System Engineering and Productivity/ 2026/ Vol.6/ No.3/ 163-188

System Engineering and Productivity

DOI: 10.22034/sep.2026.2076246.1420

Developing a Deep Learning-based Power Outage Predictive Model to Improve Resilience of Power Systems

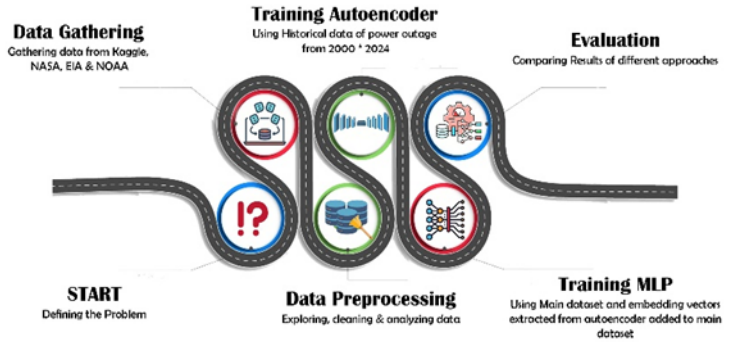
Reihane Montazeri Najafabadi [1], **Mohammadreza Shams** [2,*]

[1] M.Sc. Student, Department of Data Science, Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran
[2] Assistant Professor, Department of Computer Engineering, Faculty of Shahreza Campus, University of Isfahan, Isfahan, Iran

HIGHLIGHTS

- A transfer learning framework was proposed for power outage prediction in Maryland, USA, combining an AutoEncoder-based feature extractor with a residual MLP classifier.
- This approach eliminate dependence on geographical features such as latitude and longitude, enabling efficient generalization to unseen regions.
- Compared to traditional models, proposed model achieved higher accuracy and overall performance across different tests with other predictive approaches.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Article Type: Research paper
Received: 1 November 2025
Revised: 22 December 2025
Accepted: 1 January 2026
Available online: 1 January 2026
*Correspondence:
m.r.shams@shr.ui.ac.ir
How to cite this article:
Najafabadi, R.M., Shams, M. (2026). Developing a deep learning-based power outage predictive model to improve resilience of power systems. *System Engineering and Productivity*, 6 (3), 163-188.

Keywords:
Power Outage
Resilience of Power Systems
Deep Learning
Transfer Learning
Autoencoder

ABSTRACT

Unplanned power outages disrupt grid stability and increase operational costs, posing a major threat to power system efficiency. In this paper, we propose a robust outage prediction model that combines an Autoencoder-based feature extractor with a residual multi-layer perceptron (MLP) classifier. The novelty of our approach lies in its ability to maintain high predictive performance while eliminating reliance on geographic features such as latitude and longitude—commonly required by traditional models. We first train the Autoencoder on a rich, unlabeled dataset of weather and energy demand data collected over two decades (2000–2024) across Maryland, USA. The learned latent representations are then used to augment a supervised classification model trained on labeled outage data. Our final model achieves an F1-score of 81% even without location-based features, compared to 90% when using all features. This generalizability enables the deployment of predictive tools in previously unseen regions, directly enhancing grid flexibility, reliability, and system efficiency.

2538-2195/© 2026 The Authors. Published by University of Eyvanekey. This is an open-access article under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license.

1. Introduction

Reliable electricity supply is an essential component of modern infrastructure, yet power systems are increasingly challenged by extreme and unpredictable weather conditions. Events such as heatwaves, winter storms, strong winds, and heavy precipitation are becoming more frequent due to climate change, placing additional stress on distribution networks. These disruptions have resulted in widespread outages in regions across North America, Europe, and Asia, emphasizing the need for predictive tools that can identify outage risks before they escalate. For example, major storms in the northeastern United States during 2018 left over 2.7 million customers without power. Similar events have occurred globally, including extreme heatwaves in South Asia and Europe (Coleman, 2022).

Despite substantial research on climate effects and power system operations, many outage prediction models depend heavily on geographic variables such as latitude, longitude, or region identifiers. While these features improve localized performance, they limit model generalizability and restrict the ability to apply trained models to new areas where precise spatial data are unavailable or confidential. Additionally, most previous studies rely only on labeled outage datasets, which often cover short time spans and fail to capture the broader environmental variations that shape long-term outage behavior (Satterlee, 2024; Ghasemkhani et al., 2024)

This study addresses both challenges by developing a predictive framework that does not require geographic features and incorporates long-term environmental behavior through representation learning. The model combines a short-term labeled outage dataset with a 24-year unlabeled climate-and-demand dataset, enabling it to learn from both immediate operating conditions and historical environmental patterns. The goal is to identify low-, medium-, and high-risk outage events with strong performance across all classes, including the rare severe events that are often the most critical for grid operators.

