

A Feature Extraction Based Long-Term Electricity load forecasting Framework to Reduce the Outliers Data Effects**Mohammad Davoud Saeidi^{1,2}, M.Sc., Majid Moazzami^{1,2}, Associate Professor**¹Department of Electrical Engineering- Najafabad Branch, Islamic Azad University, Najafabad, Iran²Smart Microgrid Research Center- Najafabad Branch, Islamic Azad University, Najafabad, Iran
mdsaeidi@sel.iaun.ac.ir, m_moazzami@pel.iaun.ac.ir**Abstract**

Generally, medium-term electrical load forecasting is often used for the operation of thermal and hydropower plants, optimal time planning for maintenance of power plants and the power grids. However, long-term electrical load forecasting is used to manage on-time future demands and generation, transmission and distribution expansion planning. In this paper, a hybrid long-term load forecasting approach using wavelet transform and an outlier robust extreme learning machine is proposed. Hourly load and temperature data were extracted from the GEFCOM 2014 database and divided into two classes of training and test. The one-level wavelet transform is used to decompose data to extract properties and reduce the dimensions of the data matrix. Decomposed low-frequency component (approximations) and high-frequency component values (details) from wavelet analysis are entered into the model for training and forecasting. For comparison accuracy of the proposed method, wavelet transform is applied to the data for the other three extreme learning machines. Also data without wavelet transform entered into four other forecasting models and the load forecasting results are compared with the proposed method. The results of the above mentioned evaluation show that electrical load forecasting by using wavelet transform and outlier robust extreme learning machine improves forecasting accuracy and the MAPE reduces to 3.0966. The overall calculated error by the proposed method was the best result obtained between the three several models of extreme learning machines and without preprocessing model. The MAPE is 0.4208 less than the ELM, 0.944 less than the RELM, and 0.1353 less than the WRELM model, respectively.

Keywords: extreme learning machine, improve forecast accuracy, preprocessing, long-term load forecasting, mean absolute percentage error, wavelet transform,

Received: 19 November 2021**Revised:** 23 December 2021**Accepted:** 7 March 2022**Corresponding Author:** Dr. Majid Moazzami

Citation: M.D. Saeidi, M. Moazzami, "A feature extraction based long-term electricity load forecasting framework to reduce the outliers data effects", Journal of Intelligent Procedures in Electrical Technology, vol. 14, no. 56, pp. 1-20, January 2023 (in Persian).

یک روش پیش‌بینی بلندمدت بار الکتریکی مبتنی بر استخراج ویژگی برای کاهش اثر داده‌های خارج از محدوده

محمد داود سعیدی^{۱،۲}، دانش‌آموخته کارشناسی ارشد، مجید معظمی^{۱،۲}، دانشیار

۱- دانشکده مهندسی برق- واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران

۲- مرکز تحقیقات ریزشبکه‌های هوشمند- واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران

mmdsaedi@sel.iaun.ac.ir, m_moazzami@pel.iaun.ac.ir

چکیده: پیش‌بینی میان‌مدت بار الکتریکی اغلب برای برنامه‌ریزی عملیات نیروگاه‌های حرارتی و آبی، زمان‌بندی بهینه برای بازرسی و تعمیرات و نگهداری نیروگاه‌ها و شبکه برق استفاده می‌شود. در این مقاله یک روش ترکیبی با استفاده از تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده، برای پیش‌بینی بلندمدت بار ارائه شده است. داده‌های بار و دمای ساعتی، از پایگاه داده GEFCOM 2014 استخراج شده و به دو دسته آموزش و آزمایش تقسیم شده است. از تبدیل موجک یک سطحی برای تجزیه داده‌ها به منظور استخراج ویژگی‌ها و کاهش ابعاد ماتریس داده‌ها استفاده می‌شود. دو دسته مقادیر مؤلفه‌های فرکانس پایین (تقریب) و مقادیر مؤلفه‌های فرکانس بالا (جزئیات) حاصل از تجزیه جهت آموزش و پیش‌بینی به مدل وارد شده و خروجی مقادیر پایین با خروجی مقادیر بالای مدل جمع می‌شود تا پیش‌بینی نهایی را تشکیل دهد. جهت سنجش و مقایسه دقت و کارایی روش پیشنهادی، اعمال تبدیل موجک روی داده‌ها، برای سه مدل دیگر ماشین یادگیری شدید انجام گردیده است. همچنین داده‌ها بدون اعمال تبدیل موجک به چهار مدل پیش‌بینی دیگر نیز وارد شده و نتایج پیش‌بینی حاصل با روش پیشنهادی مورد مقایسه قرار گرفته است. نتایج ارزیابی فوق نشان می‌دهد که تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج محدوده باعث بهبود دقت پیش‌بینی می‌گردد و مقدار میانگین درصد خطای مطلق به عدد ۳/۰۹۶۶ کاهش یافته است. مقدار خطای کلی محاسبه شده روش پیشنهادی بهترین نتیجه در بین سایر مدل‌های ماشین یادگیری شدید و روش‌های بدون پیش‌پردازش بوده است. خطای فوق بر مبنای مقدار میانگین درصد خطای مطلق به ترتیب ۰/۴۲۰۸ نسبت به مدل ماشین یادگیری شدید اصلی، ۰/۱۱۹۴ نسبت به مدل تنظیم‌شده و ۰/۱۳۵۳ نسبت به مدل تنظیم‌شده و وزن‌دار، کاهش یافته است.

کلمات کلیدی: بهبود دقت پیش‌بینی، پیش‌پردازش، پیش‌بینی بلندمدت بار، تبدیل موجک، ماشین یادگیری شدید، میانگین درصد خطای مطلق

تاریخ ارسال مقاله: ۱۴۰۰/۸/۲۸

تاریخ بازنگری مقاله: ۱۴۰۰/۱۰/۲

تاریخ پذیرش مقاله: ۱۴۰۰/۱۲/۱۶

نام نویسنده‌ی مسئول: دکتر مجید معظمی

نشانی نویسنده‌ی مسئول: نجف‌آباد- بلوار دانشگاه- دانشگاه آزاد اسلامی واحد نجف‌آباد- دانشکده مهندسی برق

۱- مقدمه

مصرف بار الکتریکی به صورت پیچیده و غیرخطی تابعی از پارامترهای متعددی مانند شرایط آب و هوایی (درجه حرارت، مقدار رطوبت، روشنایی هوا، سرعت باد) است. هر روز هفته منحنی بار خاص خود را دارد. منحنی‌های مصرف بار در روزهای تعطیل و غیر تعطیل نیز از یکدیگر متمایز هستند. در فصول مختلف سال نیز با توجه به عوامل مختص هر فصل نظیر طول روز، منحنی مصرف بار تغییر می‌کند [۱].

برای برنامه‌ریزی و بهره‌برداری از سیستم‌های قدرت دو دسته متمایز از پیش‌بینی بار و بر مبنای مدت‌زمان پیش‌بینی وجود دارد. پیش‌بینی بار برای یک بازه زمانی چندماهه تا یک‌ساله میان‌مدت بوده و این نوع پیش‌بینی عمدتاً برای زمان‌بندی سوخت استفاده می‌شود. مدت زمان طولانی‌تر در پیش‌بینی بار می‌تواند از یک تا ده سال باشد (بلند-مدت) که برای تعیین موقعیت اقتصادی، نوع و اندازه نیروگاه‌های آینده استفاده می‌شود. در عملیات بهره‌برداری سیستم‌های قدرت همچنین دودسته اصلی پیش‌بینی بار کوتاه-مدت شامل پیش‌بینی بار بسیار کوتاه و کوتاه-مدت می‌تواند انجام شود. پیش‌بینی بار بسیار کوتاه-مدت برای چند دقیقه پیش رو است و برای کنترل خودکار تولید استفاده می‌شود. پیش‌بینی بار کوتاه-مدت برای یک تا ۱۶۸ ساعت آینده است. نتایج پیش‌بینی بار کوتاه-مدت عمدتاً برای اهداف برنامه‌ریزی تولید استفاده می‌شود [۲]. ارائه مدل‌های جدید با استفاده از ابزارهای هوش محاسباتی و محاسبات نرم و ترکیب مدل‌ها از جمله دقیق‌ترین و پرکاربردترین روش‌های حال حاضر به‌منظور مدل‌سازی پیچیدگی و عدم قطعیت موجود در داده‌ها هستند. هر الگو یا تکنیک، مزایا و معایبی دارد که بعضاً به‌تنهایی نتیجه مطلوب و دقت موردنیاز را نمی‌تواند داشته باشد، بنابراین، بسیاری از محققین و پژوهشگران برای حصول نتایج و کارایی بهتر، ترکیبی بهینه از مدل‌ها را استفاده می‌کنند.

۱-۱- بررسی مطالعات پیش‌بینی بلندمدت بار

در روش‌های ترکیبی اخیر، شبکه‌های عصبی یکی از اجزای اساسی بوده است و سه فاکتور زمان، شرایط آب و هوایی و شرایط اقتصادی مانند هزینه برق، عوامل مؤثر در پیش‌بینی بار هستند [۳]. در مرجع [۴] نمونه‌ای از پیش‌بینی بار، مبتنی بر مدل‌های فصلی میانگین متحرک خود همبسته^۱ و اندازه‌گیری‌های بار دوره‌ای جمع‌آوری شده توسط کنتورهای هوشمند، برای برنامه‌ریزی بهینه عملیاتی^۲ در سیستم تحت نظارت توزیع برق، ارائه شده است. برای پیش‌بینی بار از روش‌های یادگیری عمیق نیز استفاده می‌شود. در یک تعریف کلی، یادگیری عمیق، همان یادگیری ماشین است، به‌طوری‌که در سطوح مختلف نمایش یا تفکیک یادگیری را برای ماشین انجام می‌دهد. با این کار، ماشین درک بهتری از واقعیت وجودی داده‌ها پیدا کرده و می‌تواند الگوهای مختلف را شناسایی کند. در مرجع [۲] دو مدل پیش‌بینی شبکه عصبی کانولوشن وابسته به زمان^۳ و شبکه عصبی حافظه طولانی کوتاه-مدت مبتنی بر سیکل (چرخه)^۴ برای پیش‌بینی بار کوتاه-مدت و میان-مدت فقط با استفاده از بار الکتریکی دوره‌ای پیشنهاد شده است. در مرجع [۵] از یک الگوریتم ژنتیک تطبیقی بهبود یافته برای بهینه‌سازی ورودی مدل و جهت پیش‌بینی تقاضای برق ساعتی از رگرسیون بردار پشتیبان استفاده شده است. در مرجع [۶] از ساختار شبکه عصبی بازگشتی^۵ استفاده شده که داده‌های ورودی به‌منظور بهبود آموزش و برآورد عملکرد به‌طور پی‌درپی انتخاب شده‌اند. در مرجع [۷] یک شبکه عصبی کانولوشن چند مقیاسه با شناخت زمان^۶ برای پیش‌بینی بار کوتاه-مدت چند مرحله‌ای پیشنهاد شده است. پیش‌بینی بار با یک چهارچوب یادگیری عمیق ترکیبی با هدف پیش‌بینی مصرف انرژی از خانواده‌های منفرد در مرجع [۸] ارائه شده است. ابتدا، داده‌های اصلی مصارف انرژی با استفاده از تبدیل موجک ثابت^۷ به چندین سیگنال فرعی تجزیه می‌شود. سپس هر سیگنال فرعی به‌عنوان یک داده سری زمانی مستقل در نظر گرفته می‌شود. یک شبکه عصبی حافظه کوتاه-مدت به هر سیگنال فرعی متصل است و نتایج پیش‌بینی را ایجاد می‌کند. در مورد روش‌های مختلف تجزیه داده‌ها و سری‌های زمانی تقویمی بار در مرجع [۹] یک مدل ترکیبی جهت پیش‌بینی کوتاه-مدت بار ارائه گردیده است. به این صورت که ابتدا از روش تجزیه کامل حالت تجربی گروه با سر و صدای سازگار^۸ برای تجزیه توالی اصلی برای همه اجزا و سپس از روش تجزیه حالت تنوع^۹ جهت به‌دست آوردن یک توالی غیرخطی و در نهایت از روش بازسازی فاز^{۱۰} جهت بهینه‌سازی تمام اجزا استفاده شده است. پس از آن داده‌های تجزیه شده به یک ماشین بردار پشتیبان بهینه شده توسط الگوریتم بهینه‌سازی نهنگ جهت آموزش و پیش‌بینی بار وارد می‌شوند. در مرجع [۱۰]

