Research in Production and Operations Management
University of Isfahan E-ISSN: 2981-0329
Vol. 17, Issue 2, No. 45, 2026
DOI: 10.22108/pom.2026.148566.1658

(Research paper)

A Model for Predicting and Scheduling Short-Term Electrical Power Outages

Fatemeh Shahmoradi
Department of industrial engineering, Yazd university, Yazd, Iran, fatemehshahmoradi@stu.yazd.ac.ir
Majid Shakhsi-Niaei*
Department of industrial engineering, Yazd university, Yazd, Iran, m.niaei@yazd.ac.ir
Mohammad Mehdi Lotfi
Department of industrial engineering, Yazd university, Yazd, Iran, lotfi@yazd.ac.ir
Davoud Abootorabi Zarchi
Department of Electrical Engineering, Yazd University, Yazd, Iran, d.abootorabi@yazd.ac.ir

Purpose: The electricity industry is a fundamental pillar of industrial and economic development. However, the persistent imbalance between electricity supply and demand has become a major challenge in recent years, leading to planned power outages in the industrial sector and causing substantial economic losses. Interruptions in electricity supply negatively affect industrial production and development by reducing productivity, damaging equipment, delaying orders, and disrupting supply chains. In Iran, the imbalance between supply and demand—particularly during peak consumption periods such as summer and winter—has increased the likelihood of planned outages for industrial consumers. At the same time, rising electricity demand, limitations in generation capacity, and the prioritization of residential consumers by distribution companies have resulted in frequent, and in some cases unannounced, interruptions for industrial users. Therefore, short-term outage prediction can provide a valuable basis for proactive planning and help mitigate the impacts of unexpected interruptions on industrial operations. This study aims to develop a machine-learning-based approach for short-term prediction of electricity outages at the feeder and hourly levels in an industrial area in Iran.

Design/methodology/approach: The study uses data collected from industrial electricity feeders in an Iranian industrial zone over the period from March 20, 2024, to March 20, 2025. The dataset includes electricity load, outage records, weather information (temperature and wind speed), temporal and calendar variables, and

* Corresponding author, 0000-0001-8034-9955 2981-0329 / © University of Isfahan
This is an open access article under the CC-BY-NC-ND 4.0 License (https://creativecommons.org/licenses/by-nc-nd/4.0/)
How to cite: Shahmoradi, F. , Shakhsi-Niaei, M. , Lotfi, M. M. and Abootorabi Zarchi, D. (2026). A model for predicting and scheduling electrical power outages in short-term. *Research in Production and Operations Management*, 17(2), 137-156. doi: 10.22108/pom.2026.148566.1658

features derived from lagged observations. Machine learning methods were applied to model outage occurrences and predict the probability of outages over a short-term horizon. The proposed approach was designed to identify outage patterns at both the feeder and hourly levels and to generate information that can support planning and decision-making by industrial managers and electricity distribution companies.

Findings: The results show that machine learning methods provide an effective approach for predicting outages in industrial electricity networks. Among the evaluated models, LightGBM achieved an overall accuracy of approximately 99.9% and an F1-score of 0.989, indicating that it successfully reproduced observed outage patterns with high precision. The predicted outage probabilities can also serve as early-warning indicators and decision-support tools for electricity distribution companies and industrial managers. To demonstrate the practical application of the developed model, two prediction calendars were generated. The first covered March 14–20, 2025, representing the final week of the study period, while the second covered March 21–27, 2025, representing the first week outside the study period. The results showed that no outages were predicted during March 14–20, 2025, which was consistent with the actual records provided by the electricity distribution company. For March 21–27, 2025, although the predicted probabilities were close to zero, the model still identified areas with relatively higher levels of risk. This demonstrates that the model can capture temporal patterns in outage occurrence and provide precautionary warnings even when the predicted probability is low.

Research limitations/implications: A key limitation of this study is that the analysis is based on a single year of data, covering the period from March 20, 2024, to March 20, 2025. Future studies could incorporate multi-year datasets, where available, to examine medium- and long-term patterns in electricity load and outage behavior. Using multi-year data would also improve understanding of seasonal and annual variations in outage occurrence. Additionally, future research could integrate electricity consumption and outage data with economic indicators—such as industrial value added or employment levels—to investigate the direct effects of power outages on industrial production and broader economic performance. Such analyses could provide a stronger basis for prioritization in energy management and industrial planning.

Practical implications: Short-term outage predictions can help industrial managers plan production activities more effectively. By providing outage probabilities at the feeder and hourly levels, the developed approach can function as both an early-warning system and a decision-support tool. The predictions may also assist electricity distribution companies in making informed decisions related to outage management and load reduction during periods of peak demand.

Social implications: Accurate prediction of power outages can contribute to more stable industrial operations and help reduce disruptions to production and supply chains. More broadly, the use of predictive models for outage management supports more efficient energy planning and contributes to the development of smarter electricity networks and sustainable industrial growth.

Originality/value: This study develops a machine-learning-based approach for short-term electricity outage prediction using data from industrial electricity feeders in Iran. The approach integrates electricity load, outage records, weather conditions, temporal and calendar variables, and lag-based features to predict outage occurrence at both the feeder and hourly levels. The creation of prediction calendars for the final week of the study period (March 14–20, 2025) and the subsequent week outside the study period (March 21–27, 2025) illustrates how outage probabilities can be used to provide practical warnings and support decision-making in industrial electricity management.

Keywords: Short-term prediction of power outages, Machine learning models, Planned power outages, Power outages management, Data-driven decision-making

پژوهش در مدیریت تولید و عملیات، دوره ۱۷، شماره ۲، پیاپی ۴۵، ۱۴۰۵
دریافت: ۱۴۰۴/۱۲/۰۹، بازنگری: ۱۴۰۵/۰۵/۳۰ پذیرش: ۱۴۰۵/۰۶/۲۱ ص ۱۳۷-۱۵۶
DOI: 10.22108/pom.2026.148566.1658

(مقالهٔ پژوهشی)

مدل داده‌محور پیش‌بینی و زمان‌بندی خاموشی‌های برق بخش صنعت در افق کوتاه‌مدت

فاطمه شاهمرادی[۱]؛ مجید شخصی نیائی[*۲]؛ محمدمهدی لطفی[۳]؛ داود ابوترابی زارچی[۴]

۱- کارشناسی ارشد گروه مهندسی صنایع، دانشگاه یزد، یزد، ایران، fatemehshahmoradi@stu.yazd.ac.ir
۲- دانشیار گروه مهندسی صنایع، دانشگاه یزد، یزد، ایران، m.niaei@yazd.ac.ir
۳- استاد گروه مهندسی صنایع، دانشگاه یزد، یزد، ایران، lotfi@yazd.ac.ir
۴-استادیار گروه قدرت، دانشکده مهندسی برق، دانشگاه یزد، یزد، ایران، d.abootorabi@yazd.ac.ir

چکیده: صنعت برق به‌عنوان یکی از زیربناهای حیاتی توسعهٔ صنعتی و اقتصادی کشور، همواره با چالش ناترازی میان عرضه‌وتقاضا روبه‌رو بوده است. در سال‌های اخیر، این ناترازی به‌ویژه در فصول گرم و سرد، به اعمال خاموشی‌های برنامه‌ریزی‌شده در بخش صنعت منجر شده و خسارت‌های اقتصادی چشمگیری را بر جای گذاشته است. با توجه به اهمیت مدیریت انرژی و کاهش آثار خاموشی بر صنایع، پیش‌بینی کوتاه‌مدت خاموشی‌ها، ابزار کلیدی برای بهبود برنامه‌ریزی و افزایش تاب‌آوری شبکه به شمار می‌رود. در پژوهش حاضر، با تمرکز بر داده‌های فیدرهای برق صنعتی یکی از شهرهای صنعتی ایران در سال ۱۴۰۳، مدل‌سازی مبتنی بر یادگیری ماشین توسعه داده شده است که هدف آن پیش‌بینی وقوع خاموشی‌ها در افق کوتاه‌مدت است. داده‌های استفاده‌شده شامل بار مصرفی، اطلاعات آب‌وهوایی (دما و سرعت باد)، ویژگی‌های زمانی و تقویمی و متغیرهای مشتق‌شده از نوع وقفه بودند. نتایج این پژوهش، نشان می‌دهد که به‌کارگیری روش‌های یادگیری ماشین، رویکردی مؤثر برای پیش‌بینی خاموشی‌ها در شبکه‌های صنعتی است. مدل طراحی‌شده، به‌ویژه الگوریتم تقویت گرادیانی سبک، با دستیابی به صحت کلی حدود ۹۹/۹درصد و امتیاز **F1** برابر با ۰/۹۸۹، الگوهای وقوع خاموشی را به‌خوبی بازتولید کرده است. پیش‌بینی خاموشی‌ها مشخصه‌ای هشداردهنده و پشتیبان تصمیم برای شرکت‌های توزیع برق و مدیران صنایع به کار می‌رود.

واژه‌های کلیدی: پیش‌بینی خاموشی‌ها در افق کوتاه‌مدت، مدل‌های یادگیری ماشین، خاموشی‌های برنامه‌ریزی‌شده، مدیریت خاموشی، تصمیم‌گیری مبتنی بر داده

* نویسنده مسئول، 0000-0001-8034-9955

2981-0329 / © University of Isfahan
This is an open access article under the CC-BY-NC-ND 4.0 License (https://creativecommons.org/licenses/by-nc-nd/4.0/)

استناد به این مقاله: شاهمرادی، فاطمه، شخصی نیائی، مجید، لطفی، محمدمهدی و ابوترابی زارچی، داود . (۱۴۰۵). مدل داده‌محور پیش‌بینی و زمانبندی خاموشی‌های برق بخش صنعت در افق کوتاه‌مدت. *پژوهش در مدیریت تولید و عملیات* ۱۷(۲)، ۱۳۷-۱۵۶.
doi: 10.22108/pom.2026.148566.1658

۱- مقدمه

با وجود توسعهٔ فناوری‌ها و منابع انرژی تجدیدپذیر، همچنان شاهد وقوع خاموشی‌ها[1] در سراسر دنیا هستیم. در ایران نیز به‌دلیل ناترازی بین عرضه‌وتقاضای برق، به‌خصوص در فصول اوج مصرف مانند تابستان و زمستان، احتمال وقوع خاموشی‌های برنامه‌ریزی‌شده[2] در صنایع زیاد خواهد بود. در سال‌های اخیر، بیش از ۹۰درصد خاموشی‌ها در ایران، در تابستان رخ می‌دهند (Tamjid Shabestari et al., 2022). مدیریت نادرست و غیرعلمی مصرف انرژی در شبکهٔ برق، خسارات اقتصادی و عملیاتی گسترده‌ای را به‌ویژه برای مشترکان صنعتی به بار می‌آورد؛ زیرا کوچک‌ترین اختلال در تأمین برق و رخداد خاموشی، به توقف خطوط تولید و زیان‌های مالی سنگین آنها منجر می‌شود. طبق گزارش وزارت نیرو در تیرماه سال ۱۴۰۴، مصرف برق در بخش صنعتی، 37/8درصد از کل موارد مصرف بوده است. جدول (۱)، وضعیت مصرف برق کشور را در سال جاری نشان می‌دهد. همان‌گونه که در این جدول مشخص است، بخش صنعتی و خانگی بیشترین سهم مصرف را به خود اختصاص می‌دهند.