2. Problem Definition

Two main limitations in current outage prediction studies motivate this work. First, the strong reliance on geographic variables introduces model bias. Feature-importance analyses show that spatial identifiers often dominate predictions, which means the model learns location-specific structures rather than general relationships between weather, electricity demand, and outages. When transferred to new geographic contexts, such models perform poorly. Second, although long-term climate data contain valuable seasonal and interannual patterns, they are rarely incorporated into outage prediction models due to the absence of outage labels. Ignoring this information limits the model's ability to recognize structural environmental patterns that influence the probability of outages.

To overcome these issues, this study develops a location-independent predictive model that integrates long-term climate patterns through an AutoEncoder and uses an enhanced MLP classifier to analyze enriched features. This approach allows the model to use both high-frequency labeled data and long-duration unlabeled data to improve generalization.

3. Data Description and Preprocessing

Two datasets were used in the study. Dataset A, the labeled outage dataset, consists of over 67,000 hourly observations including weather variables, electricity demand, and outage counts. The outage variable shows significant skewness, with most observations corresponding to low outage levels. To address this, outage counts were categorized into three risk levels: low, medium, and high. Dataset B, the unlabeled climate dataset, contains approximately 11,900 samples spanning 24 years. It includes historical records for temperature, precipitation, wind, pressure, snowfall, and electricity demand. Although it lacks outage labels, this dataset provides the long-term environmental variability needed for effective representation learning. Preprocessing steps included KNN imputation, removal of duplicates, standard normalization, and deliberate exclusion of geographic features. Dataset A was split into training (70%), validation (15%), and test (15%) sets. Outliers—particularly in demand and outage variables—were retained because they represent real extremes that influence model learning.

4. Methodology

The modeling framework consists of two interconnected stages: a representation-learning stage performed using an AutoEncoder, followed by a supervised classification stage using an enhanced multilayer neural network.

4.1. Stage One

Representation Learning with AutoEncoder: An AutoEncoder is trained on Dataset B to capture long-term environmental patterns. The encoder compresses each sample into an 8-dimensional latent vector capturing interactions among weather variables and electricity demand. After training, these latent representations are generated for Dataset A and append to its original features. This provides the classifier with both short-term measurements and deeper environmental context.

4.2. Stage Two

Classification Using Enhanced MLP: A multilayer neural network is constructed with residual connections to stabilize training and allow deeper

architectures without gradient degradation. A feature-weighting mechanism is included to help the classifier emphasize the most relevant variables under different conditions. Weighted cross-entropy loss ensures that the rare, high-risk outage class received sufficient emphasis during training. Regularization via dropout and batch normalization prevented overfitting. The model is evaluated using 5-fold cross-validation and compared with TabNet, ensemble models, and CNN-based architectures.

5. Results & Discussion

The experimental results demonstrate that integrating long-term environmental representations significantly improves prediction performance, especially for high-risk outages. The proposed AE-enhanced MLP achieves the strongest performance among all models. The results show that it is possible to develop a location-independent outage prediction model that maintains strong performance across all risk levels. Even without geographic features, the classifier is able to capture meaningful relationships between weather variables, electricity demand, and outage probability. The integration of long-term environmental representations leads to improved generalization and stability, particularly for identifying severe outage conditions. The performance of the proposed model relative to TabNet and CNN-based approaches highlights the importance of combining environmental context with carefully designed deep learning architectures. The improvements in the high-risk class are particularly important for operational decision-making because these events often correspond to dangerous or disruptive conditions requiring early intervention.

6. Conclusion

This study presents a deep-learning–based outage prediction framework that integrates long-term climate behavior and short-term operational data to predict outage risk levels. The model performs well without relying on geographic variables, making it transferable across regions and useful in scenarios where spatial data are limited. By combining an AutoEncoder for representation learning with an enhanced MLP classifier, the framework achieves strong performance across all classes, including rare high-risk outages. Future work may extend this approach by incorporating satellite data, modeling compound climate events, or predicting outage duration in addition to risk level. As climate patterns continue to evolve, predictive models that integrate long-term environmental information will play an increasingly vital role in grid resilience and operational planning.