مدل ترکیبی تجزیه و تحلیل طیف‌ها برای پیش‌پردازش داده‌ها و الگوریتم تکاملی چند هدفه مبتنی بر تجزیه^{۱۱} را جهت پیش‌بینی بار پیشنهاد می‌کند. در مرجع [۱۱] روش پیشنهادی شامل دو مرحله است. در مرحله اول، مجموعه‌ای از مراحل پیش‌پردازش داده‌ها که شامل پاک‌سازی داده‌ها، انتخاب ایستگاه هواشناسی و ترکیب سری‌های دمایی است و مرحله دوم از یک شبکه عصبی بهبودیافته با مدل فازی تاکاگی سوگنو^{۱۲} توسعه‌یافته برای پیش‌بینی بار استفاده می‌نماید. در سوابق پژوهش‌های مبتنی بر تبدیل موجک^{۱۳} و شبکه‌های عصبی، به داده‌های ورودی برای آموزش توجه زیادی شده و هر گروه از پژوهشگران روشی را برای استفاده از داده‌های خام یا وزندهی به آن‌ها استفاده نموده‌اند. در مرجع [۱۲] یک مدل شبکه عصبی موجک کامل برای پیش‌بینی بار کوتاه-مدت استفاده شده که این مدل شامل دو روش بسته کامل تبدیل موجک و شبکه عصبی چندلایه است. در مرجع [۱۳] استفاده از فیلتر کالمن^{۱۴}، شبکه عصبی یا موجک و شبکه عصبی به‌تنهایی برای پیش‌بینی بار کوتاه-مدت استفاده شده است. در مرجع [۱۴] یک شبکه عصبی کانولوشن موجک چندگانه^{۱۵} برای پیش‌بینی بار پیشنهاد شده و یک روش چند سطحی برای انتخاب ویژگی‌ها، تجزیه بار و بازسازی موجک به کار گرفته شده است. در مرجع [۱۵] الگوی جدیدی به نام مکانیزم توجه ورودی بر مبنای وزن‌دهی لایه‌های ورودی و مکانیزم اتصال پنهان برای افزایش دقت پیش‌بینی و کارایی شبکه‌های عصبی بازگشتی و مدل‌های مبتنی بر شبکه‌های عصبی بازگشتی دوطرفه^{۱۶} برای پیش‌بینی بار کوتاه-مدت ارائه شده است. در مرجع [۱۶] یک مدل سه مرحله‌ای ارائه شده است. مرحله اول، شامل پیش‌پردازش داده‌ها و انتخاب ویژگی است. در مرحله دوم، آموزش و پیش‌بینی مبتنی بر روش یادگیری عمیق ماشین بولتزمن محدود شرطی^{۱۷} انجام شده و در مرحله سوم، الگوریتم بهینه‌سازی باد محور^{۱۸} استفاده شده است. کاربرد شبکه‌های عصبی عمیق در پیش‌بینی بار نیز در پژوهش‌ها اشاره شده که به‌عنوان نمونه‌ای از آن‌ها در مرجع [۱۷] می‌توان به ماشین یادگیری شدید عمیق^{۱۹} اشاره نمود که جهت پیش‌بینی تولید انرژی در یک نیروگاه برق ارائه شده است. در مرجع [۱۸] جنگل رگرسیون عمیق برای پیش‌بینی بار کوتاه-مدت در سیستم‌های قدرت ارائه شده است. این روش، یک الگوریتم یکپارچه شامل دو روش اسکن چند دانه^{۲۰} و روش جنگل آبشار^{۲۱} و بر اساس درخت تصمیم است. برای کاهش اثرات داده‌های خارج از محدوده نمونه‌ای در مرجع [۱۹] از روش نرمال‌سازی داده‌ها در پیش‌پردازش استفاده شده است. مجموعه داده‌های جمع‌آوری شده قبل از استفاده واقعی برای آموزش شبکه عصبی، تحت پیش‌پردازش داده‌ها قرار می‌گیرد. برای پیش‌بینی بار روش رگرسیون خطی چند متغیره و شبکه عصبی پیشنهاد شده است. در مرجع [۲۰] مجموعه‌ای از شبکه‌های عصبی با تابع شعاعی^{۲۲} پیشنهاد شده که با به حداقل رساندن خطای تعمیم موضعی^{۲۳} برای پیش‌بینی بار کوتاه-مدت و میان-مدت آموزش داده می‌شود. متغیرهای برون‌زا^{۲۴} و ویژگی‌های استخراج شده از سری بار (با شبکه‌های حافظه کوتاه-مدت و تبدیل موجک چند وضوح) در مقیاس‌های زمانی مختلف برای آموزش مجموعه شبکه‌های عصبی فوق استفاده می‌شود. در مرجع [۲۱] از سه روش معمول یادگیری ماشین شامل روش ماشین بردار پشتیبان، روش رگرسیون جنگل تصادفی^{۲۵} و روش شبکه عصبی حافظه کوتاه-مدت، برای پیش‌بینی بار استفاده شده است. با تلفیق مزایای سه روش و ترکیب با پیش‌پردازش داده‌ها، برای جداسازی داده‌های غیرعادی، یک روش پیش‌بینی ارائه شده است. در مرجع [۲۲] مدلی از تلفیق شبکه عصبی کانولوشن^{۲۶}، حافظه کوتاه-مدت طولانی^{۲۷} و تبدیل موجک گسسته^{۲۸} مطرح شده است. داده‌ها از طریق دو مسیر به مدل فوق قابل وارد شدن هستند. مسیر داده‌های خام مصارف بار مربوط به استخراج ویژگی دامنه زمان است در حالی که مسیر تبدیل موجک گسسته، مربوط به دامنه فرکانس است. در مرجع [۲۳] از روش‌های تجزیه و تحلیل طیف منفرد به عنوان فیلتر اولیه برای حذف نویزها و شبکه عصبی عمیق برای پیش‌بینی بار کوتاه مدت، استفاده شده است. در مرجع [۲۴]، برای پیش‌بینی بار بلند مدت، یک مدل یادگیری عمیق بهینه بر اساس شبکه حافظه کوتاه-مدت انباشته^{۲۹} پیشنهاد شده است. برای بهینه نمودن تابع ضرر از داده‌های حافظه طولانی کوتاه مدت و برای حذف بیش‌برازش و تنظیم از روش حذف تصادفی استفاده شده است. در این روش از الگوریتم ژنتیک^{۳۰} بهینه‌سازی مدل انجام شده است. در مرجع [۲۵] از یک روش یادگیری ماشین هوشمند با یک الگوریتم تکاملی^{۳۱} برای پیش‌بینی بار کوتاه-مدت استفاده شده است. این الگوریتم شامل مراحل مختلف از جمله تجزیه داده‌ها، پیش‌پردازش داده‌ها، انتخاب ویژگی، پیش‌بینی و تنظیم پارامتر است. از تبدیل موجک برای تجزیه سری‌های زمانی، از الگوریتم بهینه‌سازی ماهی^{۳۲} جهت انتخاب مجموعه‌ای از ویژگی‌های بهینه و سپس از الگوریتم بهینه‌سازی موج آب^{۳۳} با مدل شبکه عصبی المان^{۳۴}، برای فرآیند پیش‌بینی استفاده شده است.

۱-۲- نوآوری و ساختار مقاله

با توجه به مطالعات انجام شده می‌توان از نوآوری‌های این مقاله به موارد زیر اشاره کرد:

- استفاده از تبدیل موجک چندگانه جهت استخراج ویژگی و کاهش ابعاد ماتریس ورودی داده‌های بار الکتریکی.
- اعمال تبدیل موجک به داده‌های ورودی و استفاده از ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده که تابع ضرر نرم L1 در فرمول‌بندی آن به کار گرفته شده است. لازم به ذکر است که در فرمول‌بندی سه مدل ماشین یادگیری شدید اصلی، ماشین یادگیری شدید تنظیم‌شده و ماشین یادگیری شدید تنظیم‌شده و وزن‌دار از تابع ضرر L_2 نرم استفاده شده است و اینکه نرم L2 در برابر داده‌های خارج از محدوده ناپایدار بوده و نمی‌تواند عملکرد خوبی در برابر این دسته داده‌ها داشته باشد.
- ارائه مدل ترکیبی بهینه تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده به عنوان مدل برتر جهت پیش‌بینی بلندمدت بار الکتریکی.

ساختار مقاله به این شرح است. پس از بیان مسئله و اهمیت موضوع، در بخش ۲ روش پیشنهادی و فلوچارت مراحل اجرایی تشریح شده است. در بخش ۳ نتایج شبیه‌سازی ارائه شده است و در بخش ۴ نتیجه‌گیری روش پیشنهادی بیان می‌شود.

۲- روش پیشنهادی

روش پیشنهادی در این مقاله، ترکیب تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده^{۳۶} برای پیش‌بینی بار الکتریکی بلندمدت، است. مراحل اجرای کار شامل پیش‌پردازش داده‌ها، آموزش و اعتبارسنجی، پیش‌بینی نهایی و ارزیابی است. الگوریتم اجرایی در شکل (۱) ارائه شده است. در روش پیشنهادی، در مرحله پیش‌پردازش، استخراج ویژگی از داده‌های ورودی دما مطابق روش وانیل تائو^{۳۷} [۲۶] و با استفاده از تبدیل موجک انجام می‌شود. تبدیل موجک چندگانه یک‌بعدی در یک سطح به منظور استخراج ویژگی‌ها و کاهش ابعاد ماتریس داده‌ها، مورد استفاده قرار می‌گیرد. داده‌های بار الکتریکی نیز به دو مؤلفه فرکانس پایین و فرکانس بالا با استفاده از تبدیل موجک تجزیه شده و دو بردار هدف را تشکیل می‌دهند. سپس داده‌ها به دو دسته آموزش و آزمایش تقسیم‌بندی شده و جهت آموزش، به مدل‌های پیش‌بینی وارد می‌شوند.