جدول ۱- سهم مصرف برق بخش‌های مختلف (میلیون کیلووات ساعت)

Tabel 1- Electricity consumption of different sections in Iran

بخش	سال ۱۴۰۳	تا پایان تیرماه سال ۱۴۰۴	سهم مصرف (درصد)
خانگی	۱۰۸۴۴۱	۳۱۱۲۳	۳۰/۸
عمومی	۳۰۲۶۲	۸۱۶۵	۸/۱
کشاورزی	۴۵۷۲۰	۱۳۹۳۲	۱۳/۸
صنعتی	۱۲۲۵۱۱	۳۸۱۳۸	**۳۷/۸**
تجاری	۲۸۵۷۵	۸۶۴۱	۸/۶
روشنایی معابر	۴۶۴۸	۸۹۷	۰/۹
کل	۳۴۰۱۵۷	۱۰۰۸۹۶	۱۰۰

توسعهٔ مدل‌هایی برای پیش‌بینی کوتاه‌مدت خاموشی‌ها، به اطلاع‌رسانی به‌موقع و کاهش خسارت‌ها کمک می‌کند. در سال‌های اخیر، یادگیری ماشین[3] و الگوریتم‌های پیشرفته، دقت بالایی را در پیش‌بینی الگوهای پیچیده ازجمله بار مصرفی[4] و خاموشی‌ها فراهم کرده‌اند. توسعهٔ مدلی که بتواند در افق کوتاه‌مدت، احتمال خاموشی را به تفکیک ساعت و فیدر پیش‌بینی کند، به صنایع کمک می‌کند برنامهٔ تولید خود را با آن متناسب کنند و خسارات ناشی از توقف ناگهانی را به حداقل برسانند. همچنین این پیش‌بینی‌ها برای شرکت‌های توزیع برق، مبنای تصمیم‌گیری در اعمال خاموشی عادلانه و کاهش بار شبکه در ساعات اوج مصرف است. به همین دلیل، انجام این پژوهش نه‌تنها برای ارتقای بهره‌وری و کاهش زیان صنایع حیاتی است، گام مهمی در جهت مدیریت هوشمند شبکهٔ برق و توسعهٔ پایدار کشور تلقی می‌شود. هدف اصلی این پژوهش نیز، توسعهٔ مدلی کارآمد و ارزیابی آن در پیش‌بینی کوتاه‌مدت خاموشی‌های برق در یکی از شهرک‌های صنعتی کشور است که به درخواست شرکت توزیع برق مربوطه، از ذکر نام شهر و همچنین شهرک صنعتی آن، پرهیز شده است.

هدف اصلی این پژوهش، توسعه و ارزیابی مدلی کارآمد برای پیش‌بینی کوتاه‌مدت خاموشی‌های برق، در شهرکی صنعتی است تا براساس داده‌های تاریخی بار مصرفی، اطلاعات آب‌وهوایی و متغیرهای تقویمی، احتمال وقوع خاموشی را به تفکیک فیدر و ساعت برآورد کرد.

در بخش اول، به‌اهمیت موضوع خاموشی‌ها اشاره شد. ادامهٔ مقاله نیز به این صورت تنظیم شده است: در بخش دوم، مبانی نظری استفاده‌شده در این مقاله تصریح شده است، در بخش سوم روش انجام پژوهش و گام‌های اجرایی آن برشمرده شده‌اند، در بخش چهارم نمونهٔ موردی و داده‌های گردآوری شده و همچنین نتایج عددی حاصل از پیاده‌سازی مدل‌های پیشنهادی شرح داده شده‌اند، بخش پنجم به بحث دربارهٔ نتایج روی آورده و در بخش ششم، جمع‌بندی پژوهش و پیشنهادهایی برای ادامهٔ این پژوهش توصیه شده است.

۲- ارکان اصلی مقاله

در این بخش از مقاله، به‌منظور ایجاد درک عمیق‌تر از مسئلهٔ پژوهش و فراهم‌سازی مبنای علمی برای مراحل بعدی، مهم‌ترین پژوهش‌های مرتبط، مرور شده‌اند.

۲-۱ پیشینهٔ پژوهش

در این بخش، مهم‌ترین مقالات مرتبط با پژوهش حاضر شامل پژوهش‌های مرتبط با پیش‌بینی بار مصرفی، پیش‌بینی خاموشی‌های ناشی از رخداد حوادث و اعمال خاموشی‌ها به‌جهت تعمیرات و نگهداری مرور شده‌اند.

والسر و سور[۵] (۲۰۲۱) در پژوهشی با تأکید بر اهمیت پیش‌بینی کوتاه‌مدت بار مصرفی، با ترکیب دو روش درخت تصمیم[۶] و گونه‌ای از شبکه‌های عصبی[۷]، بار مصرفی سه کارخانه را با کاربری مختلف در آلمان پیش‌بینی کردند. ویژگی‌های استفاده‌شده در این پژوهش شامل بار مصرفی گذشته، داده‌های آب و هوایی و اطلاعات تقویمی بوده‌اند. نورالدین و همکاران[۸] (۲۰۲۳)، ارتقای عملکرد الگوریتم‌های یادگیری ماشین را در پیش‌بینی بار روزانه در کرکوک عراق برای برنامه‌ریزی بهتر، تخصیص منابع و راهبرد مناسب بررسی کرده‌اند. آنها از روش‌های درخت تصمیم، ماشین بردار پشتیبان، بگینگ[۹] و ادابوست[۱۰] برای پیش‌بینی استفاده کرده‌اند. نتایج نشان می‌دهد روش بگینگ بیشترین دقت را نسبت‌به دیگر مدل‌ها دارد و بیانگر پتانسیل استفاده از آن به‌عنوان ابزاری ارزشمند برای برنامه‌ریزان و سیاست‌گذاران شرکت‌های انرژی است. مجموعه داده‌های استفاده‌شده در این پژوهش مربوط به پارامترهای محیطی و هواشناسی بوده است. تیمور و اوستونل[۱۱] (۲۰۲۵) در پژوهشی پیش‌بینی کوتاه‌مدت بار الکتریکی را در یک کارخانهٔ صنعتی بزرگ در آدانا، ترکیه انجام داده‌اند. هدف آنها بهبود مدیریت انرژی، کاهش هزینه‌ها و تسهیل بهره‌گیری از منابع تجدیدپذیر بوده است. برای این منظور، عملکرد پنج روش یادگیری ماشین شامل رگرسیون خطی چندگانه، روش گروهی داده‌پردازی، شبکهٔ عصبی پرسپترون چندلایه، درخت تصمیم تقویتی گرادیان و برنامه‌ریزی بیان ژنی مقایسه شده است. داده‌های استفاده‌شده از سال ۲۰۲۰ تا ۲۰۲۳ و شامل داده‌های آب‌وهوایی (دما، رطوبت و میزان بارندگی)، داده‌های بار مصرفی و داده‌های تقویمی بوده است. نتایج نشان می‌دهند که مدل درخت تصمیم تقویتی گرادیان، بالاترین دقت و عملکرد را داشته است؛ هرچند زمان محاسبهٔ بیشتری نسبت‌به دیگر مدل‌ها دارد. واتسون و همکاران[۱۲] (۲۰۲۲)، مدل پیشرفته‌ای را برای پیش‌بینی خاموشی‌های ناشی از رویدادهای

شدید آب‌وهوایی، با هدف کمک به شرکت‌های برق در برنامه‌ریزی و تخصیص منابع قبل از وقوع طوفان ارائه دادند. در این پژوهش، داده‌های شبیه‌سازی عددی هواشناسی همراه با ویژگی‌های محیطی به کار گرفته شده است تا مجموعه داده‌ای بزرگ و چندبعدی تشکیل شود. مدل با استفاده از الگوریتم گرادیان بوستینگ درختی، آموزش یافته و درنهایت ارزیابی با روش اعتبارسنجی متقابل نشان داده است که مدل نهایی قادر است با ضریب تعیین R2=0.82، پیش‌بینی دقیقی از میزان خاموشی‌ها در مقیاس وسیع ارائه دهد. شارما و همکاران[۱۳] (۲۰۲۳)، روشی برای پیش‌بینی ساعتی خاموشی‌های ناشی از شرایط جوی در افق یک روزه ارائه و برای غلبه بر نبود تعادل داده‌ها، از رگرسیون لجستیک وزنی استفاده کرده‌اند. مدل ایشان با به‌کارگیری متغیرهای باد، بارش، دما و اعتبارسنجی متحرک، در سطح شهر و پست آزمایش شده و نسبت‌به روش‌های ساده‌تر، عملکرد دقیق‌تری در شناسایی خاموشی‌ها داشته است. هانگ و همکاران[۱۴] (۲۰۲۴)، مدلی ترکیبی برای پیش‌بینی مناطق خاموشی در زمان بلایای طبیعی ارائه کرده‌اند که با تحلیل هم‌زمان ویژگی‌های مکانی و زمانی داده‌های شدید آب‌وهوایی، دقت و سرعت پیش‌بینی را بهبود می‌دهد. علی‌محمدی و بهنامیان[۱۵] (۲۰۲۱)، مدل ریاضی برای را زمان‌بندی بهینهٔ تعمیرات پیشگیرانه در شبکهٔ توزیع برق ارائه کرده‌اند که با داده‌های واقعی در ایلام ارزیابی شده است. در این مدل با در نظر گرفتن شاخص‌هایی مانند انرژی توزیع‌نشده، تعداد مشترکان متأثر و ظرفیت تیم‌های عملیاتی، از دو روش ابتکاری برای حل مسئله استفاده شده است. نتایج این پژوهش، بیانگر کاهش درخور توجه انرژی توزیع‌نشده و بهبود قابلیت اطمینان نسبت‌به روش‌های سنتی بوده است که البته تمرکز اصلی آن بر مدیریت خاموشی‌های تعمیراتی است و حوزهٔ پیش‌بینی خاموشی‌ها را پوشش نمی‌دهد. توبو و همکاران[۱۶] (۲۰۲۲)، یک چارچوب داده‌محور را برای برنامه‌ریزی تعمیرات شبکه‌های قدرت با حضور انرژی‌های تجدیدپذیر پیشنهاد می‌کنند که با کاهش هزینهٔ محاسباتی، به تصمیم‌گیری سریع‌تر اپراتور کمک می‌کند. با استفاده از داده‌های بار، تولید، آب‌وهوا و تبادلات توان، امنیت بازه‌های تعمیر پیش‌بینی شده و در میان مدل‌های آزموده‌شده، جنگل تصادفی با دقت بالای ۹۰درصد، بهترین عملکرد را نشان داده است. این پژوهش همچنین دو شاخص جدید برای سنجش پیچیدگی زمان‌بندی تعمیرات معرفی می‌کند و نشان می‌دهد افزایش منابع تجدیدپذیر، بازه‌های امن تعمیرات را کاهش می‌دهد. البته تمرکز پژوهش مذکور بر بهینه‌سازی برنامهٔ تعمیرات است و پیش‌بینی مستقیم وقوع خاموشی‌ها را پوشش نمی‌دهد.