Funding

This research received no external funding.

Author contributions

All authors have had equal roles and contributions to the article.

Conflicts of interest

There are no conflicts of interest associated with this research.

Acknowledgments

We are grateful to all colleagues who provided insights and expertise that greatly assisted this research. We also thank the anonymous reviewers for their valuable suggestions to improve the paper.

References

Coleman, J. (2022). Climate change made South Asian heatwave 30 times more likely. *Nature*. https://doi.org/10.1038/d41586-022-01444-1

Ghasemkhani, B., Kut, R. A., Yilmaz, R., Birant, D., Arıkök, Y. A., Güzelyol, T. E., & Kut, T. (2024). Machine learning model development to predict power outage duration (POD): A case study for electric utilities. *Sensors, 24*(13), 4313. https://doi.org/10.3390/s24134313

Satterlee, K. (2024). Combining big data and machine learning to predict power outages and help consumers prepare. *Texas A&M University News*.

مهندسی سیستم و بهره‌وری/ سال ۱۴۰۵/ دوره ۶/ شماره ۳/ صفحه ۱۶۳-۱۸۸

نشریه مهندسی سیستم و بهره‌وری

DOI: 10.22034/sep.2026.2076246.1420

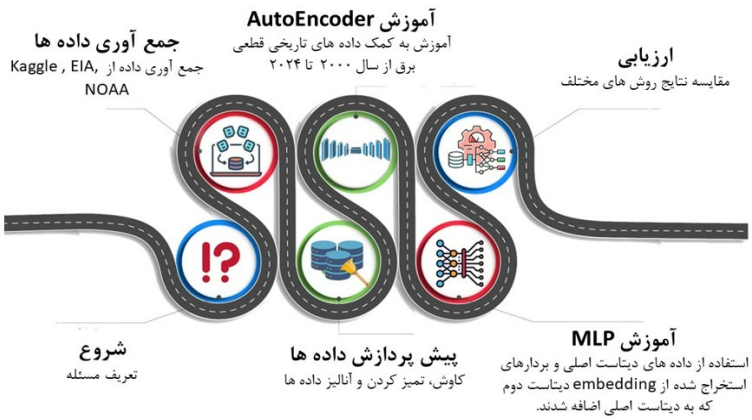
توسعه مدل پیش‌بینی قطعی برق مبتنی بر یادگیری عمیق برای بهبود تاب‌آوری در سیستم‌های قدرت

ریحانه منتظری نجف‌آبادی[۱]، محمدرضا شمس[۲]،*[*]

[۱] دانشجوی کارشناسی ارشد، گروه علوم داده، دانشکده مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان، ایران

[۲] استادیار، گروه مهندسی کامپیوتر، مرکز آموزش عالی شهرضا، دانشگاه اصفهان، اصفهان، ایران

چکیده گرافیکی



برجسته‌ها

- رویکرد یادگیری انتقالی برای پیش‌بینی قطعی برق در سیستم‌های قدرت پیشنهاد شد که ترکیبی از یک استخراج‌کننده ویژگی مبتنی بر خودرمزگذار و یک طبقه‌بند MLP با ساختار باقی‌مانده است.
- وابستگی به ویژگی‌های جغرافیایی مانند عرض و طول جغرافیایی برای افزایش قابلیت تعمیم مدل حذف گردید.
- مدل پیشنهادی دقت و عملکرد کلی بالاتری نسبت به روش‌های مرسوم از خود نشان داده است.