شکل (۱): فلوچارت روش پیشنهادی

Figure (1): Flowchart of the proposed method

علاوه بر مدل ماشین مقاوم به داده‌های خارج از محدوده، سه مدل دیگر ماشین یادگیری شدید مانند ماشین یادگیری شدید اصلی^{۳۸}، ماشین یادگیری شدید تنظیم‌شده^{۳۶} و ماشین یادگیری شدید تنظیم‌شده وزن‌دار^{۴۰}، برای پیش‌بینی بار در نظر گرفته شده تا کارایی آن‌ها پس از پیش‌پردازش با تبدیل موجک، بررسی و بهترین روش مشخص گردد. برای آموزش مدل‌های پیش‌بینی، مقادیر بالا و پایین استخراج شده از مجموعه داده‌ها مورد استفاده قرار گرفته‌اند و برای هر یک از بردارهای هدف شامل مقادیر بالا و مقادیر پایین و با استفاده از هر یک از روش‌ها یک مدل مجزا آموزش داده شده است. در نهایت برای پیش‌بینی نیز ماتریس ویژگی‌ها به‌عنوان ورودی به هر یک از مدل‌ها داده شده و خروجی مدل مقادیر پایین با خروجی مدل مقادیر بالا جمع می‌شود تا پیش‌بینی نهایی را تشکیل دهد. اعتبارسنجی مدل‌ها، از طریق محاسبه معیارهای ارزیابی و با مقایسه مقدار پیش‌بینی شده خروجی مدل و مقادیر واقعی بار الکتریکی داده‌های آزمایش انجام می‌گیرد.

۱-۲- انتخاب داده‌ها

در این تحقیق داده‌های ساعتی بار و دما مربوط به رقابت جهانی پیش‌بینی انرژی در سال ۲۰۱۴^{۴۱} مورد استفاده قرار گرفته است. افق پیش‌بینی در این مطالعه افق سالانه است و به همین منظور از داده‌های سال‌های ۲۰۰۵ و ۲۰۰۶ و ۲۰۰۷ به‌عنوان داده‌های آموزش و داده‌های سال ۲۰۰۸ به‌عنوان داده‌های آزمایش انتخاب شده‌اند. ماتریس اصلی داده‌های ورودی جهت پیش‌بینی بار از مجموعه بار و ۲۵ ایستگاه هواشناسی استخراج شده است.

۲-۲- آماده‌سازی ماتریس ورودی

در این قسمت نحوه آماده‌سازی ماتریس ورودی بیان می‌شود.

۱-۲-۲- مدل وانیل تائو برای دما

در ابتدا با استفاده از موجک مادر دابچیز^{۴۲} ۱ موسوم به هار^{۴۳} داده‌های ورودی تجزیه شده و به دو بخش مؤلفه‌های فرکانس پایین به‌عنوان تقریب و مؤلفه‌های فرکانس بالا به‌عنوان جزئیات تقسیم می‌شوند. سپس استخراج و کاهش ابعاد ویژگی از داده‌ها با استفاده از تبدیل موجک و بر اساس مدل رگرسیون وانیل تائو انجام شده تا پس از آن ماتریس ورودی تشکیل و داده‌ها وارد چهار مدل ماشین یادگیری شدید شوند. این مدل یک مدل رگرسیون خطی چندگانه به‌صورت رابطه (۱) است [۲۷]:

$$y_t = \beta_0 + \beta_1 \text{Trend}_t + \beta_2 M_t + \beta_3 W_t + \beta_4 H_t + \beta_5 W_t H_t + f(T_t) \quad (1)$$

y_t پیش‌بینی بار برای زمان t (ساعت) است؛ β_i ضرایبی هستند که با استفاده از روش حداقل مربعات معمولی برآورد می‌شوند. متغیرهای طبقه‌بندی M_t ، W_t و H_t به‌ترتیب متغیرهای طبقه‌بندی ماه از سال، روز هفته و ساعت هستند. T_t دما در زمان t است و $f(T_t)$ از رابطه (۲) به‌دست می‌آید.

$$f(T_t) = \beta_6 T_t + \beta_7 T_t^2 + \beta_8 T_t^3 + \beta_9 T_t M_t + \beta_{10} T_t^2 M_t + \beta_{11} T_t^3 M_t + \beta_{12} T_t H_t + \beta_{13} T_t^2 H_t + \beta_{14} T_t^3 H_t \quad (2)$$

این مدل، فقط از متغیرهای دما و تقویم استفاده می‌کند [۲۷]. همبستگی آشکاری بین دما و ساعت روز یا ماه و سال وجود دارد. رابطه بین روز هفته و دما مشخص نیست. برای به‌دست آوردن این همبستگی‌ها اثرات متقابل در مدل گنجانده شده است.

۲-۲-۲- اثر تجمعی تأخر^{۴۴}

با درج دمای ساعات قبل در مدل، ویژگی‌های برجسته گم‌شده‌ای مانند تأخیر دما و بار الکتریکی و اختلاف بین بار پیش‌بینی شده و بار واقعی در برخی از زمان‌ها، در نظر گرفته می‌شود. دو شکل تأخیر دمای ساعتی $(T_{t-h}, h=1,2,\dots,72)$ و دمای میانگین متحرک ۲۴ ساعته (یا روزانه) که معمولاً مورد استفاده قرار می‌گیرد، انتخاب شده است. میانگین دمای متحرک روزانه روز d را می‌توان به‌صورت رابطه (۳) نوشت و سپس مدل معیار وانیل تائو را می‌توان به‌صورت رابطه (۴) گسترش داد [۲۷]. بر مبنای رابطه (۴) بردار ساعت، بردار هفته، بردار ماه و بردار سال با متوسط‌گیری از دمای ۲۵ ایستگاه هواشناسی ساخته شده و با تهیه

توان ۲ و توان ۳ دما و لحاظ نمودن اثر تأخر بر مبنای ساعت-دما ساخت بردار نهایی دما انجام می‌شود تا داده‌های ورودی آماده شوند.

$$\tilde{T}_{t,d} = \frac{1}{24} \sum_{h=24d-23}^{24d} T_{t-h}, \quad d=1,2,\dots,7. \quad (3)$$

$$y_t = \beta_0 + \beta_1 \text{Trend} + \beta_2 M_t + \beta_3 W_t + \beta_4 H_t + \beta_5 W_t H_t + f(T_t) + \underbrace{\sum_d f(\tilde{T}_{t,d}) + \sum_h f(T_{t-h})}_{\text{Recency Effect Terms}} \quad (4)$$

۲-۳- پیش‌پردازش

در مرحله پیش‌پردازش از تبدیل موجک یک سطحی برای تجزیه داده‌ها استفاده می‌شود. به این ترتیب که در ابتدا از تبدیل موجک چندگانه یک‌بعدی در یک سطح و با استفاده از موجک دابچیز به منظور استخراج ویژگی‌ها و کاهش ابعاد ماتریس ویژگی‌ها استفاده می‌شود. داده‌های هدف نیز با استفاده از موجک دابچیز به دو سطح مقادیر بالا (تقریب^{۴۵}) و مقادیر پایین (جزئیات^{۴۶}) تقسیم شده و دو بردار هدف را تشکیل می‌دهند. تعداد ویژگی‌های موجود در ماتریس ویژگی‌ها بر اساس سعی و خطا انتخاب شده است و پس از هر بار استخراج ویژگی مدل‌های پیش‌بینی آموزش داده شده و مورد آزمایش قرار گرفته‌اند و تعداد ویژگی‌هایی که کمترین خطای محاسبه شده را ارائه داده‌اند در نهایت انتخاب شده‌اند. فرآیند تجزیه موجک یک سطحی به صورت شکل (۲) آورده شده است.

۲-۴- تبدیل موجک

اگر به مجموعه اندازه‌گیری‌های انجام شده بر روی یک پدیده، اندیس زمان یا مکان داده شود، به سری اطلاعاتی خواهیم رسید که می‌توان از آن اطلاعات خاصی را استخراج نمود. در تبدیل موجک، تقسیم سیگنال به بخش‌های مختلف با استفاده از مقیاس‌گذاری و انتقال دادن یک تابع، انجام می‌شود. این تابع در طول سری اطلاعاتی انتقال پیدا می‌کند و برای هر موقعیت آن، طیف سری اطلاعاتی محاسبه می‌شود. این مراحل برای توابعی با مقیاس‌های مختلف تکرار می‌شود و در نهایت نتیجه حاصل به صورت مجموعه‌ای از اطلاعات آرگومان-فرکانس به دست می‌آید. تبدیل موجک اطلاعات را به صورت هم‌زمان در دو حوزه زمان و فرکانس ارائه می‌کند. در تبدیل موجک برخلاف تبدیل فوری که سیگنال یا سری اطلاعاتی را بر روی توابع سینوسی و کسینوسی و هارمونیک‌های آن‌ها تجزیه می‌نماید، سیگنال بر روی یک دسته از توابع برگرفته از موج مادر که موجک نامیده می‌شوند، تصویر می‌گردد. واژه مادر نیز به این منظور به کار برده می‌شود که تمامی نسخه‌های انتقال یافته و مقیاس شده همگی از روی یک تابع اولیه به دست می‌آید که اصطلاحاً موجک مادر نامیده می‌شود. برخلاف توابع سینوسی و کسینوسی در تبدیل فوری تابع موجک در فضای زمان محدود بوده و بعد از چند ارتعاش به سرعت به سمت صفر میل می‌کند [۲۸].

۲-۴-۱- تبدیل موجک پیوسته^{۴۷}

در رابطه (۵) پارامترهای τ و s به ترتیب پارامترهای انتقال و مقیاس هستند. انتقال یک موجک به معنای تأخیر انداختن آن است و باعث می‌شود موجک مورد نظر به سمت راست انتقال پیدا کرده و از اول تا انتهای سیگنال را طی کند.

$$\text{CWT}_x^\psi = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-\tau}{s}\right) dt \quad (5)$$



شکل (۲): فرآیند تبدیل موجک یک سطحی

Figure (2): The Process of one level wavelet transforms

۲-۴-۲- تبدیل موجک گسسته^{۴۸}

در تبدیل موجک گسسته، سیگنال از یک سری فیلترهای بالا گذر برای آنالیز فرکانس‌های بالا و از یک سری فیلترهای پایین گذر برای آنالیز فرکانس‌های پایین، عبور داده می‌شود. سیگنال به دو بخش تقسیم می‌شود: بخش حاصل از عبور سیگنال از فیلتر بالا گذر که شامل اطلاعات فرکانس بالا مانند نویز است و جزئیات نام دارد و بخش حاصل از عبور سیگنال از فیلتر پایین گذر که شامل اطاعات فرکانس پایین و در برگیرنده مشخصات هویتی سیگنال است و کلیات نامیده می‌شود. به بیان دیگر سیگنال اصلی در هر مرحله به دو سیگنال تقریب و جزئیات تجزیه می‌شود. سیگنال تقریب، حاوی مؤلفه‌های مقیاس بالا (فرکانس پایین) و سیگنال جزئیات، حاوی مؤلفه‌های مقیاس پایین و یا فرکانس بالای سیگنال است. به عبارت دیگر در هر بار تجزیه یک باند فرکانسی از سیگنال اصلی جدا شده و باقیمانده در سیگنال تقریب ذخیره می‌شود، بنابراین در حالت کلی می‌توان سیگنال اصلی را از حاصل جمع سیگنال‌های به دست آمده به دست آورد [۲۸]. اگر سیگنال اصلی با $a_0(t)$ نشان داده شود، با استفاده از تبدیل گسسته موجک می‌توان آن را در سطوح مختلف تجزیه نمود. سیگنال اصلی از جمع سیگنال‌های حاصل شده به دست می‌آید. به عبارت دیگر اگر سیگنال مورد نظر به n سطح تجزیه گردد سیگنال اولیه را می‌توان از جمع سیگنال تقریب در آخرین سطح و جمع توابع جزئیات در سطوح مختلف به صورت رابطه (۶) به دست آورد. بنابراین عملیات تجزیه را می‌توان تا سطحی که اطلاعات قابل ملاحظه‌ای در تقریب باقی نماند ادامه داد. این کار با مقایسه بی‌نظمی در دو سیگنال تقریب و جزئیات قابل حصول است، بنابراین با توجه به اینکه اطلاعات فراوانی در تقریب باقی نمی‌ماند، می‌توان سیگنال اولیه را با استفاده از سیگنال‌های جزئیات مطابق رابطه (۷) به‌تنهایی بازسازی نمود بدون آنکه اطلاعات قابل ملاحظه‌ای از دست رود.