به‌منظور ترسیم تصویر کلی از جایگاه مطالعهٔ حاضر در بین پژوهش‌های اشاره‌شده، جدول (۲) ارائه شده است.

همان‌گونه که در جدول (۲) مشخص است، اگرچه مطالعات متعددی در زمینهٔ پیش‌بینی بار و خاموشی‌های اضطراری انجام شده است، شکاف پژوهشی مهمی در حوزهٔ پیش‌بینی خاموشی‌های ناشی از ناترازی وجود دارد و با دانش فعلی نویسندگان این مقاله، تا به امروز پژوهشی در این زمینه انجام نشده است. با توجه به اینکه کشورمان با تغییرات الگوی مصرف و اعمال سیاست‌های جدید در سال‌های اخیر روبه‌رو شده است، به توسعهٔ مدلی نیاز داریم که احتمال وقوع خاموشی‌های برنامه‌ریزی‌شده را با توجه به ویژگی‌های بار، شرایط آب‌وهوایی و تقویمی پیش‌بینی کند. پژوهش حاضر با تمرکز بر این خلأ علمی، تلاش دارد چارچوب نوینی را برای پیش‌بینی خاموشی‌های برنامه‌ریزی‌شده ارائه دهد که مکمل اطلاع‌رسانی شرکت‌های توزیع برق است و صنایع را در کاهش خسارات ناشی از این خاموشی‌ها یاری می‌دهد.

جدول ۲- مرور و مقایسهٔ پژوهش‌های موجود با پژوهش حاضر

Tabel 2- A review of recent research

نویسنده/ نویسندگان	سال انتشار	حوزهٔ پژوهش	روش/ مدل (های) استفاده‌شده
والسر و سور	۲۰۲۱	پیش‌بینی بار مصرفی	درخت تصمیم و شبکهٔ عصبی کانولوشنی
نورالدین و همکاران	۲۰۲۳	پیش‌بینی بار مصرفی	درخت تصمیم، ماشین بردارپشتیبان، بگینگ و ادابوست
تیمور و اوستونل	۲۰۲۵	پیش‌بینی بار مصرفی	رگرسیون خطی چندگانه، روش گروهی داده‌پردازی، شبکهٔ عصبی پرسپترون چندلایه، درخت تصمیم تقویتی گرادیان و برنامه‌ریزی بیان ژنی
واتسون و همکاران	۲۰۲۲	خاموشی ناشی از رخداد	گرادیان بوستینگ درختی
شارما و همکاران	۲۰۲۳	خاموشی ناشی از رخداد	رگرسیون لجستیک
هانگ و همکاران	۲۰۲۴	خاموشی ناشی از رخداد	ترکیبی از شبکهٔ عصبی کانولوشنی و حافظهٔ کوتاه‌مدت و بلندمدت
علی‌محمدی و بهنامیان	۲۰۲۱	خاموشی برنامه‌ریزی‌شده (ناشی از تعمیرات)	مدل‌سازی ریاضی زمان‌بندی ماشین‌های موازی + الگوریتم‌های ابتکاری
توبو و همکاران	۲۰۲۲	خاموشی برنامه‌ریزی‌شده (ناشی از تعمیرات)	جنگل تصادفی، ماشین بردار پشتیبان، گرادیان بوستینگ، نایو بیز
پژوهش حاضر	۲۰۲۴	خاموشی برنامه‌ریزی‌شده ناشی از ناترازی	رگرسیون لجستیک، تقویت گرادیانی شدید[۱۷] و ماشین تقویت گرادیانی سبک[۱۸]

به‌منظور پاسخگویی به شکاف‌های تحقیقاتی شناسایی‌شده، نوآوری‌های زیر در پژوهش حاضر ارائه شده‌اند:

- توسعهٔ مدل‌های مبتنی بر یادگیری ماشین برای پیش‌بینی کوتاه‌مدت خاموشی برق در سطح فیدر و مقایسهٔ نتایج هر مدل؛
- استفادهٔ هم‌زمان از داده‌های بار مصرفی، متغیرهای آب‌وهوایی و شاخص‌های تقویمی برای استخراج الگوی وقوع خاموشی؛
- ایجاد تقویم پیش‌بینی خاموشی‌های هفتگی به‌عنوان خروجی کاربردی مدل، به‌جهت تصمیم‌سازی برای صنایع و مدیران شبکه؛
- استفاده از داده‌های واقعی شبکهٔ توزیع یک شهرک صنعتی به‌جای داده‌های شبیه‌سازی‌شده.

۳- روش‌شناسی پژوهش

در این بخش، رویکرد پیشنهادی پژوهش برای پیش‌بینی کوتاه‌مدت خاموشی فیدرهای شهرکی صنعتی، تشریح می‌شود. هدف از ارائهٔ این بخش، توضیح مراحل گام‌به‌گام انجام پژوهش از جمع‌آوری داده‌ها تا پیاده‌سازی و ارزیابی مدل‌های یادگیری ماشین است. در ابتدا، فرآیند آماده‌سازی و پاک‌سازی آنها بیان می‌شود؛ سپس، روش‌های مدل‌سازی شامل رگرسیون‌لجستیک، تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک معرفی و نحوهٔ آموزش و اعتبارسنجی آنها توضیح داده می‌شود. در این راستا، مراحل زیر شرح داده شده‌اند:

- جمع‌آوری و آماده‌سازی داده‌ها[۱۹]؛
- پیش‌پردازش و پاک‌سازی داده‌ها[۲۰]؛
- مدل‌سازی؛
- اعتبارسنجی و تست مدل[۲۱]؛
- تولید تقویم پیش‌بینی خاموشی.

۳-۱ جمع‌آوری و آماده‌سازی داده‌ها

داده‌های این پژوهش شامل بار مصرفی ساعتی ۱۶ فیدر صنعتی، اطلاعات خاموشی‌های برنامه‌ریزی‌شده، داده‌های تقویم رسمی و متغیرهای هواشناسی مربوط به ابتدا تا انتهای سال ۱۴۰۳ هستند. اولین گام، یکپارچه‌سازی این منابع، در قالب یک پایگاه داده‌ی یکتاست تا هر ردیف نشان‌دهندهٔ یک ساعت از یک فیدر مشخص باشد و علاوه بر بار مصرفی، وضعیت خاموشی و ویژگی‌های تقویمی و آب‌وهوایی را نیز در بر بگیرد. برای این منظور، داده‌های بار از قالب ستونی (۲۴ ستون برای ساعات شبانه‌روز) به قالب ستونی _ بلند[۲۲] تبدیل شدند تا هر ساعت، یک رکورد مجزا داشته باشد؛ سپس اطلاعات خاموشی با استفاده از تطبیق تاریخ و بازهٔ زمانی به این رکوردها ملحق و متغیر برچسب خاموشی ساخته شده است که در صورت وقوع خاموشی، مقدار یک و در غیر این صورت، صفر می‌گیرد.

۳-۲ پیش‌پردازش و پاک‌سازی داده‌ها

در مرحلهٔ پیش‌پردازش، مقادیر غیرواقعی بار مصرفی اصلاح شدند؛ به‌طوری که مقادیر منفی ناشی از بازگشت توان و همچنین صفرهای غیر مرتبط با خاموشی، با میانگین همسایگی زمانی و میانه جایگزین شدند تا روند واقعی مصرف، حفظ شود. تنها صفرهایی بدون تغییر نگه داشته شدند که هم‌زمان با وقوع خاموشی ثبت شده بودند؛ زیرا نقش کلیدی را در برچسب‌گذاری صحیح داده دارند. پس از پاک‌سازی، مجموعه‌ای از ویژگی‌های مؤثر بر الگوهای مصرف و رخداد خاموشی ایجاد شد. این ویژگی‌ها شامل اطلاعات تقویمی (ساعت، روز هفته، ماه و فصل)، نوع روز (عادی، تعطیل و روز جمعه) و ویژگی‌های تأخیری بار و خاموشی برای مدل‌سازی وابستگی‌های کوتاه‌مدت و روزانه بودند. همچنین، ویژگی تغییرات لحظه‌ای بار به‌جهت بازتاب نوسانات ناگهانی در مصرف محاسبه شد. در ادامه، متغیرهای طبقه‌ای با روش وان‌هات اینکدینگ و داده‌های پیوسته شامل بار مصرفی، دما و سرعت باد، با استانداردسازی مقیاس‌بندی شدند تا تمامی ورودی‌ها در یک مقیاس مقایسه‌پذیر قرار گیرند. یکی از چالش‌های اصلی، نامتوازن‌بودن توزیع داده‌ها بود؛ به‌گونه‌ای که کمتر از یک درصد نمونه‌ها، شامل خاموشی بودند. برای جلوگیری از تمایل مدل به پیش‌بینی کلاس اکثریت، رویکرد ترکیبی شامل افزایش نمونه‌های اقلیت با SMOTE[۲۳] و کاهش نمونه‌های اکثریت به کار گرفته شد. این روش موجب شد مدل، توانایی تشخیص خاموشی‌های واقعی را کسب کند و عملکرد آن به شکل چشمگیری بهبود یابد. شایان ذکر است که تمام مراحل انجام‌شده مثل استاندارسازی و ساخت ویژگی‌ها، تنها بر داده‌های آموزش محاسبه شده است. در تقسیم داده‌ها نیز به‌جای روش تصادفی، تفکیک براساس ترتیب زمانی انجام شد تا از نشت اطلاعات جلوگیری و شرایط واقعی بهره‌برداری شبیه‌سازی شود. مجموعهٔ آزمون به‌گونه‌ای انتخاب شد که هم شامل دوره‌هایی دارای خاموشی و هم فاقد خاموشی باشد تا ارزیابی مدل ازنظر مثبت‌های واقعی و کاذب معتبر باشد. همچنین عملیات متوازن‌سازی، تنها برای داده‌های آموزش انجام شد تا استقلال و اعتبار نتایج تست حفظ شود؛ درنتیجه، داده‌ها پس از مجموعه‌ای از فرایندهای پاک‌سازی، مهندسی ویژگی، مقیاس‌بندی و متوازن‌سازی، در قالبی مناسب و اتکاپذیر برای مدل‌سازی یادگیری ماشین آماده شدند و بستر لازم برای یک پیش‌بینی دقیق و کاربردی فراهم شد.