چکیده

قطعی‌های ناگهانی برق، نه‌تنها پایداری شبکه را تحت تأثیر قرار می‌دهند، بلکه موجب افزایش چشمگیر هزینه‌های عملیاتی نیز می‌شوند؛ مسئله‌ای که چالشی اساسی برای بهره‌وری سیستم‌های برقی به شمار می‌رود. در این پژوهش، مدلی کارآمد مبتنی بر یادگیری عمیق برای پیش‌بینی قطعی برق ارائه می‌شود که تلفیقی از یک ویژگی‌یاب مبتنی بر خودرمزگذار و یک طبقه‌بند چندلایه با ساختار باقی‌مانده است. آنچه این مدل را از روش‌های مرسوم متمایز می‌سازد، توانایی آن در حفظ دقت بالای پیش‌بینی بدون اتکا به ویژگی‌های مکانی نظیر عرض و طول جغرافیایی است، عواملی که اغلب در مدل‌های سنتی نقش کلیدی داشتند. در مرحله نخست روش پیشنهادی، یک خودرمزگذار بر روی مجموعه داده‌ای غنی، فاقد برچسب، شامل اطلاعات آب‌وهوایی و تقاضای انرژی در بازه زمانی بیست‌وچهارساله (۲۰۰۰ تا ۲۰۲۴) از ایالت مریلند ایالات‌متحده آموزش داده می‌شود. نمایش‌های نهفته به‌دست‌آمده از این مدل، به‌عنوان ویژگی‌های استخراج‌شده، به مدل طبقه‌بندی آموزش‌دیده بر داده‌های برچسب‌دار قطعی‌های برق افزوده می‌شوند. مدل نهایی، حتی بدون استفاده از داده‌های مکانی، موفق به کسب امتیاز F1 معادل ۸۱٪ شده است؛ درحالی‌که عملکرد آن با بهره‌گیری از تمامی ویژگی‌ها (شامل مختصات جغرافیایی) به ۹۰٪ می‌رسد. این سطح از قابلیت تعمیم‌پذیری، امکان استقرار مدل در مناطق فاقد داده‌های مکانی را فراهم کرده و درنهایت، به ارتقاء انعطاف‌پذیری، قابلیت اطمینان و بهره‌وری سیستم قدرت منجر می‌شود.

مشخصات مقاله

تاریخچه مقاله:

نوع مقاله: پژوهشی

دریافت: ۱۴۰۴/۰۸/۱۰

بازنگری: ۱۴۰۴/۱۰/۰۱

پذیرش: ۱۴۰۴/۱۰/۱۱

ارائه برخط: ۱۴۰۴/۱۰/۱۱

*نویسنده مسئول:

m.r.shams@shr.ui.ac.ir

کلیدواژه‌ها:

قطعی برق

پایداری سیستم‌های قدرت

یادگیری عمیق

یادگیری انتقالی

خودرمزگذار

©۲۰۲۶ نویسندگان. منتشرشده توسط دانشگاه ایوان‌کی. این یک مقاله با دسترسی آزاد تحت شرایط و ضوابط مجوز Creative Commons Attribution (CC BY 4.0) است.

۱- مقدمه

جامعه مدرن به شبکه‌های برق به‌عنوان یک زیرساخت حیاتی وابسته است تا بتواند انرژی مداوم موردنیاز مصرف‌کنندگان را تأمین کند. بااین‌حال، با توجه به افزایش شدت و فراوانی رویدادهای اقلیمی شدید نظیر موج‌های گرما، طوفان‌ها، طوفان‌های زمستانی، خشک‌سالی و بارش‌های سنگین، پایداری این شبکه‌ها به‌طور فزاینده‌ای به چالش کشیده می‌شود. این پدیده‌ها، ریسک‌های اقتصادی-اجتماعی مهمی را به همراه دارند که می‌تواند منجر به اختلالات گسترده و هزینه‌های هنگفت شود. به‌عنوان‌مثال، در مارس ۲۰۱۸، زمانی که شمال شرقی ایالات‌متحده در تنها ۱۴ روز توسط سه طوفان زمستانی مورد اصابت قرار گرفت، قطعی‌های برق در سراسر منطقه نیوانگلند بیش از ۲,۷۵۵,۰۰۰ مشتری را تحت تأثیر قرار داد و زیان‌های اقتصادی کلی ۴ میلیارد دلار، شامل ۳ میلیارد دلار زیان بیمه‌شده، به بار آورد.