$$a_0(t) = a_n(t) + \sum_{j=1}^n d_j(t) \quad (6)$$

$$a_0(t) = \sum_{j=1}^n d_j(t) \quad (7)$$

۲-۵- آموزش ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده

پس از استخراج ویژگی‌ها، داده‌های آماده شده برای آموزش به مدل ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده وارد می‌گردند. مؤلفه‌های خروجی مقادیر پایین مدل‌ها با مؤلفه‌های خروجی مقادیر بالای مدل‌ها جمع می‌شود تا پیش‌بینی نهایی را تشکیل دهند. به‌طور کلی نقاط خارج از محدوده، درصد کمی از کل نمونه‌های آموزش را تشکیل می‌دهند. این ویژگی را می‌توان با پراکندگی و برای خطای آموزش e ، توصیف نمود. پراکندگی نیز، با استفاده از نرم ℓ_0 به جای نرم ℓ_2 بهتر قابل توصیف است [۲۹]. بنابراین، وزن خروجی β با نرم ℓ_2 کوچک به‌عنوان جایگزین در نظر گرفته می‌شود به‌طوری‌که خطای آموزش e پراکندگی است. خطای آموزش طبق رابطه (۸) قابل تعریف است:

$$\min_{\beta} C \|e\|_0 + \|\beta\|_2^2 \quad \text{Subject to } y - H\beta = e \quad (8)$$

با این‌حال، رابطه (۸) یک مسئله برنامه‌نویسی غیرمحدب است. یک راه آسان برای حل این مشکل این است که آن را به شکل برنامه‌ریزی محدب قابل انعطاف بدون از دست دادن ویژگی پراکندگی بازسازی نمود. با توجه به میزان فشردگی و تجزیه و تحلیل مؤلفه اصلی قوی، مدت‌زمان پراکندگی را می‌توان با نرم ℓ_1 به دست آورد. به‌راحتی می‌توان دریافت که جایگزینی نرم ℓ_0 با نرم ℓ_1 در رابطه (۸) نه تنها ویژگی پراکندگی را تضمین می‌کند بلکه به حداقل‌سازی تحدب کلی منجر می‌شود. در نتیجه، رابطه برنامه‌ریزی محدب^{۴۹} قابل انعطاف را به دست می‌آوریم که فرم آن به شرح رابطه (۹) است:

$$\min_{\beta} \|e\|_1 + \frac{1}{C} \|\beta\|_2^2 \quad \text{Subject to } y - H\beta = e \quad (9)$$

رابطه (۹) یک مسئله بهینه‌سازی محدب محدود است و کاملاً متناسب با دامنه قابل استفاده از روش ضریب لاگرانژ تقویت‌شده^{۵۰} است. به همین ترتیب، تابع لاگرانژ تقویت‌شده توسط رابطه (۱۰) ارائه می‌شود:

$$L_{\mu}(e, \beta, \lambda) = \|e\|_1 + \frac{1}{C} \|\beta\|_2^2 + \lambda^T (y - H\beta - e) + \frac{\mu}{2} \|y - H\beta - e\|_2^2 \quad (10)$$

که در آن λ یک بردار از ضریب لاگرانژ است و μ یک پارامتر جریمه است. مطابق مرجع [۳۰] در این مقاله نیز پارامتر جریمه به صورت $\mu = 2N/\|y\|_1$ انتخاب شده است. الگوریتم ضریب لاگرانژ تقویت‌شده راه‌حل بهینه (e, β) و ضریب لاگرانژ λ را با به حداقل رساندن تابع لاگرانژ تقویت شده، به صورت سعی و خطا تخمین می‌زند. با توجه به (λ_k, μ) طرح روش سعی و خطا برای تکرار الگوریتم فوق به صورت رابطه (۱۱) است. حداقل نمودن خطا در تکرار این الگوریتم، در مرحله اول [به‌عنوان مثال، معادله (۱۱) - a] می‌تواند با استفاده از روش تناوبی که به‌طور گسترده‌ای دو مجهول e و β را به‌هنگام‌سازی می‌کند، اجرا شود [۲۹]. به عبارت دقیق‌تر، هدف الگوریتم ضریب لاگرانژ تقویت‌شده حل مسئله به صورت رابطه (۱۲) برای ایجاد تکرار جدید است.

$$\begin{cases} (e_{k+1}, \beta_{k+1}) = \arg \min_{e, \beta} L_{\mu}(e, \beta, \lambda_k) & (a) \\ \lambda_{k+1} = \lambda_k + \mu(y - H\beta_{k+1} - e_{k+1}) & (b) \end{cases} \quad (11)$$

$$\begin{cases} \beta_{k+1} = \arg \min_{\beta} L_{\mu}(e_k, \beta, \lambda_k) & (a) \\ e_{k+1} = \arg \min_e L_{\mu}(e, \beta_{k+1}, \lambda_k) & (b) \\ \lambda_{k+1} = \lambda_k + \mu(y - H\beta_{k+1} - e_{k+1}) & (c) \end{cases} \quad (12)$$

برای β_{k+1} راه‌حل (۱۳) در نظر گرفته می‌شود و برای e_{k+1} رابطه (۱۴) تعیین می‌شود:

$$\beta_{k+1} = (H^T H + 2/C\mu I)^{-1} H^T (y - e_k + \frac{\lambda_k}{\mu}) \quad (13)$$

$$e_{k+1} = \text{Shrink}(y - H\beta_{k+1} + \lambda_k/\mu, 1/\mu) \square \max\{|y - H\beta_{k+1} + \lambda_k/\mu| - 1/\mu, 0\} \circ \text{sign}(y - H\beta_{k+1} + \lambda_k/\mu) \quad (14)$$

در رابطه (۱۴)، (\circ) نشان‌دهنده ضرب عناصر است. به راحتی می‌توان دریافت که مقدار محاسبه هر تکرار عمده‌تاً برای حل معکوس $H^T H + 2/c\mu I$ استفاده می‌شود. خوشبختانه در هر تکرار، همان ماتریس $(H^T H + 2/C\mu I)^{-1} H^T$ وجود دارد که می‌تواند قبل از شروع تکرارهای کلی از قبل محاسبه شود، بنابراین محاسبات را خیلی سریع انجام می‌دهد. در همین حال، با تغییر μ و با تکرار کمتر، همگرایی تسریع می‌شود [۳۱، ۳۲]. با این حال، این نیاز به محاسبات مختلف $(H^T H + 2/C\mu I)^{-1} H^T$ دارد که به مقدار بیشتری از عملیات محاسبه‌ای منجر می‌شود. همچنین، باید توجه داشت که اگرچه این روش تکراری لزوماً سریع‌ترین روش برای دستیابی به دقت بسیار بالا در هنگام آموزش با استفاده از داده‌ها، فاقد داده‌های خارج از محدوده نیست، اما مسلماً در به‌دست آوردن بهترین دقت در مواردی که داده‌ها زیاد هستند، مؤثرترین روش است. برای سهولت وقتی تعداد تکرارها به حداکثر آستانه از پیش تعیین شده می‌رسد، تکرارها را خاتمه می‌دهیم. سایر معیارهای احتمالی خاتمه را می‌توان در مرجع‌های [۳۳] و [۳۴] یافت.

۲-۶- اعتبار سنجی و مقایسه مدل

برای سنجش و مقایسه کارایی و دقت روش پیشنهادی، داده‌های آماده شده برای آموزش سه مدل دیگر شامل ماشین یادگیری شدید، ماشین یادگیری شدید تنظیم‌شده، ماشین یادگیری شدید تنظیم‌شده و وزن‌دار نیز استفاده شده و پیش‌بینی بار انجام گردیده است. معیارهای میانگین درصد خطای مطلق، میانگین مربعات خطا، جذر میانگین مربعات خطا، ضریب تعیین و میانگین مطلق خطا برای ارزیابی عملکرد هر یک از مدل‌ها به کار گرفته شده است. روابط ریاضی محاسبه هر یک از معیارهای ارزیابی مطابق روابط (۱۵) تا (۱۹) بیان شده است.

الف- میانگین مربعات خطا^{۵۱}

$$MSE = \frac{\sum_{i=1}^N (y_i - \lambda(x_i))^2}{n} \quad (15)$$

ب- ضریب تعیین^{۵۲}

$$R^2 = \frac{(\sum_{i=1}^N (O_i - O_i^-)(P_i - P_i^-))^2}{\sum_{i=1}^N (O_i - O_i^-)^2 \sum_{i=1}^N (P_i - P_i^-)^2} \quad (16)$$

ج- جذر میانگین مربعات خطا^{۵۳}

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{N}} \quad (17)$$

د- میانگین درصد مطلق خطا^{۵۴}

$$MAPE = \frac{1}{n} \sum_{i=1}^N \left| \frac{A_i - F_i}{A_i} \right| \quad (18)$$

ه- میانگین مطلق خطا^{۵۵}

$$MAE = \frac{1}{n} \sum_{i=1}^N |x_i - \hat{x}| \quad (19)$$

در روابط (۱۶) و (۱۷) مقدار مشاهده شده بار (واقعی) و O_i مقدار متوسط بار مشاهده شده (واقعی)، P_i مقدار بار پیش‌بینی شده، P_i^- مقدار متوسط بار پیش‌بینی شده و N تعداد نمونه‌ها هستند. در رابطه (۱۵)، y_i مقدار هدف واقعی برای نمونه آزمایش x_i ، $\lambda(x_i)$ مقدار هدف پیش‌بینی شده برای نمونه آزمایش x_i و n تعداد نمونه‌های آزمون بوده و در رابطه (۱۸) مقدار واقعی و F_i مقدار پیش‌بینی شده و در رابطه (۱۹) مقدار اندازه‌گیری شده و x مقدار واقعی، هستند.

۳- نتایج شبیه‌سازی

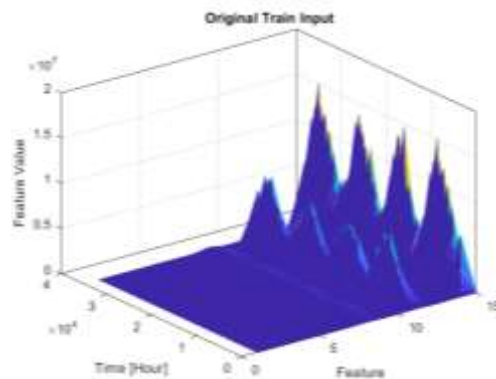
تمامی شبیه‌سازی‌های انجام شده در این مقاله در نسخه متلب ۲۰۱۸ انجام شده است و همه آزمایش‌ها و شبیه‌سازی‌ها با رایانه با ویندوز ۱۰، ۶۴ بیت، پردازشگر Core i-7 7500U، حافظه ۱۶ گیگابایت (RAM 16G) انجام شده‌اند.