۳-۳ مدل‌سازی

هدف اصلی این پژوهش، پیش‌بینی کوتاه‌مدت خاموشی با اتکا به داده‌های عملیاتی شبکهٔ توزیع است. این داده‌ها شامل ویژگی‌های زمانی، بار مصرفی لحظه‌ای و تأخیری، وضعیت پیشین خاموشی و متغیرهای اقلیمی‌اند. با توجه به نامتوازن‌بودن شدید توزیع کلاس‌ها، مدل انتخابی باید توانایی یادگیری الگوهای پیچیده و در عین حال مقاومت در برابر ناتوازنی داده‌ها را داشته باشد. برای پاسخ به این نیاز، دو سطح از مدل‌های یادگیری ماشین به کار رفت. در سطح پایه، رگرسیون لجستیک خط مبنا انتخاب شد تا امکان تفسیرپذیری و تحلیل تأثیر متغیرها فراهم شود. در سطح پیشرفته، دو مدل تقویت گرادیانی شامل تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک به کار گرفته شدند؛ این مدل‌ها با ساختار مبتنی بر درخت تصمیم و توانایی کشف الگوهای غیرخطی و آثار متقابل متغیرها، گزینهٔ مناسبی برای مسائل طبقه‌بندی نامتوازن به شمار می‌روند و قابلیت ارائهٔ احتمال پیش‌بینی‌شده را دارند که برای تحلیل ریسک و برنامه‌ریزی عملیاتی، ضروری است. داده‌ها براساس زمان به دو بخش آموزش و آزمون تقسیم شدند؛ به‌گونه‌ای که داده‌های آزمون متعلق به هفته‌های پایانی سال ۱۴۰۳ باشند. این تقسیم زمانی، باعث می‌شود مدل روی داده‌هایی آموزش ببیند که پیش از بازهٔ پیش‌بینی رخ داده‌اند و از نشت اطلاعات جلوگیری شود.

۳-۴ اعتبار سنجی و تست مدل

عملکرد مدل‌ها با معیارهای کلاسیک طبقه‌بندی شامل دقت، بازخوانی، امتیاز F1 و صحت سنجیده شده‌اند. محاسبهٔ این معیارها براساس عناصر ماتریس سردرگمی انجام شده است که تعاریف آنها به شرح زیر است:

- True Positive (TP): نمونه‌های مثبت که به‌درستی مثبت پیش‌بینی شده‌اند.
- True Negative (TN): نمونه‌های منفی که به‌درستی منفی پیش‌بینی شده‌اند.
- False Positive (FP): نمونه‌های منفی که به‌اشتباه، مثبت پیش‌بینی شده‌اند.
- False Negative (FN): نمونه‌های مثبت که به‌اشتباه، منفی پیش‌بینی شده‌اند.

معیار دقت در رابطهٔ (۱)، نسبت نمونه‌های مثبت پیش‌بینی‌شده به‌درستی، به کل نمونه‌هایی است که مدل آنها را مثبت پیش‌بینی کرده است.

$$Precision = \frac{TP}{TP + FP} \quad (۱)$$

معیار بازخوانی در رابطهٔ (۲)، نسبت نمونه‌های مثبت شناسایی‌شده به‌درستی، به کل نمونه‌های واقعاً مثبت در مجموعه داده است.

$$\mathrm{Re}call = \frac{TP}{TP + FN} \quad (۲)$$

امتیاز F1 در رابطهٔ (۳)، میانگین هارمونیک معیارهای دقت و بازخوانی است که به‌منظور ارزیابی هم‌زمان این دو معیار به کار می‌رود.

$$F1 = 2 * \frac{Precision \times Recall}{Precision + Recall} \quad (۳)$$

معیار صحت در رابطهٔ (۴)، نسبت تعداد پیش‌بینی‌های صحیح مدل به کل نمونه‌های ارزیابی را نشان می‌دهد.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (۴)$$

به‌دلیل نامتوازن‌بودن داده‌ها، مساحت زیر منحنی دقت_ بازخوانی[۲۴] معیار اصلی انتخاب مدل در نظر گرفته شده است. این معیار تمرکز بیشتری بر تشخیص کلاس مثبت خاموشی (مدل خاموشی را پیش‌بینی کرده و در واقعیت هم خاموشی رخ داده است) دارد و نسبت‌به داده‌های نامتوازن حساسیت کمتری نشان می‌دهد. علاوه بر این، مساحت زیرمنحنی نمودار مشخصهٔ عملکرد[۲۵] برای ارزیابی کلی عملکرد، محاسبه شده است. برای ارزیابی دقیق عملکرد مدل و اطمینان از بروزنیافتن بیش‌برازش، از اعتبارسنجی متقاطع [۲۶]مبتنی بر زمان استفاده شده است. در این روش، داده‌های آموزشی به چندین بازهٔ زمانی متوالی تقسیم شده‌اند و مدل در هر گام، بر بازه‌های قبلی آموزش دیده و سپس روی بازهٔ بعدی ارزیابی شده است. این رویکرد تضمین می‌کند که ترتیب زمانی داده‌ها حفظ و ارزیابی مدل براساس داده‌هایی انجام می‌شود که در آینده رخ می‌دهند. نتایج این فرآیند، شامل مقادیر میانگین و انحراف معیار معیارهای ارزیابی، نشان می‌دهد که مدل در پیش‌بینی رخدادهای جدید پایدار است و عملکرد یکنواختی در تمام بازه‌های زمانی دارد. این امر، اعتبار نتایج را افزایش می‌دهد و قابلیت تعمیم مدل را برای استفاده در شرایط عملی تقویت می‌کند. در بخش چهارم مقاله، نتایج روش اعتبارسنجی متقاطع زمانی آورده شده است.

۳-۵ تولید تقویم پیش‌بینی خاموشی

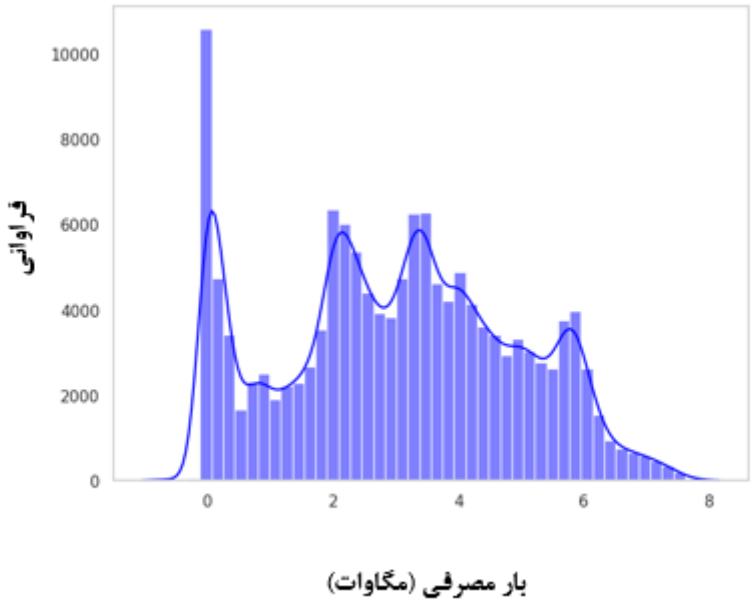
خروجی مدل‌ها برای افق‌های زمانی (هفت‌روزه، آخرین هفتهٔ اسفندماه و هفتهٔ اول فروردین‌ماه) محاسبه شده است. هدف این مرحله، ایجاد یک تقویم پیش‌بینی خاموشی است که در آن برای هر فیدر، هر روز و هر ساعت، احتمال وقوع خاموشی گزارش شود؛ سپس این احتمال‌ها براساس آستانهٔ انتخاب‌شده به برچسب‌های دودویی (خاموشی/ نبود خاموشی) تبدیل شدند. این بازنمایی ساختاریافته کمک می‌کند تا تصمیم‌گیرندگان صنعتی و مدیران بهره‌برداری به شکل بصری و ساده، وضعیت ریسک خاموشی را در هفتهٔ آینده مشاهده و برنامه‌ریزی عملیاتی خود را بهینه کنند.

۳- مطالعهٔ کاربردی و یافته‌ها

داده‌های استفاده‌شده در این پژوهش، از یکی از شرکت‌های توزیع برق کشور، به‌صورت محرمانه و با رعایت اصول حریم خصوصی دریافت شده‌اند. با توجه به تعهدات پژوهشگران دربارهٔ انتشارنیافتن منبع و جزییات داده، از ذکر جزئیات مربوط به سازمان خودداری شده است.

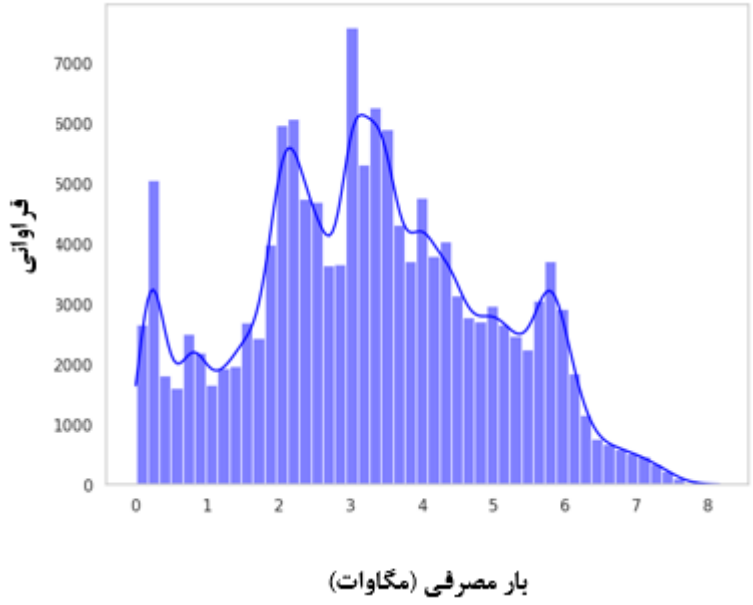
در ادامه، یافته‌های این پژوهش به شکل مختصر آورده شده‌اند. شکل‌های (۱) و (۲) و جدول (۳)، کیفیت داده‌های بار مصرفی استفاده‌شده در پژوهش، که دارای مقادیر صفر و منفی بودند، را قبل و بعد از پاک‌سازی پیش‌پردازش داده‌ها مقایسه می‌کنند. با توجه به مقایسهٔ شکل‌های (۱) و (۲)، پاک‌سازی داده‌ها امری ضروری در مدل‌سازی تلقی می‌شود.

