابی بهتر روش پیشنهادی کمک کنند. با این حال، با توجه به محرمانه بودن داده‌های مصرف برق کشور، دسترسی به این داده‌ها برای استفاده در پژوهش‌های علمی با محدودیت‌هایی همراه است. علاوه بر این، آلودگی و نویز زیاد در داده‌های مصرف برق داخلی ایران، نیاز به پاکسازی و ارتقای کیفیت داده به کمک کارشناسان برق را ایجاب می‌کند. در تحقیقات آتی، با دسترسی به داده‌های با کیفیت داخلی و پیش‌پردازش مناسب آن‌ها، می‌توان روش پیشنهادی را بر روی داده‌های واقعی ایران پیاده‌سازی و ارزیابی کرد تا بتوان از مزایای آن در برنامه‌ریزی تولید و توزیع برق در کشور بهره‌مند شد.



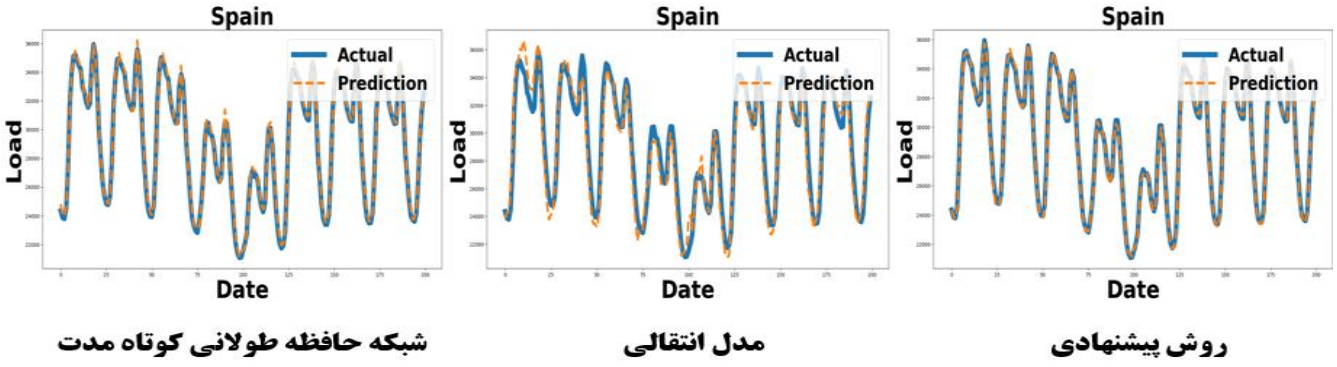
شکل ۱۱- نتایج پیش بینی روش‌های مورد ارزیابی در سناریو تک متغیره بر روی داده مونته نگرو .

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]



شکل ۱۲- نتایج پیش بینی روش‌های مورد ارزیابی در سناریو تک متغیره بر روی داده اسپانیا.



شکل ۱۳- نتایج پیش بینی روش‌های مورد ارزیابی در سناریو چند متغیره بر روی داده اسپانیا.

۶- مراجع

A. Xu, M.-W. T., B. Firouzi, K. A. Alattas, A. Mohammadzadeh, and E. Ghaderpour, "A new deep learning Restricted Boltzmann Machine for energy consumption forecasting," Sustainability, vol. 14, no. 16, p. 10081, 2022.

Al Mamun, A., Sohel, M., Mohammad, N., Sunny, M. S. H., Dipta, D. R., & Hossain, E. (2020). A comprehensive review of the load forecasting techniques using single and hybrid predictive models. *IEEE Access*, *8*, 134911-134939.

Amalou, I., Mouhni, N., & Abdali, A. (2022). Multivariate time series prediction by RNN architectures for energy consumption forecasting. *Energy Reports*, *8*, 1084-1091.

Cui, C., He, M., Di, F., Lu, Y., Dai, Y., & Lv, F. (2020). Research on power load forecasting method based on LSTM model. 2020 IEEE 5th information technology and mechatronics engineering conference (ITOEC),

Farsi, B., Amayri, M., Bouguila, N., & Eicker, U. (2021). On short-term load forecasting using machine learning techniques and a novel parallel deep LSTM-CNN approach. *IEEE Access*, *9*, 31191-31212.

L'Heureux, A., Grolinger, K., & Capretz, M. A. (2022). Transformer-based model for electrical load forecasting. *Energies*, *15*(14), 4993.

Lu, W., Li, J., Wang, J., & Qin, L. (2021). A CNN-BiLSTM-AM method for stock price prediction. *Neural Computing and Applications*, *33*, 4741-4753.

Mahjoub, S., Chrifi-Alaoui, L., Marhic, B., & Delahoche, L. (2022). Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks. *Sensors*, *22*(11), 4062.

Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing*, *452*, 48-62.

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]

Ozcan, A., Catal, C., & Kasif, A. (2021). Energy load forecasting using a dual-stage attention-based recurrent neural network. *Sensors*, *21*(21), 7115.

Shah, I., Iftikhar, H., & Ali, S. (2022). Modeling and forecasting electricity demand and prices: A comparison of alternative approaches. *Journal of Mathematics*, *2022*.

Song, X., Liu, Y., Xue, L., Wang, J., Zhang, J., Wang, J., Jiang, L., & Cheng, Z. (2020). Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. *Journal of Petroleum Science and Engineering*, *186*, 106682.

Tarmanini, C., Sarma, N., Gezegin, C., & Ozgonenel, O. (2023). Short term load forecasting based on ARIMA and ANN approaches. *Energy Reports*, *9*, 550-557.

Wang, C., Wang, Y., Ding, Z., Zheng, T., Hu, J., & Zhang, K. (2022). A transformer-based method of multienergy load forecasting in integrated energy system. *IEEE Transactions on Smart Grid*, *13*(4), 2703-2714.

Wang, D., Gan, J., Mao, J., Chen, F., & Yu, L. (2023). Forecasting power demand in China with a CNN-LSTM model including multimodal information. *Energy*, *263*, 126012.

Wang, H., Zhang, Y., Liang, J., & Liu, L. (2023). DAFA-BiLSTM: Deep Autoregression Feature Augmented Bidirectional LSTM network for time series prediction. *Neural Networks*, *157*, 240-256.

توان, م., حاجیانی, پ., & پارسا, ح. (1396). ارزیابی الگوهای سری زمانی و فازی برای پیش بینی مصرف برق بخش های مختلف ایران تا افق 1410 دومین همایش بین المللی انسجام مدیریت و اقتصاد در توسعه , https://civilica.com/doc/715782

[Downloaded from ieijqp.ir on 2026-06-13]

[DOI: 10.61186/ieijqp.13.2.2]