۳-۱- ویژگی‌های داده‌ها قبل از پیش‌پردازش

بر مبنای داده‌های بار و دمای ساعتی استخراج شده از پایگاه داده GEFCOM2014 [۳۵]، عناوین ۱۵ ستون ماتریس داده‌ها بر مبنای مدل وانیل تائو عبارت‌اند از: ستون ۱- عدد ۱، ستون ۲- روز ماه، ستون ۳- روز هفته، ستون ۴- ساعت روز، ستون ۵- حاصل‌ضرب روز هفته در ساعت، ستون ۶- دما، ستون ۷- توان دوم دما، ستون ۸- توان سوم دما، ستون ۹- حاصل‌ضرب دما در روز ماه، ستون ۱۰- حاصل‌ضرب توان دوم دما در روز ماه، ستون ۱۱- حاصل‌ضرب توان سوم دما در روز ماه، ستون ۱۲- حاصل‌ضرب دما در ساعت، ستون ۱۳- حاصل‌ضرب توان دوم دما در ساعت، ستون ۱۴- حاصل‌ضرب توان سوم دما در ساعت، ستون ۱۵- اثر تأخر. نمودار تعداد و مقادیر ویژگی‌های داده‌های آموزش قبل از پیش‌پردازش بر اساس ماتریس اصلی ورودی با عناوین فوق، در شکل (۳) نشان داده شده است. محور X در شکل (۳) نشان‌دهنده تعداد ویژگی‌های^{۵۶} داده‌های ورودی (تعداد ستون‌های ماتریس ورودی) مرحله آموزش برحسب زمان و محور Y زمان‌های اندازه‌گیری (به‌صورت ساعتی) و محور Z مقدار عددی هر یک از ویژگی‌ها در هر زمان است. در اینجا تعداد ویژگی‌ها ۱۵ عدد و شامل داده‌های ورودی اولیه دمای هوا از ۲۵ ایستگاه هواشناسی و داده‌های زمانی شامل ساعت، روز، هفته و ماه و سال است.

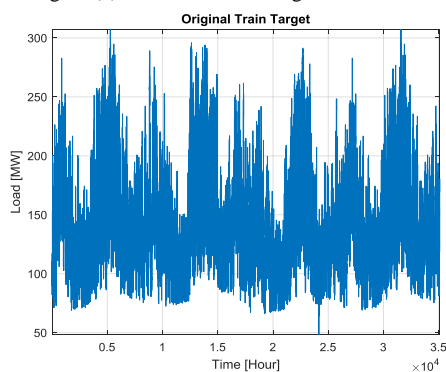
۳-۲- نمودار داده‌های بار الکتریکی برای آموزش مدل‌ها

شکل‌های (۴) و (۵) داده‌های بار الکتریکی ساعتی ورودی اصلی قبل از پیش‌پردازش در سال‌های ۲۰۰۵، ۲۰۰۶، ۲۰۰۷ و ۲۰۰۸ را نشان می‌دهند که برای آموزش و آزمایش مدل‌ها استفاده می‌شوند.



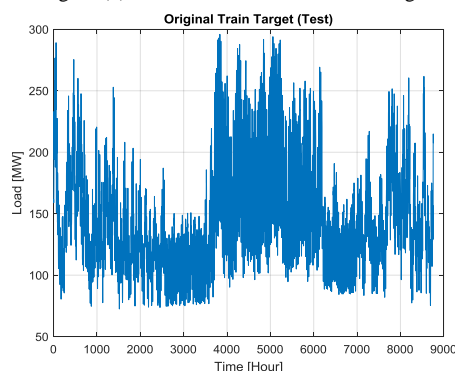
شکل (۳): ویژگی‌های داده‌های اصلی ورودی آموزش

Figure (3): The Features of original train data



شکل (۴): داده‌های بار الکتریکی مرحله آموزش

Figure (4): Electrical Load data in train stage



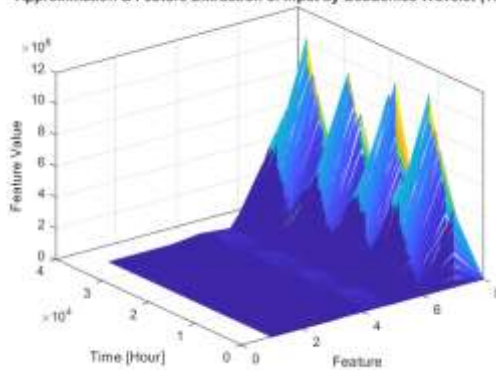
شکل (۵): نمودار بار الکتریکی در مرحله آزمایش

Figure (5): The Chart of electrical load in test stage

۳-۳- استخراج ویژگی و کاهش ابعاد برای داده‌های آموزش مدل‌ها

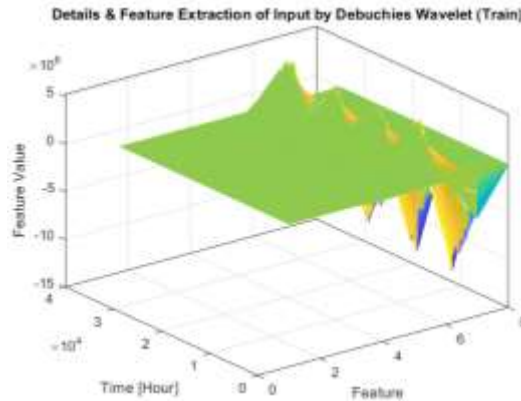
در این مقاله تبدیل موجک گسسته بر روی داده‌های ورودی و از روش نمونه‌گیری پایین^{۵۷} استفاده می‌شود و موجک استفاده شده هر بوده و در یک سطح انجام شده است. با اعمال این روش، استخراج ویژگی از داده‌های ورودی اولیه انجام و تعداد ویژگی‌ها از ۱۵ عدد به ۸ عدد تقلیل داده شده تا هر یک از بردارهای دما و زمان ساخته شود. ماتریس ورودی که ۱۵ ستون دارد به ۸ ستون کاسته شده و سطرهای ماتریس داده‌ها نیز به دو دسته تجزیه می‌شود. تعداد ۱۷۵۴۴ سطر به داده‌های فرکانس پایین اختصاص پیدا نموده و ۱۷۵۴۴ سطر نیز جهت داده‌های فرکانس بالا اختصاص یافته تا ماتریس ورودی برای آموزش مدل‌های پیش‌بینی آماده شود. در شکل‌های (۶) و (۷) نمودارهای مؤلفه‌های فرکانس پایین و فرکانس بالا داده‌های آموزش با استفاده از استخراج ویژگی نشان داده شده است. در دو شکل (۶) و (۷) محور X نشان‌دهنده ۸ ویژگی جهت داده‌های فرکانس پایین و فرکانس بالا و محور Y تعداد ساعات اندازه‌گیری شده و محور Z نشان‌دهنده مقادیر آن ۸ ویژگی است.

Approximation & Feature Extraction of Input by Debuchies Wavelet (Train)



شکل (۶): نمودار مؤلفه‌های فرکانس پایین استخراج ویژگی‌ها از داده‌های ورودی در مرحله آموزش

Figure (6): The Chart of approximations components of feature extraction from input train data

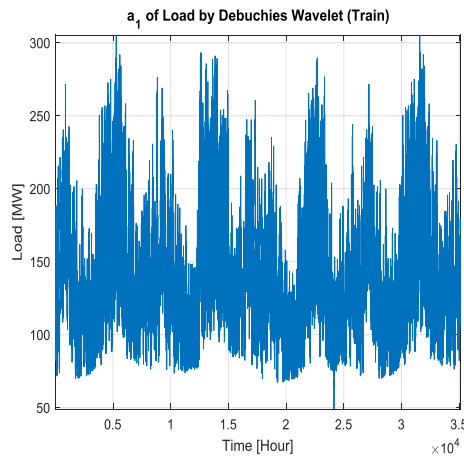


شکل (۷): نمودار مولفه‌های فرکانس بالا استخراج ویژگی‌ها از داده‌های ورودی در مرحله آموزش

Figure (7): The Chart of details components of feature extraction from input train data

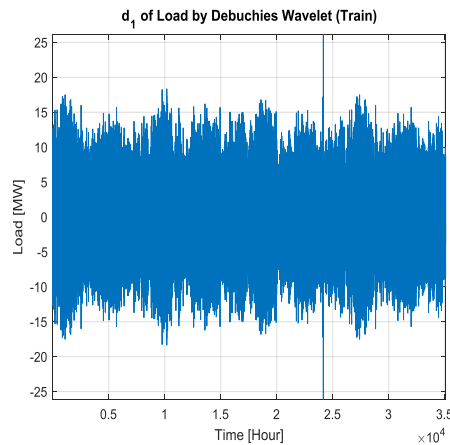
۳-۴- تجزیه داده‌های بار الکتریکی برای مرحله آموزش

داده‌های بار الکتریکی مورد استفاده برای آموزش مدل‌ها را با استفاده از تبدیل موجک گسسته به دو مؤلفه فرکانس پایین (تقریب) و فرکانس بالا (جزئیات) تجزیه می‌کنیم. در شکل‌های (۸) و (۹) نمودارهای مؤلفه‌های فرکانس پایین (a_1) و فرکانس بالا (d_1) که در یک سطح انجام شده نشان داده شده است.



شکل (۸): نمودار مؤلفه فرکانس پایین (a_1) بار الکتریکی داده‌های مرحله آموزش

Figure (8): The Chart of approximations components (a_1) of electrical load in train Data



شکل (۹): نمودار مؤلفه فرکانس بالا (d_1) بار الکتریکی داده‌های مرحله آموزش

Figure (9): The Chart of details components (d_1) of electrical load in train Data

۳-۵- تجزیه داده‌های بار الکتریکی برای مرحله آزمایش مدل

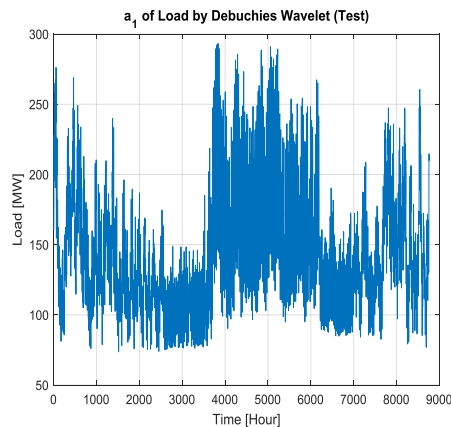
از داده‌های مرحله آزمایش نیز باید تبدیل موجک گرفت. در شکل‌های (۱۰) و (۱۱) نمودارهای مؤلفه‌های فرکانس پایین (a_1) و فرکانس بالا (d_1) از داده‌های آزمایش نشان داده شده است.