جدول (۳) نیز بیانگر این است که مقادیر منفی و مقادیر صفر بار، که در آن لحظه خاموشی رخ نداده بود، با میانگین همسایگی جایگزین شدند و مقدارهای صفر همراه با خاموشی، ثابت باقی ماندند.



شکل ۱- هیستوگرام بار مصرفی سال ۱۴۰۳ قبل از پاک‌سازی

Fig.1- Load consumption histogram before cleaning



شکل ۲- هیستوگرام بار مصرفی سال ۱۴۰۳ بعد از پاک‌سازی

Fig.2 - Load consumption histogram after cleaning

جدول ۳- مقایسهٔ کیفیت داده‌ها قبل و بعد از پاک‌سازی

Tabel 3- Comparison of data quality before and after cleaning

شاخص	قبل از پاک‌سازی	بعد از پاک‌سازی
تعداد کل رکوردها	۱۴۰۵۴۴	۱۴۰۵۴۴
تعداد مقادیر منفی بار مصرفی	۱۸۴	۰
تعداد بارهای صفر (بدون خاموشی)	۹۶۳۸	۰
تعداد بارهای صفر (با خاموشی)	۳۷۷	۳۷۷

فرآیند مدل‌سازی داده‌های آماده‌سازی‌شده تشریح شد و نتایج اولیهٔ سه مدل منتخب شامل رگرسیون لجستیک، تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک نشان داد که هر سه مدل از توانایی بالایی در شناسایی الگوهای خاموشی برخوردارند. جدول (۴)، عملکرد سه مدل را ازنظر معیارهای ارزیابی در آستانهٔ پایه نشان می‌دهد. با توجه به جدول (۴)، هر سه مدل نتایج و عملکرد بالایی را نشان می‌دهند.

جدول ۴- مقایسهٔ عملکرد مدل‌ها

Tabel 4- Model performance comparison

مدل	Precision	Recall	F1-Score	Accuracy	ROC-AUC	PR-AUC	نرخ خطای مثبت کاذب
LR	۰/۳۹۶	۰/۹۳۶	۰/۵۵۷	۰/۹۹۱	۰/۹۹۶	۰/۷۰۷	۰/۰۰۲۴
XGBOOST	۰/۹۵۷	۰/۹۵۷	۰/۹۵۷	۰/۹۹۹	۱	۰/۹۹۷	۰/۰۰۰۰
LightGBM	۰/۹۷۹	۱	۰/۹۸۹	۰/۹۹۹	۱	۰/۹۹۶	۰/۰۰۰۰

در ابتدا عملکرد مدل‌ها با آستانهٔ تصمیم ۰,۵ ارزیابی شد؛ سپس، با توجه به ماهیت نامتوازن داده‌ها و اهمیت شناسایی رخدادهای خاموشی، اثر تغییر آستانهٔ تصمیم بر دقت و بازخوانی بررسی شد. برای تعیین یک نقطهٔ عملیاتی مناسب، آستانه‌ای انتخاب شد که در آن بازخوانی حداقل ۹۰درصد باشد و نتایج آن نشان داد آستانهٔ بهینه به ترتیب برای رگرسیون لجستیک، تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک به ۷۰، ۴۵ و ۴۰درصد تغییر می‌کند. این پراکندگی نشان می‌دهد که آستانهٔ بهینه بیشتر به نوع مدل و توزیع احتمالات پیش‌بینی‌شده از سوی آن وابسته است. به همین دلیل، برای حفظ چارچوب مقایسه‌ای یکسان و منصفانه بین مدل‌ها و همچنین نبود تفاوت چشمگیر میان دیگر معیارهای ارزیابی، در جدول (۵)، از آستانهٔ پایه استفاده و نتایج بهینه‌سازی آستانه به‌صورت تحلیل مکمل ارائه شد.

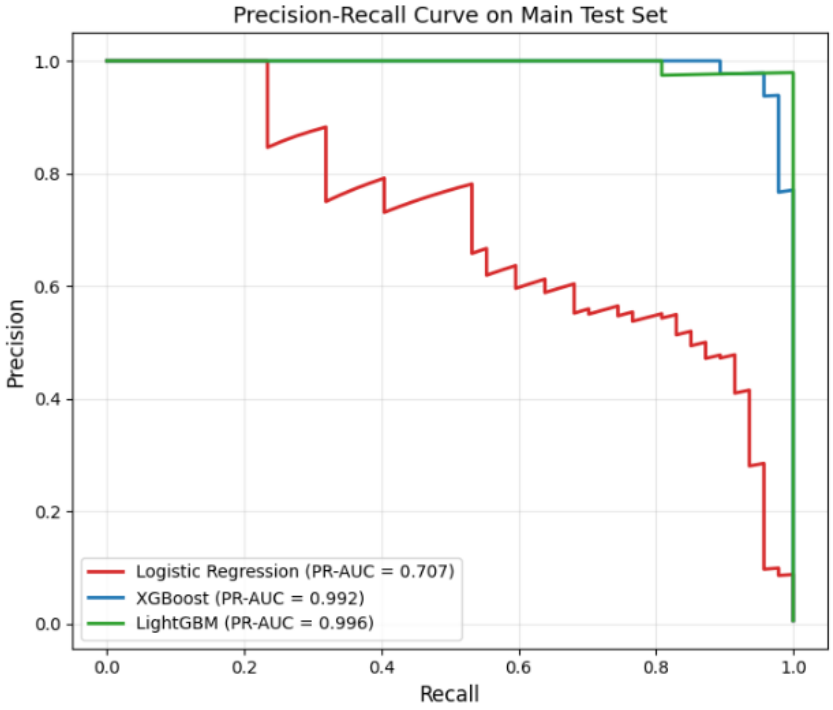
جدول ۵- مقایسهٔ عملکرد مدل‌ها بعد از بهینه‌سازی آستانه

Tabel 5- Model performance comparison after threshold optimization

مدل	آستانه بهینه	Precision	Recall	F1-Score	Accuracy	ROC-AUC	PR-AUC	FPR
LR	۰/۷۰	۰/۷۵۰	۰/۳۱۹	۰/۴۴۸	۰/۹۹۵	۰/۹۹۶	۰/۷۰۷	۰/۰۰۰۶۲
XGBOOST	۰/۴۵	۰/۹۷۸	۰/۹۵۷	۰/۹۶۸	۰/۹۹۹	۱	۰/۹۹۷	۰/۰۰۰۱۲
LightGBM	۰/۴۰	۰/۹۸۰	۱	۰/۹۹۹	۰/۹۹۹	۱	۰/۹۹۷	۰/۰۰۰۰

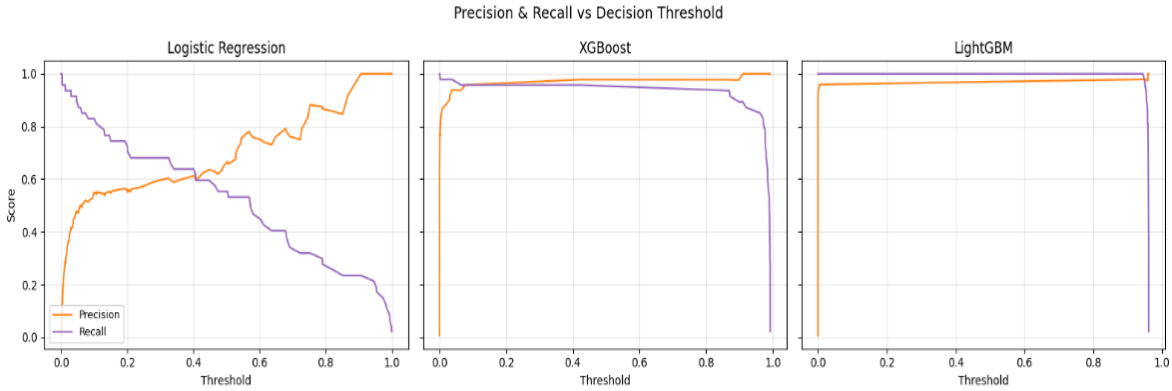
برای بررسی دقیق‌تر رفتار مدل‌ها در آستانه‌های مختلف تصمیم‌گیری، منحنی دقت_ بازخوانی برای هر سه مدل بر مجموعه تست اصلی رسم شد (شکل۳). همان‌گونه که در شکل (۳) مشاهده می‌شود، مدل رگرسیون لجستیک نسبت‌به دو مدل دیگر روند نزولی محسوس‌تری بین دقت و بازخوانی نشان می‌دهد؛ در حالی که تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک در بازهٔ وسیعی از مقادیر، بازخوانی و دقت نزدیک به یک را حفظ می‌کنند. بررسی تغییرات دقت و بازخوانی برحسب آستانهٔ تصمیم در شکل (۴) نیز، نشان می‌دهد که برای رگرسیون لجستیک، افزایش آستانه از مقادیر پایین به‌سمت آستانهٔ بهینهٔ گزارش‌شده در جدول (۵) (۰,۷۰)، با افت محسوس بازخوانی همراه است که این با نتایج ارائه‌شده در جدول (۵) کاملاً هم‌خوانی دارد. در مقابل، برای مدل‌های تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک، تغییر آستانه در بازهٔ ۰/۱ تا ۰/۸ تأثیر محدودی بر عملکرد دارد و

افت شدید تنها در آستانه‌های بسیار بالا مشاهده می‌شود؛ این رفتار نیز با آستانه‌های بهینهٔ نسبتاً پایین این دو مدل در جدول (۵) (به‌ترتیب ۰/۴۵ و ۰/۴۰) سازگار است. این تحلیل امکان می‌دهد تا صنعتگران آستانهٔ مناسب عملیاتی خود را بسته به نسبت هزینهٔ هشدار کاذب به هزینهٔ خاموشی پیش‌بینی‌نشده، انتخاب کنند.