رویدادهای اقلیمی اخیر ابعاد جهانی این چالش را برجسته ساخته‌اند. در سال ۲۰۲۲، امواج گرمایی مرگبار هند را درگیر کرد، درحالی‌که کشورهای اروپایی مانند بریتانیا و ایتالیا نیز شدت و مدت رو به رشد امواج گرما را تجربه کردند (Falabretti et al., 2020; Atrigna et al., 2023). در جنوب آسیا، گزارش شده است که احتمال وقوع موج گرما به دلیل تغییرات آب و هوایی ۳۰ برابر بیشتر شده است (Coleman, 2022). هم‌زمان، چین امواج گرمای رکوردشکن و طولانی‌مدتی را در سال ۲۰۲۲ تحمل کرد. این واقعیت در حالی است که تأثیر مستقیم گرما و باران سنگین بر عملکرد فیزیکی شبکه برق، که عامل اصلی بخش قابل‌توجهی از قطعی‌هاست (۴۴ تا ۷۸ درصد قطعی‌های گزارش‌شده در ایالات‌متحده مربوط به آب‌وهواست) (Campbell & Lowry, 2012)، کمتر موردتوجه قرار گرفته است. در ایران، در بسیاری از مناطق، به‌ویژه در مناطق دورافتاده و روستایی، دسترسی به انرژی الکتریکی و سایر منابع انرژی به‌طور مکرر با مشکلاتی مواجه است. این عدم دسترسی می‌تواند تأثیرات جدی بر روند ساخت‌وساز و همچنین کیفیت زندگی ساکنان داشته باشد (Rahbaripour et al., 2025; Razzaghy et al., 2024). برق نه‌تنها زیرساخت حیاتی برای توسعه کشور محسوب می‌شود، بلکه ابزاری راهبردی برای همکاری منطقه‌ای و امنیت انرژی آینده به شمار می‌آید (Avazpour et al., 2025). این خلأ در درک و مدل‌سازی، لزوم پژوهش‌های عمیق‌تر را بیش‌ازپیش نمایان می‌سازد.

علی‌رغم بینش‌هایی در مورد چگونگی افزایش تقاضا برای خنک‌کننده و محدودیت عرضه انرژی (Romitti & Sue Wing, 2022)، شکافی در کمی‌سازی جامع این دو تأثیر وجود دارد. شواهد تجربی مرتبط با امواج گرما و قطعی برق مهم است، با توجه به اینکه قطعی برق ناشی از آب‌وهوا پیامدهای گسترده‌ای دارد، به‌ویژه برای جوامع آسیب‌پذیری که مستعد بیماری‌های مرتبط با گرما هستند (Stone et al., 2021; Stone et al., 2023). مطالعه ما با ارائه یک ارزیابی تجربی از قابلیت اطمینان عرضه برق در ایالت مریلند آمریکا، با استفاده از داده‌های قطعی برق و اطلاعات آب و هوایی و اقلیمی این شکاف را پر می‌کند. چنین مطالعاتی در طول رویدادهای آب و هوایی شدید مکرر یا طولانی‌مدت به‌ویژه حیاتی می‌شوند. درنتیجه، ارزیابی دقیق تاب‌آوری شبکه برق برای تخمین خسارات ناشی از آب‌وهوای شدید، انجام واکنش کوتاه‌مدت به فاجعه برای کاهش زیان‌ها، برنامه‌ریزی برای افزایش تاب‌آوری بلندمدت، و تدوین سیاست انرژی از اهمیت بالایی برخوردار است.

بدین منظور، تحقیق پیش رو سعی در پیش‌بینی احتمال قطعی برق با توجه به رویدادهای آب و هوایی داشته و باهدف ساخت یک مدل پیش‌بینی تعمیم‌پذیر، توسعه داده‌شده است. با توجه به شرایط اقلیمی و میزان نیاز به انرژی و صرف‌نظر از موقعیت مکانی، مدل پیشنهادی می‌تواند با دقت ۸۱٪ میزان قطعی را به‌درستی در سه دسته احتمال قطعی کم، متوسط و بالا پیش‌بینی کند. در این پژوهش یک چارچوب دومرحله‌ای مبتنی بر یادگیری عمیق پیشنهادشده که باهدف حذف وابستگی مدل به موقعیت جغرافیایی و ارتقاء توان تعمیم‌پذیری آن طراحی‌شده است. این چارچوب بر پایه ترکیب دو مجموعه داده با ساختار و ماهیت متفاوت بناشده است. مجموعه نخست شامل بیش از ۶۷۰۰۰ رکورد برچسب‌خورده از یک بازه زمانی کوتاه است که انواع خاموشی‌های ثبت‌شده را شامل می‌شود و به‌عنوان مجموعه هدف برای آموزش مدل نهایی استفاده شده است. در مقابل، مجموعه دوم شامل حدود ۱۱۹۰۰ نمونه

بدون برچسب از داده‌های آب‌وهوایی و تقاضای انرژی در یک بازه بلندمدت ۲۴ ساله است که علی‌رغم نبود برچسب، ازنظر تنوع زمانی، فصلی و محیطی بسیار غنی است.