۳-۶- آموزش مدل پیش‌بینی

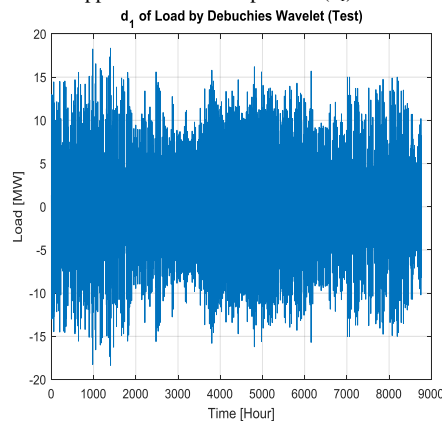
تعداد لایه‌های پنهان در مدل ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده، ۵۰ لایه و تابع فعال‌ساز^{۵۸} مدل نیز از نوع سیگموئید^{۵۹} است. برای آموزش مدل، داده‌های بخش آموزش توسط تبدیل موجک به دو مؤلفه فرکانس پایین و فرکانس بالا تجزیه شده و به مدل وارد می‌شود. در این مرحله مقدار مؤلفه‌های بار الکتریکی خروجی مدل با مؤلفه‌های بار الکتریکی داده‌های اصلی مقایسه می‌شود. برای اعتبار سنجی^{۶۰}، مؤلفه‌های فرکانس پایین و بالای بار الکتریکی پیش‌بینی شده توسط مدل با مؤلفه‌های فرکانس‌های پایین و بالای داده‌های آزمایش مقایسه می‌شود. برآورد کارایی هر مدل در مرحله آموزش و اعتبارسنجی با معیار جذر میانگین مربعات خطا انجام شده است.

۳-۷- نتایج ارزیابی آموزش مؤلفه‌های فرکانس پایین مدل

داده‌های مؤلفه‌های فرکانس پایین برای آموزش مدل ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده و سه مدل دیگر مورد استفاده قرار گرفته و نتایج اعتبار سنجی چهار مدل در جدول (۱) آورده شده است.



شکل (۱۰): نمودار مؤلفه فرکانس پایین (a_1) بار الکتریکی داده‌های مرحله آزمایش
Figure (10): The Chart of approximations components (a_1) of electrical load in test data



شکل (۱۱): نمودار مؤلفه فرکانس بالا (d_1) بار الکتریکی داده‌های مرحله آزمایش
Figure (11): The Chart of details components (d_1) of electrical load in test data

Table (1): The Results of validation error with trained model with approximations components
جدول (۱): نتایج خطای اعتبارسنجی آموزش مدل با داده‌های مؤلفه‌های فرکانس پایین

	Training RMSE	Validation RMSE	Training Time
ELM	۹۵۹۱/۶۹۱	۵/۷۲۲۵۹	۰/۰۴۸۳۵
RELM	۹۵۷۲/۳۶۶	۵/۷۹۶۰۸	۰/۰۴۶۴۲
WRELM	۹۵۷۷/۱۰۴	۵/۷۹۷۵۶	۰/۰۷۷۲۷
ORELM	۹۵۵۶/۴۰۴	۶/۰۳۲۵۸	۰/۰۹۴۷۱

برای آموزش مدل‌های یادگیری ماشین روش سعی و خطا و نرم خطای L^2 (نرم اقلیدسی^{۴۱}) استفاده می‌شود که حداقل نمودن مقدار خطا؛ هدف است. همان‌گونه که در جدول بالا مشخص است، از جذر میانگین مربعات خطا به‌عنوان معیار ارزیابی خطای مرحله آموزش استفاده شده است. مقدار این معیار در مرحله آموزش همه مدل‌ها با استفاده از داده‌های فرکانس پایین در حدود ۹۵۵۰ تا ۹۵۹۰ است. مقدار بالای خطای فوق به علت قرار گرفتن نرم اقلیدسی خطا در فرمول محاسبه آن بوده که نشان‌دهنده پراکندگی زیاد داده‌های آموزش ناشی از وجود داده‌های خارج از محدوده است. با این وجود، با استفاده از داده‌های آزمایش، جذر میانگین مربعات خطای مدل‌ها بسیار کاهش یافته و در محدوده ۵/۶ تا ۶/۰۴ قرار گرفته که مناسب بودن مدل را برای مقدار پیش‌بینی نشان می‌دهد.

۳-۸- نتایج ارزیابی آموزش مؤلفه‌های فرکانس بالای مدل

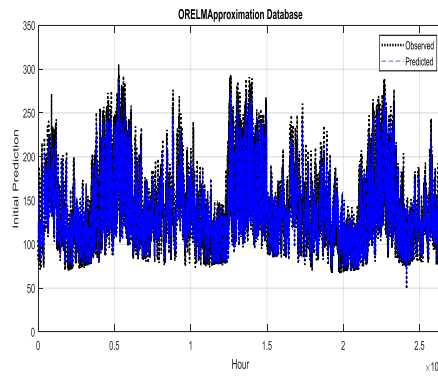
نتایج اعتبارسنجی آموزش مدل ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده و سایر مدل‌های مورد مقایسه با داده‌های مؤلفه‌های فرکانس بالا در جدول (۲) آورده شده است. مقادیر جدول (۲) شامل جذر میانگین مربعات خطای مرحله آموزش مدل اصلی و مدل‌های مورد مقایسه با استفاده از داده‌های فرکانس بالا از حدود ۱۱۵۰ تا ۱۱۶۰ است. مقدار این معیار در داده‌های فرکانس بالا نسبت به داده‌های فرکانس پایین بسیار کمتر است ولی همچنان از نظر عددی بالاست که در اینجا نیز نشان‌دهنده پراکندگی زیاد داده‌های آموزش است. در این مرحله نیز با استفاده از داده‌های آزمایش، جذر میانگین مربعات خطای مدل‌ها بسیار کاهش یافته و در محدوده ۰/۷۷ تا ۰/۷۹ قرار گرفته که مناسب بودن مدل را برای مقدار پیش‌بینی نشان می‌دهد.

۳-۹- نتایج آموزش و ارزیابی ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده

مدل ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده با داده‌های فرکانس پایین و فرکانس بالا آموزش داده می‌شود. در این مقاله پارامتر تنظیم C برای این مدل برابر 2^{-40} در نظر گرفته شده است. شکل‌های (۱۲) تا (۱۷) نمودارهای آموزش، اعتبارسنجی و پیش‌بینی نهایی مدل تبدیل مویک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده با داده‌های فرکانس پایین و فرکانس بالا را نشان می‌دهند.

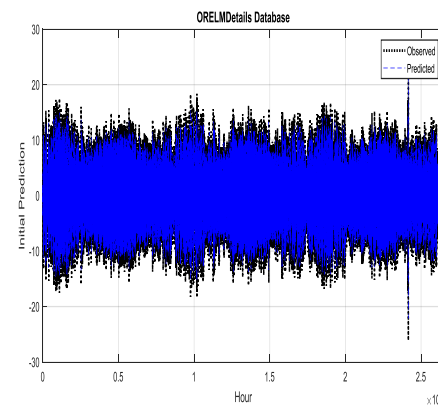
Table (2): The Results of validation error in trained model with details components
جدول (۲): نتایج خطای اعتبارسنجی آموزش مدل با داده‌های مؤلفه‌های فرکانس بالا

	Training RMSE	Validation RMSE	Training Time
ELM	۱۱۵۲/۷۳۷	۰/۷۹۶۵۶	۰/۰۳۴۰۶
RELM	۱۱۵۷/۸۷۶	۰/۷۷۲۹۲	۰/۰۳۶۹۱
WRELM	۱۱۵۴/۷۱۲	۰/۷۸۹۳۷	۰/۰۵۹۶۳
ORELM	۱۱۵۷/۵۵۳	۰/۷۸۱۹۶	۰/۰۷۱۹۱



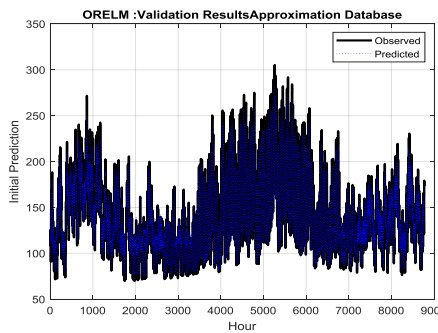
شکل (۱۲): نمودار پایگاه داده مؤلفه‌های فرکانس پایین تخمین زده‌شده توسط ORELM

Figure (12): The Chart of approximations database estimated by ORELM



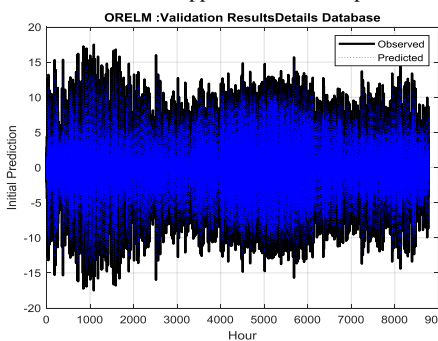
شکل (۱۳): نمودار پایگاه داده مؤلفه‌های فرکانس بالا تخمین زده‌شده توسط ORELM

Figure (13): The Chart of details database estimated by ORELM



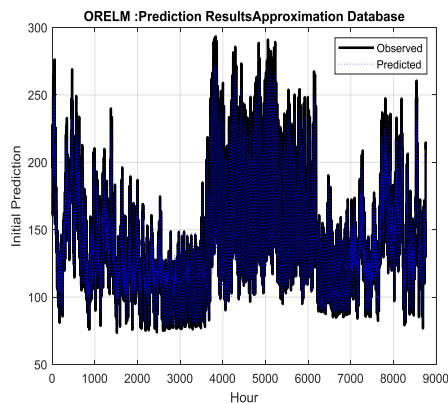
شکل (۱۴): نمودار اعتبارسنجی مؤلفه‌های فرکانس پایین تخمین زده‌شده توسط ORELM

Figure (14): The Chart of validation of approximations components estimated by ORELM

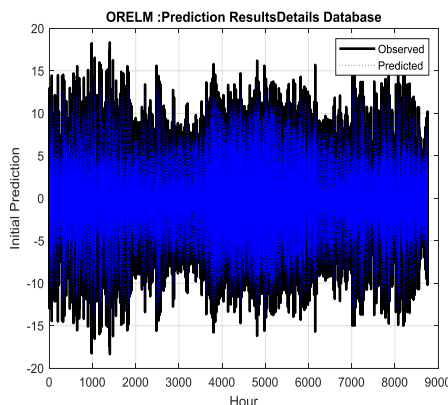


شکل (۱۵): نمودار اعتبارسنجی مؤلفه‌های فرکانس بالا تخمین زده‌شده توسط ORELM

Figure (15): The Chart of validation of details components estimated by ORELM



شکل (۱۶): نمودار پیش‌بینی بار نهایی مؤلفه‌های فرکانس پایین تخمین زده‌شده توسط ORELM
 Figure (16): The Chart of finally load forecast of approximations components estimated by ORELM



شکل (۱۷): نمودار پیش‌بینی بار نهایی مؤلفه‌های فرکانس بالا تخمین زده‌شده توسط ORELM
 Figure (17): The Chart of finally load forecast of details components estimated by ORELM