شکل ۳- منحنی Precision-Recall برای سه مدل بر مجموعه تست اصلی

Fig.3. - Precision-Recall curve for the three models on the main test set



شکل ۴- تغییرات دقت و بازخوانی برحسب آستانهٔ تصمیم برای هر مدل

Fig.4 - Precision and Recall versus decision threshold for each model

با توجه به ماهیت زمانی داده‌ها و وابستگی مشاهدات به ترتیب زمانی، برای بررسی پایداری عملکرد مدل کاندید نهایی، یعنی ماشین تقویت گرادیانی سبک، از اعتبارسنجی متقاطع زمانی استفاده شد. در این روش، داده‌ها به شش پنجرهٔ زمانی متوالی تقسیم شدند؛ به‌گونه‌ای که در هر تکرار، مدل تنها با استفاده از داده‌های مربوط به بازه‌های زمانی گذشته آموزش داده و سپس بر بازهٔ زمانی بعدی ارزیابی شد؛ بنابراین، اطلاعات مربوط به آینده در فرآیند آموزش هر فولد وارد نمی‌شود و احتمال نشت اطلاعات ناشی از ترتیب زمانی داده‌ها کاهش می‌یابد. عملکرد مدل در هر فولد، با استفاده از معیارهای موجود در جدول (۶) ارزیابی شد. نتایج شش فولد نشان داد که ماشین تقویت

گرادیانی سبک در بازه‌های زمانی مختلف، عملکرد پذیرفتنی داشته است. با وجود تغییرات طبیعی عملکرد میان فولدها، پراکندگی معیارها محدود بود و شواهدی از ناپایداری شدید مدل در طول زمان مشاهده نشد. نتایج کامل اعتبارسنجی متقاطع زمانی مدل ماشین تقویت گرادیانی سبک در جدول (۶) ارائه شده است.

جدول ۶- نتایج اعتبارسنجی زمانی برای مدل LightGBM

Tabel 6- Time cross validation results for the LightGBM model

فولد	Precision	Recall	F1-Score	ROC-AUC	PR-AUC	FPR
۱	۰/۸۸	۰/۷۸	۰/۸۳	۰/۹۹۰۰	۰/۸۷۱۷	۰/۰۰۰۸
۲	۰/۹۳	۰/۷۱	۰/۸۰	۰/۹۹۰۶	۰/۸۸۱۵	۰/۰۰۰۸
۳	۰/۹۰	۰/۶۴	۰/۷۵	۰/۹۹۶۸	۰/۸۶۷۹	۰/۰۰۰۷
۴	۰/۸۳	۰/۸۳	۰/۸۳	۰/۹۹۷۹	۰/۸۹۸۰	۰/۰۰۰۴
۵	۰/۹۱	۰/۷۵	۰/۸۲	۰/۹۹۷۰	۰/۸۹۷۸	۰/۰۰۰۴
۶	۰/۸۵	۰/۸۰	۰/۸۲	۰/۹۸۱۴	۰/۸۴۸۵	۰/۰۰۰۵
MEAN ± STD	۰/۸۸۹±۰/۰۳۸	۰/۷۵۵±۰/۰۷۰	۰/۸۱۳±۰/۰۳۲	۰/۹۹۲±۰/۰۰۶	۰/۸۷۷±۰/۰۱۹	۰/۰۰۰۷±۰/۰۰۰۲

به‌منظور بررسی قابلیت تعمیم مدل‌ها به یک بازهٔ زمانی کاملاً مستقل، عملکرد سه مدل نهایی شامل رگرسیون لجستیک، تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک در ۲۸ روز پایانی دورهٔ مطالعه‌شده به‌صورت مستقل ارزیابی شد. در فرآیند آموزش، انتخاب ویژگی، تنظیم مدل و انتخاب آستانه دخالت داده نشد. مجموعهٔ آزمون مستقل شامل ۱۰۷۵۲ نمونه بود که ۱۳۹ مورد آن به وقوع خاموشی مربوط می‌شد. نتایج نشان داد که هر سه مدل توانایی مناسبی در تشخیص خاموشی‌های برنامه‌ریزی‌شده در بازهٔ زمانی کاملاً مستقل دارند. XGBoost با F1 برابر با ۰/۸۰۸ و PR-AUC برابر با ۰/۹۳۰، بهترین توازن کلی میان دقت و بازخوانی را نشان داد؛ در حالی که ماشین تقویت گرادیانی سبک با دقت برابر با ۰/۹۶۸ و نرخ خطای مثبت کاذب ۰/۰۰۰۳، کمترین میزان هشدار کاذب را داشت. رگرسیون لجستیک نیز با وجود بازخوانی برابر با ۰/۷۱۹، نسبت‌به دو مدل درختی، امتیاز F1 و مقدار مساحت زیر منحنی دقت_ بازخوانی پایین‌تری به دست آورد. نتایج این ارزیابی مستقل، عملکرد مناسب مدل‌ها را در مواجهه با داده‌های زمانی مشاهده‌نشده تأیید می‌کند. نتایج به‌طور کامل در جدول (۷) آورده شده است.

جدول ۷- نتایج نهایی آزمون ۲۸ روزه

Tabel 7- Final 28-day test results

مدل	Precision	Recall	F1-Score	Accuracy	ROC-AUC	PR-AUC	نرخ خطای مثبت کاذب
LR	۰/۵۴۶	۰/۷۱۹	۰/۶۲۱	۰/۹۸۸	۰/۹۶۹	۰/۷۳۱	۰/۰۰۷۸
XGBOOST	۰/۹۳۴	۰/۷۱۲	۰/۸۰۸	۰/۹۹۵	۰/۹۹۷	۰/۹۳۰	۰/۰۰۰۰۷
LightGBM	۰/۹۶۸	۰/۶۶۹	۰/۷۹۱	۰/۹۹۵	۰/۹۹۶	۰/۹۱۹۸	۰/۰۰۰۰۳

در گام نهایی این بخش، تقویم پیش‌بینی خاموشی ارائه می‌شود که هدف آن تبدیل خروجی مدل به ابزاری استفاده‌شدنی برای برنامه‌ریزی عملیاتی صنایع است. براساس نتایج جدول (۷)، مدل ماشین تقویت گرادیانی سبک، مدل برگزیده انتخاب شد؛ زیرا در مرحلهٔ ارزیابی بر داده‌های آزمون (۹۰ روز پایانی سال ۱۴۰۳)، با صحت کلی نزدیک به ۹۹/۹درصد، نرخ یادآوری ۱۰۰درصد و مساحت زیر منحنی ROC نزدیک به ۱، دارای عملکردی پایدار و

مطمئن است؛ سپس برای ارزیابی عملی خروجی مدل، پیش‌بینی خاموشی‌ها برای دو بازهٔ زمانی متوالی شامل هفتهٔ پایانی اسفند ۱۴۰۳ و هفتهٔ اول فروردین ۱۴۰۴ انجام شد که نتایج حاصل به‌صورت تجمیعی در جدول (۸) آورده شده است. براساس اطلاعات جدول (۶)، پیش‌بینی مدل با الگوی داده‌های واقعی همخوانی دارد.

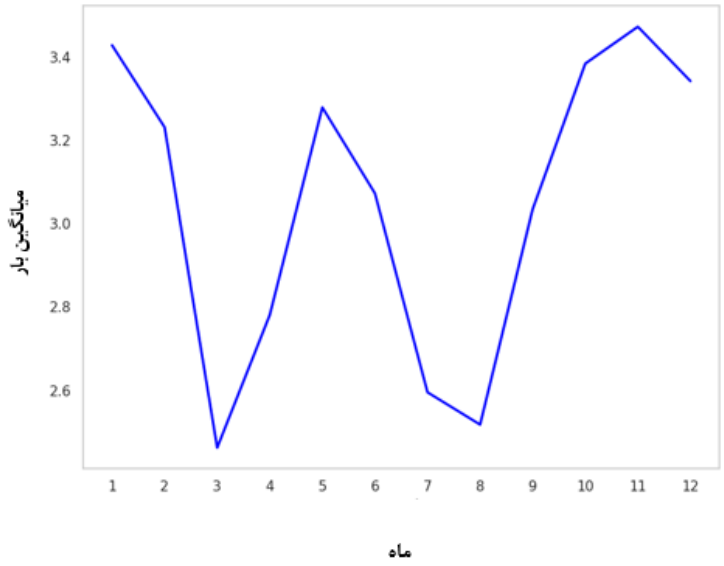
جدول ۸- خروجی تجمیعی تقویم‌های خاموشی

Tabel 8- Combined outage calender output

بازهٔ زمانی	تعداد رکوردها	مجموع ساعات پیش‌بینی‌شدهٔ خاموشی	درصد کل	توضیح
هفتهٔ آخر اسفند ۱۴۰۳	۲۶۸۸	صفر	٪۰	مدل هیچ ساعت پرریسکی پیش‌بینی نکرد. بررسی داده‌های واقعی هم تأیید می‌کند که خاموشی در این بازه رخ نداده است.
هفتهٔ اول فروردین ۱۴۰۴	۲۶۸۸	صفر	٪۰	مشابه هفتهٔ قبل، خاموشی پیش‌بینی نشده است که این با الگوی کاهش بار صنایع در بهار همخوانی دارد.

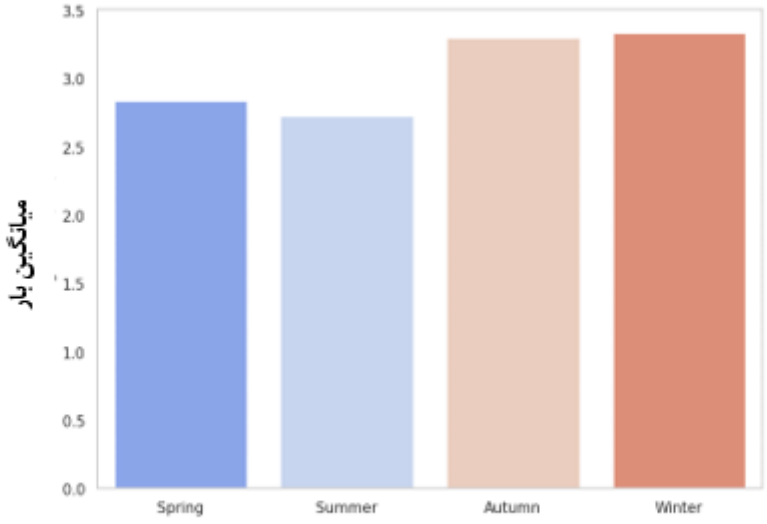
۵- بحث

پیش از آغاز فرآیند پاک‌سازی و آماده‌سازی داده‌ها، لازم بود شناخت دقیقی از ماهیت و الگوهای موجود در داده‌های خام به دست آید. این مرحله که تحلیل اکتشافی[۲۷] داده‌ها شناخته می‌شود، کمک می‌کند تا توزیع ویژگی‌ها، الگوهای زمانی، رفتار رخداد خاموشی‌ها و وجود مقادیر غیرعادی شناسایی شوند. هدف از این مرحله، ایجاد بینشی عمیق دربارهٔ رفتار سیستم و فراهم‌کردن مبنای علمی برای تصمیم‌گیری در مراحل بعدی مدل‌سازی است. الگوهای زمانی بار در مقیاس‌های مختلف بررسی شد. تحلیل سالانه و ماهانه و فصلی بار نشان داد که بیشترین مصرف در فصل‌های پاییز و زمستان رخ می‌دهد. این در حالی است که در تابستان و به‌ویژه در ماه‌های تیر و مرداد، افت مصرف محسوس است (شکل‌های ۵ و۶). این افت با الگوهای خاموشی هم‌بستگی داشت و احتمالاً ناشی از اعمال سیاست‌های مدیریت بار از سوی شرکت توزیع برق بود.