۳-۱۰- ارزیابی و مقایسه نتایج پیش‌بینی بار الکتریکی روش پیشنهادی

ارزیابی پیش‌بینی بار الکتریکی با روش ترکیب تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده و همچنین سه مدل مورد مقایسه دیگر شامل ماشین یادگیری شدید، ماشین یادگیری شدید تنظیم‌شده و ماشین یادگیری شدید تنظیم‌شده و وزن‌دار نیز انجام شده است. آموزش همه مدل‌ها، با داده‌های دو مؤلفه‌ای فرکانس پایین و فرکانس بالا مرحله آزمایش انجام و نتایج دو مؤلفه جمع شده و پیش‌بینی نهایی حاصل شده است. مقدار خطا، از مقایسه مقدار پیش‌بینی و مقدار واقعی بار الکتریکی داده‌های آزمایش، محاسبه می‌شود. مقایسه نتایج ارزیابی با استفاده از معیار میانگین درصد خطای مطلق در جدول (۳) درج شده است. در پیش‌بینی بار الکتریکی توسط مدل ترکیبی تبدیل موجک و ماشین یادگیری شدید اصلی مقدار میانگین درصد خطای مطلق برابر $3/5174$ بوده که با اضافه نمودن منظم‌سازی^{۶۲} در مدل ترکیبی تبدیل موجک و ماشین یادگیری شدید تنظیم‌شده به $3/216$ کاهش یافته است. وزن‌دهی به خطا و اعمال تبدیل موجک و آموزش در مدل ترکیبی تبدیل موجک و ماشین یادگیری شدید تنظیم‌شده و وزن‌دار منجر شده که مقدار میانگین درصد خطای مطلق $3/2319$ حاصل گردد. استفاده از تبدیل موجک و مدل مقاوم به داده‌های خارج محدوده تا حد امکان میانگین درصد خطای مطلق را کاهش داده و این معیار عدد $3/0966$ را حاصل نموده که بهترین نتیجه است. مقدار این معیار در مقایسه با نتیجه پیش‌بینی با سه مدل دیگر به ترتیب $0/4208$ نسبت به مدل ماشین یادگیری شدید اصلی، $0/1194$ نسبت به مدل تنظیم‌شده و $0/1353$ نسبت به مدل تنظیم‌شده و وزن‌دار، کمتر است. تأثیر تبدیل موجک و مقایسه کارایی مدل پیشنهادی ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده نسبت به روش‌های دیگر از طریق انجام پیش‌بینی بار با وارد نمودن داده‌های بدون اعمال تبدیل موجک به چهار روش

رگرسیون بردار پشتیبان^{۶۳}، درخت تصمیم، درخت تقویت‌شده^{۶۴} و درخت کیسه‌ای^{۶۵} مورد ارزیابی قرار گرفته که نتایج در جدول (۴) درج شده است.

روش‌های طبقه‌بندی و جایگزین نمودن آن‌ها با تبدیل موجک به‌عنوان روش‌های مورد مقایسه در پیش‌بینی بار الکتریکی در نظر گرفته شده‌اند. بهترین مقدار معیار میانگین درصد خطای مطلق عدد ۳/۰۹۶۶ است که توسط تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده (مدل پیشنهادی) به‌دست آمده است. مقدار معیار فوق در پیش‌بینی بار با مدل پیشنهادی به نسبت مدل‌های مورد مقایسه در جدول (۴)، به ترتیب مقادیر ۱/۴۴۱۲، ۰/۴۵۲۵، ۵/۸۳۱۱ و ۱/۰۹۹۱ کمتر شده است. نزدیک‌ترین نتیجه بر مبنای این معیار عدد ۳/۵۴۹۱، از روش درخت تصمیم حاصل شده، در حالیکه مقدار خطای پیش‌بینی سایر روش‌ها (بر مبنای معیار فوق)، بالاتر از این مقدار بوده که نشان‌دهنده کارایی و دقت کمتر این روش‌ها برای پیش‌بینی بار الکتریکی است. بنابراین این‌گونه نتیجه می‌گردد که اعمال تبدیل موجک و کاهش اثر داده‌های خارج از محدوده، باعث بهبود پیش‌بینی بار می‌گردد. کاربرد تبدیل موجک و ترکیب با روش‌های مختلف در اغلب پژوهش‌ها، برای پیش‌بینی بار در بازه کوتاه-مدت گزارش شده و مطالعات کمتری در بازه بلند-مدت انجام شده است. لذا، با توجه به تعداد بسیار بیشتر داده‌ها برای پیش‌بینی بلندمدت، خطای پیش‌بینی بیشتری برای روش‌ها مورد انتظار است. با این وجود، نتایج به‌دست آمده از روش پیشنهادی با برخی از جدیدترین پژوهش‌های انجام شده در سال‌های اخیر که به نوعی در آن‌ها از تبدیل موجک برای پیش‌پردازش داده‌ها استفاده شده مورد مقایسه قرار گرفته و نتایج آن در جدول (۵) ارائه شده است.

در روش‌های منفردی که در مرجع [۱۳] انجام شده است از داده‌های نوآسکوشیا در کانادا استفاده شده و ترکیب تبدیل موجک و شبکه عصبی به‌کار رفته است. در مرجع [۳۶] ترکیب تبدیل موجک مکرر و ماشین بردار پشتیبان با کرنل گوسی به‌کار گرفته شده است. همان‌گونه که در جدول (۵) مشخص است، روش ترکیبی پیشنهادی در این مقاله با توجه به این‌که بار الکتریکی یک ساله را تخمین می‌زند در مقایسه با روش‌های فوق که برای پیش‌بینی کوتاه-مدت ارائه شده، نیز بهتر عمل کرده است. خطای روش پیشنهادی با معیار میانگین درصد خطای مطلق، نسبت به روش‌های معرفی شده در مرجع [۱۳] به‌ترتیب ۱/۱۲۸۶ و ۰/۸۰۶۰۸ کمتر بوده و برای روش معرفی شده در مرجع [۳۶] به میزان ۲/۲۷۳۴ کمتر است.

Table (3): The Results of load forecasting with several extreme learning machine models based on Mean Absolute Percentage Error
جدول (۳): نتایج پیش‌بینی بار مدل‌های مختلف ماشین یادگیری شدید بر اساس معیار ارزیابی میانگین درصد خطای مطلق

	W-ELM	W-RELM	W-WRELM	W-ORELM
MAPE	۳/۵۱۷۴	۳/۲۱۶	۳/۲۳۱۹	۳/۰۹۶۶

Table (4): The Comparison of results with four Extreme Learning Machines Methods and the other methods
جدول (۴): مقایسه نتایج ارزیابی چهار روش ماشین یادگیری شدید و سایر روش‌ها

	MAE	MAPE	MSE	RMSE	R ²
Gaussian SVM	۶/۷۰۱۷	۴/۵۳۷۸	۸۳/۹۰۴	۹/۱۵۹۹	۰/۹۶۰۲۱
Decision Tree	۵/۱۷۹۷	۳/۵۴۹۱	۴۹/۷۲۶	۷/۰۵۱۶	۰/۹۷۶۴۲
Boosted Tree	۱۳/۷۸۱	۸/۹۲۷۷	۳۰۷/۶۳	۱۷/۵۳۹	۰/۸۵۴۱۳
Bagged Tree	۶/۲۵۱۹	۴/۱۹۵۷	۷۰/۷۳۱	۸/۳۸۸۷	۰/۹۶۶۶۳
W-ORELM	۴/۷۲۳۴	۳/۰۹۶۶	۳۹/۰۱۴	۶/۲۴۶۱	۰/۹۸۱۵

Table (5): The Comparison of proposed method and short-term load forecasting methods based on Mean Absolute Percentage Error
جدول (۵): مقایسه روش پیشنهادی با روش‌های ارائه شده جهت پیش‌بینی کوتاه-مدت بر اساس معیار ارزیابی میانگین درصد خطای

	ANN [13]	WNN [13]	RWT-SVM(Gaussian) [36]	Proposed Method
MAPE	۴/۲۲۵۲	۳/۹۰۲۶۸	۵/۳۷	۳/۰۹۶۶

۴- نتیجه‌گیری

در این مقاله ترکیب تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده، به‌عنوان روشی جدید و کارآمد و با دقت، با به‌کارگیری داده‌های مسابقات جهانی پیش‌بینی انرژی در سال ۲۰۱۴ شامل بار الکتریکی و دمای هوا به‌صورت ساعتی برای پیش‌بینی بار در افق سالانه ارائه شده است. نتایج شبیه‌سازی نشان داده که آموزش مدل پیش‌بینی با داده‌های مؤلفه‌های فرکانس پایین و سپس با داده‌های مؤلفه‌های فرکانس بالا، به‌تنهایی در مرحله ارزیابی با در نظر گرفتن معیارها نتایج متفاوت و گاه بیش‌برازش به‌دست می‌دهند. تجمیع نتایج خروجی دو مؤلفه فرکانس پایین و بالای مدل، نتایج بهتری را حاصل می‌نماید. از مقایسه ارزیابی‌های چهار مدل ماشین یادگیری شدید، ماشین یادگیری شدید تنظیم‌شده، ماشین یادگیری شدید تنظیم‌شده وزن‌دار و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده، مشاهده شد که مدل ترکیبی موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده با معیار ارزیابی میانگین درصد خطای مطلق به مقدار $۳/۰۹۶۶$ توانسته بهترین نتیجه را ارائه نماید. در مقایسه با مدل‌های منفرد و بدون پیش‌پردازش، مانند رگرسیون بردار پشتیبان، درخت تصمیم، درخت تقویت‌شده و درخت کیسه‌ای نیز مدل پیشنهادی توانسته است بهترین نتایج با معیارهای ارزیابی میانگین مطلق خطا، میانگین درصد خطای مطلق، میانگین مربعات خطا، جذر میانگین مربعات خطا و ضریب تعیین را به‌ترتیب برابر با مقادیر $۴/۷۲۳۴$ و $۳/۰۹۶۶$ و $۳۹/۰۱۴$ و $۶/۲۴۶۱$ و $۰/۹۸۱۵$ ایجاد نماید. به این ترتیب مدل ترکیبی تبدیل موجک و ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده نسبت به مدل‌های ارائه شده پیش‌بینی بلندمدت بار، خطای پیش‌بینی کمتری را ارائه داد. زمینه‌های دیگری که برای دقیق‌تر شدن روش پیشنهادی در آینده باید توسط پژوهشگران مورد تحقیق و بررسی قرار گیرند، به‌صورت زیر پیشنهاد می‌شود:

- تأثیر اعمال سطوح دیگر تبدیل موجک و سایر موجک‌های مادر با مدل ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده و مقایسه خطای پیش‌بینی بلندمدت بار الکتریکی.
- بررسی روش‌های دیگر پیش‌پردازش داده‌ها مانند تجزیه و تحلیل طیف منفرد و الگوریتم‌های بهینه‌سازی و ترکیب با ماشین یادگیری شدید مقاوم به داده‌های خارج از محدوده.
- افزودن داده‌های مربوط به سرعت باد و انرژی خورشیدی به پیش‌بینی بار بلندمدت.

References

مراجع

- [1] M. Shahidehpour, H. Yamin, Z. Li, "Market operations in electric power systems: forecasting, scheduling, and risk management", Wiley-IEEE Press, vol. 40, April 2002 (doi: 10.1002/047122412X).
- [2] L. Han, Y. Peng, Y. Li, B. Yong, Q. Zhou, L. Shu, "Enhanced deep networks for short-term and medium-term load forecasting", IEEE Access, vol. 7, pp. 4045–4055, 2019 (doi: 10.1109/ACCESS.2018.2888978).
- [3] A.A. Mamun, M. Sohel, N. Mohammad, M.S.H. Sunny, D.R. Dipta, E. Hossain, "A comprehensive review of the load forecasting techniques using single and hybrid predictive models", IEEE Access, vol. 8, pp. 134911–134939, 2020 (doi: 10.1109/ACCESS.2020.3010702).
- [4] J.C. Lopez, M.J. Rider, Q. Wu, "Parsimonious short-term load forecasting for optimal operation planning of electrical distribution systems", IEEE Trans. on Power Systems, vol. 34, no. 2, pp. 1427–1437, March 2019 (doi: 10.1109/TPWRS.2018.2872388).
- [5] G. Zhang, J. Guo, "A novel method for hourly electricity demand forecasting", IEEE on Trans. Power Systems, vol. 35, no. 2, pp. 1351–1363, March 2020 (doi: 10.1109/TPWRS.2019.2941277).
- [6] S.M. Baek, "Mid-term load pattern forecasting with recurrent artificial neural network", IEEE Access, vol. 7, pp. 172830–172838, 2019 (doi: 10.1109/ACCESS.2019.2957072).
- [7] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, Z. Zhu, "Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting", IEEE Access, vol. 7, pp. 88058–88071, 2019 (doi: 10.1109/ACCESS.2019.2926137).
- [8] K. Yan, W. Li, Z. Ji, M. Qi, Y. Du, "A hybrid LSTM neural network for energy consumption forecasting of individual households", IEEE Access, vol. 7, pp. 157633–157642, 2019 (doi: 10.1109/ACCESS.2019.29490-65).
- [9] W. Li, Q. Shi, M. Sibtain, D. Li, D.E. Mbanze, "A hybrid forecasting model for short-term power load based on sample entropy, two-phase decomposition and whale algorithm optimized support vector regression", IEEE Access, vol. 8, pp. 166907–166921, 2020 (doi: 10.1109/ACCESS.2020.3023143).

- [10] H. Bo, Y. Nie, J. Wang, "Electric load forecasting use a novelty hybrid model on the basic of data preprocessing technique and multi-objective optimization algorithm", *IEEE Access*, vol. 8, pp. 13858–13874, 2020 (doi: 10.1109/ACCESS.2020.2966641).
- [11] Z. Wen, L. Xie, Q. Fan, H. Feng, "Long term electric load forecasting based on TS-type recurrent fuzzy neural network model", *Electric Power Systems Research*, vol. 179, Article Number: 106106, Feb. 2020 (Doi: 10.1016/j.epsr.2019.106106).
- [12] M. El-Hendawi, Z. Wang, "An ensemble method of full wavelet packet transform and neural network for short term electrical load forecasting", *Electric Power Systems Research*, vol. 182, Article Number: 106265, May 2020 (doi: 10.1016/j.epsr.2020.106265).
- [13] H.H.H. Aly, "A proposed intelligent short-term load forecasting hybrid models of ANN, WNN and KF based on clustering techniques for smart grid", *Electric Power Systems Research*, vol. 182, Article Number: 106191, May 2020 (doi: 10.1016/j.epsr.2019.106191).
- [14] Z. Liao, H. Pan, X. Fan, Y. Zhang, L. Kuang, "Multiple wavelet convolutional neural network for short-term load forecasting", *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9730-9739, June 2021 (doi: 10.1109/JIOT.2020.3026733).
- [15] M. Zhang, Z. Yu, Z. Xu, "Short-term load forecasting using recurrent neural networks with input attention mechanism and hidden connection mechanism", *IEEE Access*, vol. 8, pp. 186514-186529, 2020 (doi: 10.1109/ACCESS.2020.3029224).
- [16] G. Hafeez, K.S. Alimgeer, I. Khan, "Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid", *Applied Energy*, vol. 269, Article Number: 114915, July 2020 (doi: 10.1016/j.apenergy.2020.114915).
- [17] S. Abbas, M.A. Khan, L.E. Falcon-Morales, A. Rehman, Y. Saeed, M. Zareei, A. Zeb, E.M. Mohamed, "Modeling, simulation and optimization of power plant energy sustainability for IoT enabled smart cities empowered with deep extreme learning machine", *IEEE Access*, vol. 8, pp. 39982-39997, 2020 (doi: 10.1109/ACCESS.2020.2976452).
- [18] L. Yin, Z. Sun, F. Gao, H. Liu, "Deep forest regression for short-term load forecasting of power systems", *IEEE Access*, vol. 8, pp. 49090-49099, 2020 (doi: 10.1109/ACCESS.2020.2979686).
- [19] M.V. Selvi, S. Mishra, "Investigation of performance of electric load power forecasting in multiple time horizons with new architecture realized in multivariate linear regression and feed-forward neural network techniques", *IEEE Trans. on Industry Applications*, vol. 56, no. 5, pp. 5603-5612, Sept.-Oct. 2020 (doi: 10.1109/TIA.2020.3009313).
- [20] C.S. Lai, Y. Yang, K. Pan, J. Zhang, H.L. Yuan HL, W.WY. Ng, Y. Gao, Z. Zhao, T. Wang, M. Shahidehpour, L.L. Lai, "Multi-view neural network ensemble for short and mid-term load forecasting", *IEEE Trans. on Power Systems*, vol. 36, no. 4, pp. 2992-3003, July 2021 (doi: 10.1109/TPWRS.2020.3042389).
- [21] W. Guo, L. Che, M. Shahidehpour, X. Wan, "Machine-learning based methods in short-term load forecasting", *The Electricity Journal*, vol. 34, no. 1, Article Number: 106884, Jan./Feb. 2021 (doi: 10.1016/j.t-ej.2020.106884).
- [22] X. Shao, C. Pu, Y. Zhang, C.S. Kim, "Domain fusion CNN-LSTM for short-term power consumption forecasting", *IEEE Access*, vol. 8, pp. 188352-188362, 2020 (doi: 10.1109/ACCESS.2020.3031958).
- [23] M.H. Pham, M.N. Nguyen, Y.K. Wu, "A novel short-term load forecasting method by combining the deep learning with singular spectrum analysis", *IEEE Access*, vol. 9, pp. 73736-73746, 2021 (doi: 10.1109/ACCESS.2021.3078900).
- [24] T.A. Farrag, E.E. Elattar, "Optimized deep stacked long short-term memory network for long-term load forecasting", *IEEE Access*, vol. 9, pp. 68511-68522, 2021 (doi: 10.1109/ACCESS.2021.3077275).
- [25] I.M. Mehedi, H. Bassi, M.J. Rawa, M. Ajour, A. Abusorrah, M.T. Vellingiri, Z. Salam, M.P. Bin, "Intelligent machine learning with evolutionary algorithm based short term load forecasting in power systems", *IEEE Access*, vol. 9, pp. 100113-100124, 2021 (doi: 10.1109/ACCESS.2021.3096918).
- [26] J. Luo, T. Hong, S.C. Fang, "Robust regression models for load forecasting", *IEEE Trans. on Smart Grid*, vol. 10, no. 5, pp. 5397-5404, Sept. 2019 (doi: 10.1109/TSG.2018.2881562).
- [27] P. Wang, B. Liu, T. Hong, "Electric load forecasting with recency effect: A big data approach", *Int Journal of Forecasting*, vol. 32, no. 3, pp. 585-597, July/Sept. 2016 (doi: 10.1016/j.ijforecast.2015.09.006).
- [28] A.J.R. Reis, A.P.A. Silva, "Feature extraction via multiresolution analysis for short-term load forecasting", *IEEE Trans. on Power Systems*, vol. 20, no. 1, pp. 189-198, Feb. 2005 (doi: 10.1109/TPWRS.2004.840380).
- [29] K. Zhang, M. Luo, "Outlier-robust extreme learning machine for regression problems", *Neurocomputing*, vol. 151, no. 3, pp. 1519-1527, March 2015 (doi: 10.1016/j.neucom.2014.09.022).
- [30] J. Yang, Y. Zhang, "Alternating direction algorithms for ℓ_1 -problems in compressive sensing", *SIAM Journal of Scientific Computing*, vol. 33, no. 1, pp. 250-278, 2011 (doi: 10.1137/090777761).
- [31] Z. Lin, M. Chen, Y. Ma, "The augmented lagrange multiplier method for exact Recovery of corrupted low-rank matrices", *Mathematics, Optimization and Control* pp. 2-28, Sept. 2010 (doi: 10.1016/j.jsb.2012.10.010).

- [32] W. Zuo, D. Meng, L. Zhang, X. Feng, D. Zhang, "A generalized iterated shrinkage algorithm for non-convex sparse coding", *Proceeding of the IEEE/ICCV*, pp. 217–224, Sydney, NSW, Australia, Dec. 2013 (doi: 10.1109/ICCV.2013.34).
- [33] D.L. Donoho, Y. Tsaig, "Fast solution of ℓ_1 -Norm minimization problems when the solution may be sparse", *IEEE Trans. on Information Theory*, vol. 54, no. 11, pp. 4789-4812, Nov. 2008 (doi: 10.1109/TIT.2008.92-9958).
- [34] M.A.T. Figueiredo, R.D. Nowak, S.J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems", *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586-597, Dec. 2007 (doi: 10.1109/JSTSP.2007.910281).
- [35] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, R.J. Hyndman, "Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond", *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, July/Sept. 2016 (doi: 10.1016/j.ijforecast.2016.02.001).
- [36] Aasim, S.N. Singh, A. Mohapatra, "Data driven day-ahead electrical load forecasting through repeated wavelet transform assisted SVM model", *Applied Soft Computing*, vol. 111, Article Number: 107730, Nov. 2021 (doi: 10.1016/j.asoc.2021.107730).

زیرنویس‌ها

1. Auto-regressive integrated moving average
2. Optimal operational planning
3. Time-dependency convolutional neural network (TD-CNN)
4. Cycle-based long short-term memory (C-LSTM)
5. Recurrent artificial neural network (RANN)
6. Multi-Scale convolutional neural network with Time-cognition (TCMS-CNN)
7. Stationary wavelet transform
8. Complete ensemble empirical mode decomposition with Adaptive noise (CEEMDAN)
9. Variational mode decomposition (VMD)
10. Phase space reconstruction (PSR)
11. Singular spectrum analysis-multi-objective evolutionary algorithm based on decomposition
12. Takagi sugenu (TS)
13. Wavelet-transform (WT)
14. Kalman filter
15. Multiple wavelet convolutional neural network (MWCNN)
16. Bidirectional recurrent neural network (Bi-RNN)
17. Factored conditional restricted boltzmann machine (FCRBM)
18. Genetic wind driven optimization (GWDO)
19. Deep extreme learning machine (DELm)
20. Multi-grained scanning
21. Cascade forest
22. Radial basis function neural networks (RBFNN)
23. Localized generalization error
24. Exogenous features
25. Random forest (RF)
26. Convolutional neural network (CNN)
27. Long short-term memory (LSTM)
28. Discrete wavelet transform (DWT)
29. Stacked long short-term memory network (SLSTMN)
30. Genetic algorithm (GA)
31. Intelligent machine learning with evolutionary algorithm (IMLEA)
32. Artificial fish swarm optimization algorithm (OAFSA)
33. Water wave optimization (WWO)
34. Elman neural networks (ENN)
35. Loss function
36. Outlier robust extreme learning machine (ORELM)
37. Vanilla- Tao's
38. Extreme learning machine (ELM)
39. Regularized extreme learning machine (RELM)
40. Weighted regularized extreme learning machine (WRELM)
41. Global energy forecasting competition 2014(GEFCOM2014)

41. Daubechies (db1)
43. Haar
44. Recency effect at an aggregated level
45. Approximation
46. Details
47. Continues wavelet transform (CWT)
48. Discrete wavelet transform (DWT)
49. Convex relaxation
50. Augmented lagrange multiplier (ALM)
51. Mean square error (MSE)
52. R squared (R^2)
53. Root mean square error (RMSE)
54. Mean absolute percentage error (MAPE)
55. Mean absolute error (MAE)
56. Features
57. Down-sampling
58. Activation function
59. Sigmoid
60. Validation
61. Frobenius norm
62. Regularization
63. Support vector regression
64. Boosted tree
65. Bagged tree