شکل ۵- نمودار بار مصرفی ماهانهٔ سال ۱۴۰۳[۲۸]

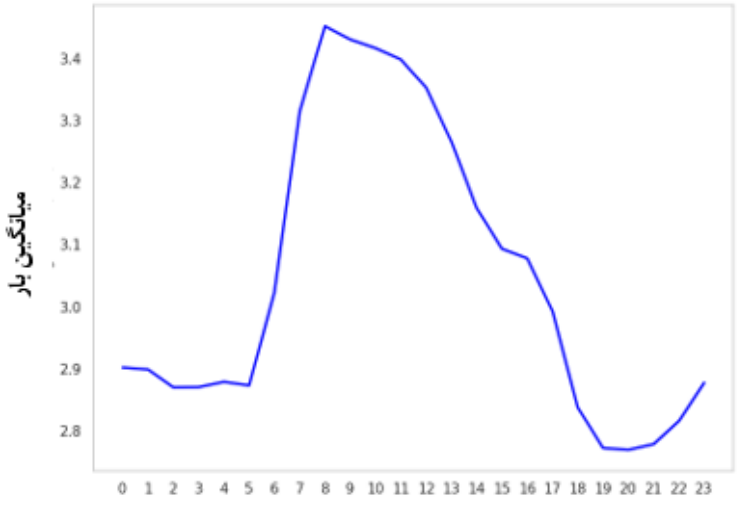
Fig.5 - Monthly load profile for 20



شکل ۶- نمودار بار مصرفی فصلی در سال ۱۴۰۳

Fig.6 - Seassonl load profile for 2024

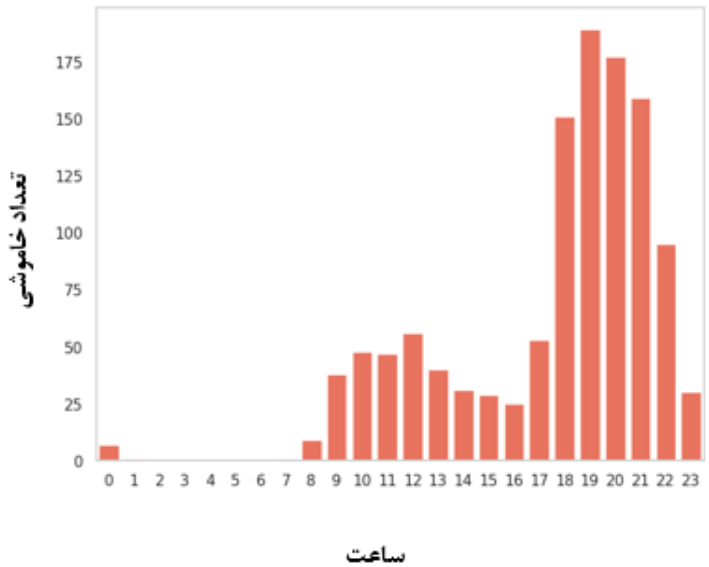
در مقیاس ساعتی نیز، مشخص شد که پیک مصرف در بازهٔ ۷ تا ۱۴ (ساعات کاری صنایع) رخ می‌دهد و بار از عصر به بعد کاهش می‌یابد (شکل ۷).



شکل ۷- نمودار میانگین سالیانهٔ بار مصرفی ساعتی در سال ۱۴۰۳

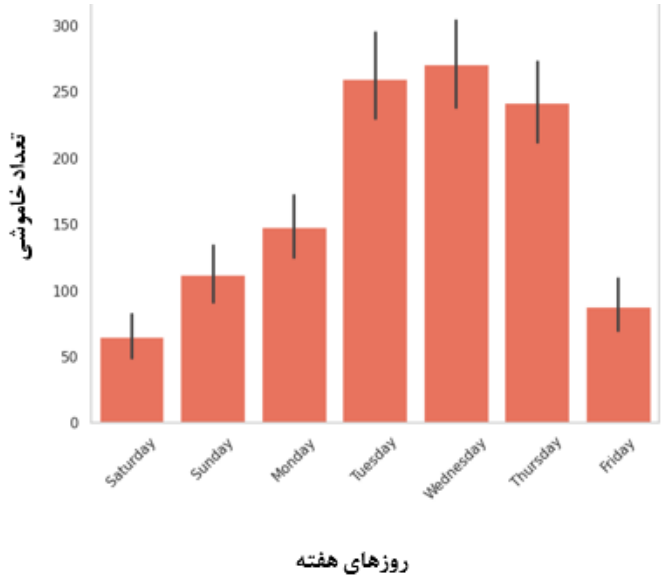
Fig.7 - Average hourly load consumption

تحلیل خاموشی‌ها نیز بخش مهمی از این مرحله بود. نمودار توزیع خاموشی‌ها برحسب ساعت و روز هفته، نشان داد که بیشترین فراوانی خاموشی در بازهٔ ۱۸ تا ۲۱ (هم‌زمان با پیک خانگی) و در روزهای میانی هفته و همچنین روزهای کاری اتفاق می‌افتد (شکل‌های ۸ و ۹). این الگو بیانگر این است که خاموشی‌ها تنها تصادفی نبوده و ابزاری برای متعادل‌سازی بار شبکه است.



شکل ۸- نمودار خاموشی‌های ساعتی در سال ۱۴۰۳

Fig.8 - Hourly outage profile



شکل ۹- نمودار خاموشی‌های روزهای هفته در سال ۱۴۰۳

Fig.9 - Daily outage profile

یکی از مهم‌ترین مراحل این پژوهش، اطمینان از کیفیت داده‌ها پیش از مدل‌سازی بود. داده‌های جمع‌آوری‌شده شامل مقادیر بار مصرفی، دما، سرعت باد، تعطیلات رسمی ایران و برچسب خاموشی‌ها در بازهٔ زمانی یک‌ساله بودند. بررسی‌های اولیه نشان داد که بخشی از داده‌های بار مصرفی، دچار مشکلاتی از قبیل مقادیر صفر غیرواقعی، مقادیر منفی ناشی از برگشت توان و ثبت‌های نویزی‌اند که دقت مدل را به‌شدت تحت تأثیر قرار می‌دهند؛ از این رو، فرآیند پاک‌سازی داده‌ها یک گام اساسی طراحی شد تا داده‌ها نمایانگر رفتار واقعی شبکه باشند. در اولین گام، کلیهٔ مقادیر بار منفی شناسایی و با میانگین متحرک ساعتی جایگزین شدند تا پیوستگی پروفایل بار حفظ شود؛ سپس مقادیر صفرِ بدون خاموشی که احتمالاً ناشی از خطای اندازه‌گیری بودند، با روش‌های آماری جایگزین شدند. این اصلاحات باعث شد نسبت مقادیر صفر کاهش یابد و تمامی مقادیر منفی حذف شود. علاوه بر این، بررسی‌های همبستگی نشان داد که ویژگی‌های اصلی نظیر بار مصرفی و روندهای فصلی، همچنان رفتار مشابهی نسبت‌به قبل

دارند و فرآیند پاک‌سازی به حذف الگوهای واقعی منجر نشده است. این مرحله تضمین کرد که ورودی‌های مدل فاقد داده‌های غیرمنطقی‌اند و مدل‌های یادگیری ماشین، الگوهای واقعی بار و خاموشی را می‌آزمونند. پس از آموزش و ارزیابی سه مدل رگرسیون لجستیک، تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک تحت شرایط پایه (آستانهٔ احتمال ۰٫۵)، نتایج نشان داد که هر سه مدل عملکرد پذیرفتنی در پیش‌بینی خاموشی‌های برنامه‌ریزی‌شده دارند. مقادیر ROC-AUC برای هر سه مدل بالاتر از ۰/۹۹ و بیانگر توانایی بالای مدل‌ها در جداسازی کلاس خاموشی و نبود آن است. همچنین مقادیر PR-AUC که برای ارزیابی کارایی مدل در داده‌های نامتوازن مناسب‌تر است، نشان دادند که مدل‌های مبتنی بر درخت (تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک) در مقایسه با مدل خطی لجستیک، عملکرد بهتری دارند و موارد مثبت را با دقت بالاتری شناسایی می‌کنند. تحلیل خطای مدل‌ها نشان داد که رگرسیون لجستیک گرچه بسیاری از موارد مثبت را شناسایی می‌کند، میزان نرخ خطای مثبت کاذب[۲۹] بالاتری در مقایسه با دو مدل دیگر دارد که به هشدارهای نادرست منجر می‌شود. در مقابل، دو مدل تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک با دقت نزدیک به ۱۰۰درصد موارد مثبت را شناسایی کردند و تقریباً هیچ هشدار نادرستی تولید نکردند (FPR ≈ 0.0). این یافته نشان‌دهندهٔ توانایی بالای مدل‌های مبتنی بر درخت در یادگیری الگوهای پیچیده و غیرخطی در داده‌های خاموشی است. درمجموع، نتایج این بخش نشان داد که مدل ماشین تقویت گرادیانی سبک بهترین عملکرد را در بین سه مدل بررسی‌شده دارد. همچنین پنج ویژگی مهم در مدل منتخب به ترتیب ساعت خاموشی، اختلاف بار مصرفی یک ساعت قبل، دما، بار مصرفی و اختلاف بار مصرفی ۲۴ ساعت قبل بوده‌اند. در تقویم خاموشی مدل، در هیچ‌یک از بازه‌ها ساعت پرریسکی پیش‌بینی نشده است. این نتیجه با داده‌های واقعی خاموشی در همین بازه هم‌خوانی دارد و نشان می‌دهد که مدل از ایجاد هشدارهای کاذب، اجتناب می‌کند. همچنین با الگوی تاریخی کاهش بار و خاموشی در آغاز فصل بهار هم سازگار است و صحت عملکرد مدل را در شرایط عملیاتی تأیید می‌کند. یافته‌های این پژوهش نشان می‌دهد که مدل‌های بوستینگ به‌ویژه ماشین تقویت گرادیانی سبک در شناسایی رخدادهای نادر خاموشی، عملکرد بهتری نسبت‌به مدل پایهٔ رگرسیون لجستیک دارند. این برتری به‌ویژه در شاخص‌های حساسیت و F1-Score نمایان است و نشان می‌دهد روابط غیرخطی و آثار کوتاه‌مدت بار و خاموشی گذشته، نقش مهمی در وقوع خاموشی‌های برنامه‌ریزی‌شده دارند. استفاده از ویژگی‌های زمانی نظیر نوع روز و الگوهای تکرارشوندهٔ مصرف نیز، به افزایش قدرت تفکیک مدل کمک و ثابت کرده است که تصمیم‌گیری‌های عملیاتی شبکه به‌شدت از رفتار مصرف‌کنندگان و الگوی تقاضا تأثیر می‌پذیرد. همچنین رویکرد ترکیبی متوازن‌سازی داده‌ها باعث شد مدل توانایی خود را در شناسایی خاموشی‌های واقعی حفظ کند و تنها به پیش‌بینی کلاس اکثریت بسنده نکند؛ بنابراین، بهره‌گیری از الگوریتم‌های بوستینگ براساس داده‌های واقعی بهره‌برداری، امکان پیش‌بینی اتکاپذیر و عملیاتی خاموشی‌ها را در افق کوتاه‌مدت فراهم می‌کند و زیرساخت مناسبی برای تصمیم‌گیری ریسک‌محور در شرکت‌های توزیع برق است. بررسی منابع نشان می‌دهد با دانش فعلی نویسندگان این مقاله، تاکنون الگوریتم‌های یادگیری ماشین برای پیش‌بینی مستقیم خاموشی‌های برنامه‌ریزی‌شده با داده‌های واقعی بهره‌برداری شرکت توزیع به کار گرفته نشده‌اند؛ بنابراین این پژوهش، خلأ موجود را پوشش می‌دهد. با توجه به نتایج به دست آمده و عملکرد مطلوب مدل پیشنهادی در پیش‌بینی کوتاه‌مدت خاموشی‌های برنامه‌ریزی‌شده، به‌کارگیری این مدل در فرآیند تصمیم‌سازی شرکت توزیع،

پیامدهای مدیریتی درخور توجهی به‌همراه دارد. بهره‌برداری از نتایج مدل در قالب یک تقویم پیش‌بینی خاموشی، امکان اطلاع‌رسانی هدفمند و زودهنگام به صنایع واقع در فیدرهای تحت محدودیت را فراهم و درنتیجه، زمینهٔ برنامه‌ریزی مناسب خطوط تولید و کاهش خسارات ناشی از توقف فعالیت‌ها را ایجاد می‌کند. همچنین، استفاده از پیش‌بینی‌ها در زمان‌بندی فعالیت‌های تعمیرات و مانورهای شبکه، از هم‌پوشانی این فعالیت‌ها با دوره‌های پرمصرف جلوگیری و مدیریت محدودیت را کارآمدتر می‌کند. به‌طور کلی، استقرار یک سامانهٔ هشدار و برنامه‌ریزی مبتنی بر پیش‌بینی، علاوه بر افزایش قابلیت اتکای شبکه و بهبود تعامل با مشترکان صنعتی، به کاهش نارضایتی ناشی از خاموشی‌های ناگهانی و افزایش اعتماد مشترکان به تصمیمات شرکت توزیع منجر می‌شود.

۶- نتیجه‌گیری

در این پژوهش، رویکردی مبتنی بر یادگیری ماشین برای پیش‌بینی کوتاه‌مدت خاموشی‌های ناشی از ناترازی در شبکهٔ توزیع صنایع یکی از شهرهای ایران ارائه شد؛ پس از پیش‌پردازش دقیق داده‌ها، مهندسی ویژگی و رفع نامتوازنی کلاس‌ها، سه مدل طبقه‌بندی شامل رگرسیون لجستیک، تقویت گرادیانی شدید و ماشین تقویت گرادیانی سبک آموزش و مقایسه شدند. نتایج نشان داد که مدل ماشین تقویت گرادیانی سبک با دستیابی به دقت ۹۹٫۹درصد و مقدار F1-Score برابر با ۰/۹۸۹ عملکردی برتر را در پیش‌بینی وقوع خاموشی صنعتی ارائه می‌کند. مهم‌ترین دستاورد تحقیق، تولید تقویم پیش‌بینی خاموشی برای افق کوتاه‌مدت است که به‌عنوان یک ابزار هشدار زودهنگام در اختیار تصمیم‌گیرندگان قرار می‌گیرد. نوآوری پژوهش در تمرکز بر خاموشی‌های برنامه‌ریزی‌شده، ناشی از محدودیت شبکه و استفاده از رویکرد داده‌محور در سطح فیدر و ساعت است؛ حوزه‌ای که تاکنون کمتر مطالعه شده است. این نوآوری نقش مهمی در ارتقای قابلیت اتکای شبکه و کاهش خسارت صنایع در شرایط ناترازی برق دارد. این پژوهش با وجود نتایج به دست آمده، با محدودیت‌هایی همراه بوده است که لازم است در تفسیر نتایج مدنظر قرار گیرند. اولین محدودیت به سطح دسترسی داده‌ها بازمی‌گردد. داده‌های مصرف برق تنها در سطح فیدر و به‌صورت تجمیعی موجود بودند و امکان تفکیک آنها در سطح واحدهای صنعتی وجود نداشت؛ درنتیجه، متغیرهای مهمی مانند نوع صنعت، ظرفیت تولید، تعداد شیفت‌های کاری و الگوهای مصرف اختصاصی کارخانه‌ها در مدل وارد نشدند. این موضوع سبب شد بخشی از تنوع واقعی رفتار مصرف صنایع در مدل‌سازی قرار نگیرد. همچنین باید توجه داشت که پیش‌بینی خاموشی حتی در افق کوتاه‌مدت، ذاتاً با عدم قطعیت همراه است. وقوع یا وقوع‌نیافتن خاموشی تنها به الگوی بار یا شرایط آب‌وهوایی وابسته نیست و تحت تأثیر تصمیمات مدیریتی، محدودیت‌های فنی لحظه‌ای و رخدادهای غیرمنتظره مانند خرابی تجهیزات نیز قرار دارد. با توجه به اینکه این عوامل در داده‌های ورودی مدل پوشش داده نشدند، نتایج مدل باید با در نظر گرفتن این محدودیت‌ها تفسیر شود.

برای پژوهش‌های آینده پیشنهاد می‌شود بازهٔ زمانی داده‌ها توسعه یابد و داده‌های عملیاتی شبکه در مدل لحاظ شود. نتایج این پژوهش نشان داد که برخی فیدرها سهم بیشتری در بار شبکه دارند. پژوهش‌های آتی تمرکز خود را بر این فیدرهای کلیدی قرار دهند و با ترکیب داده‌های مصرف انرژی با شاخص‌های اقتصادی (مانند ارزش افزوده تولید یا میزان اشتغال)، اثر مستقیم خاموشی‌ها را بر تولید و اقتصاد صنایع متصل به آنها را بررسی کنند. این تحلیل مبنای تصمیم‌گیری‌های اولویت‌بندی در مدیریت بار برای مدیران شرکت‌های توزیع قرار می‌گیرد.

References

Alimohammadi, M., & Behnamian, J. (2021). Preventive maintenance scheduling of electricity distribution network feeders to reduce undistributed energy: A case study in Iran. *Electric Power Systems Research*, *201*, 107509. https://doi.org/10.1016/j.epsr.2021.107509.

Huang, W., Zhang, W., Chen, Q., Feng, B., & Li, X. (2024). Prediction Algorithm for Power Outage Areas of Affected Customers Based on CNN-LSTM. *IEEE Access*, *12*, 15007–15015. https://doi.org/10.1109/ACCESS.2024.3355484.

Nooruldeen, O., Baker, M. R., Aleesa, A. M., Ghareeb, A., & Shaker, E. H. (2023). Strategies for predictive power: Machine learning models in city-scale load forecasting. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, *6*, 100392. https://doi.org/10.1016/j.prime.2023.100392.

Sharma, V., Hong, T., Cecchi, V., Hofmann, A., & Lee, J. Y. (2023). Forecasting weather-related power outages using weighted logistic regression. *IET Smart Grid*, *6*(5), 470–479. https://doi.org/10.1049/stg2.12109.

Tamjid Shabestari, S., Kasaeian, A., Vaziri Rad, M. A., Forootan Fard, H., Yan, W. M., & Pourfayaz, F. (2022). Techno-financial evaluation of a hybrid renewable solution for supplying the predicted power outages by machine learning methods in rural areas. *Renewable Energy*, *194*, 1303–1325. https://doi.org/10.1016/J.RENENE.2022.05.160.

Timur, O., & Üstünel, H. Y. (2025). Short-Term Electric Load Forecasting for an Industrial Plant Using Machine Learning-Based Algorithms. In *Energies*, *18*, 5. https://doi.org/10.3390/en18051144.

Toubeau, J. F., Pardoen, L., Hubert, L., Marenne, N., Sprooten, J., De Grève, Z., & Vallée, F. (2022). Machine learning-assisted outage planning for maintenance activities in power systems with renewables. *Energy*, *238*, 121993. https://doi.org/10.1016/J.ENERGY.2021.121993.

Walser, T., & Sauer, A. (2021). Typical load profile-supported convolutional neural network for short-term load forecasting in the industrial sector. *Energy and AI*, *5*, 100104. https://doi.org/10.1016/j.egyai.2021.100104.

Watson, P. L., Spaulding, A., Koukoula, M., & Anagnostou, E. (2022). Improved quantitative prediction of power outages caused by extreme weather events. *Weather and Climate Extremes*, *37*, 100487. https://doi.org/10.1016/j.wace.2022.100487.

[1] Power Outage
[2] Planned Power Outage
[3] Machine learning
[4] Power Load
[5] Walser & Sauer
[6] Decision Tree
[7] Convolutional Neural Network (CNN)
[8] Nooruldeen et al.
[9] Bagging
[10] AdaBoost
[11] Timur & Üstünel
[12] Watson et al.
[13] Sharma et al.
[14] Huang et al.
[15] Alimohammadi & Behnamian
[16] Toubeau et al.
[17] Extreme Gradient Boosting (XGBOOST)
[18] LightGBM
[19] Data collection and preparation
[20] Data preprocessing and cleaning
[21] Model validation and testing
[22] Long format
[23] Synthetic Minority Over-sampling Technique
[24] Area Under the Precision Recall Curve (PR-AUC)
[25] Receiver Operating Characteristic Curve (ROC)
[26] Cross Validation
[27] Exploratory Data Analysis (EDA)
[۲۸] ماه‌های سال به فرمت میلادی‌اند.
[29] False Positive Rate (FPR)