در گام نخست، از یک خودرمزگذار برای استخراج ویژگی‌های سطح بالا و تعمیم‌پذیر از مجموعه بلندمدت استفاده شد. خودرمزگذار به‌عنوان یک معماری غیرنظارتی، داده‌ها را بدون نیاز به برچسب‌زنی، به فضای میان‌بعدی فشرده‌ای تبدیل می‌کند که مهم‌ترین اطلاعات را حفظ و نویزها را حذف می‌کند. این بردارهای فشرده‌شده که حاصل آموزش خودرمزگذار بر روی داده‌های نرمال‌شده و مشترک میان دو مجموعه بودند، به نمونه‌های برچسب‌خورده در مجموعه هدف الصاق شدند و درنتیجه، مدل نهایی نه‌تنها به اطلاعات لحظه‌ای، بلکه به‌نوعی حافظه تاریخی از رفتار سیستم طی دو دهه گذشته نیز دسترسی پیدا کرد.

در مرحله دوم، معماری نهایی مدل به‌صورت یک شبکه عصبی مبتنی بر MLP طراحی شد که به کمک مکانیزم توجه (Attention) و بلوک‌های باقیمانده (Residual) تقویت شده است. این طراحی باعث شد مدل بتواند در میان حجم زیادی از ویژگی‌های عددی، به‌صورت تطبیقی بر مهم‌ترین آن‌ها تمرکز کند، از ناپدید شدن گرادیان‌ها در لایه‌های عمیق جلوگیری شود و از بیش‌برازش در مراحل یادگیری پیشگیری گردد.

نتایج تجربی نشان دادند که این رویکرد ترکیبی توانست عملکرد بسیار قابل قبولی را حتی بدون استفاده از ویژگی‌های مکانی به دست آورد. مدل نهایی موفق شد نتایج خوبی مانند f1-score کلی معادل ۸۱ درصد را ثبت کند که نسبت به مدل پایه افزایش چشمگیری محسوب می‌شود. به‌ویژه در طبقه‌بندی خاموشی‌های شدید که معمولاً دقت پیش‌بینی در آن پایین‌تر است، نیز بهبود قابل‌توجهی حاصل شد.

این دستاورد نشان می‌دهد که انتقال دانش از داده‌های غنی ولی بدون برچسب، در قالب بردارهای فشرده‌شده حاصل از خودرمزگذار، می‌تواند ساختارهای پنهان و روابط پیچیده‌تری را در داده‌ها برای مدل آشکار کند. به‌بیان‌دیگر، این چارچوب پیشنهادی توانسته است با بهره‌گیری از یادگیری انتقالی و طراحی معماری عمیق، مدلی دقیق، مستقل از مکان و قابل توسعه ارائه دهد که پتانسیل بالایی برای استفاده در سامانه‌های هوشمند مدیریت انرژی دارد و می‌تواند نقش مؤثری در بهبود آمادگی، توزیع بار و پیشگیری از اختلالات ایفا کند.

ادامه مقاله به این صورت سازمان‌دهی شده است: بخش ۲ به‌مرور ادبیات و تحقیقات پیشین درزمینه تأثیرات آب‌وهوایی و مدل‌سازی قطعی‌های برق می‌پردازد و شکاف‌های تحقیقاتی حیاتی را برجسته می‌سازد. در بخش ۳ روش‌شناسی و چارچوب دومرحله‌ای مبتنی بر خودرمزگذار و MLP شرح داده می‌شود. بخش ۴ نتایج تجربی و ارزیابی عملکرد مدل را ارائه می‌دهد. سرانجام، بخش ۵ نتیجه‌گیری و مسیرهای تحقیقاتی آتی را بیان می‌کند.

۲- مرور ادبیات

در مواجهه با چالش‌های فوق، تحلیل داده‌های بزرگ و به‌کارگیری تکنیک‌های یادگیری ماشین و یادگیری عمیق به‌عنوان ابزارهایی قدرتمند برای پیش‌بینی و مدیریت قطعی برق مطرح شده‌اند. پژوهش‌های اخیر در این حوزه پیشرفت‌های چشمگیری را نشان می‌دهند. برخی مطالعات با استفاده از داده‌های گسترده نشان داده‌اند که از طریق شناسایی و تقویت نقاط حیاتی زیرساخت، می‌توان تعداد قطعی‌های برق را تا ۵۰ درصد کاهش داد؛ موضوعی که بر نقش کلیدی شناسایی و تقویت نقاط آسیب‌پذیر شبکه تأکید دارد (Zhu et al., 2025).

در پژوهشی دیگر، پتانسیل روش‌های مبتنی بر یادگیری ماشین در بهبود زمان واکنش و عملیات اضطراری شبکه برق موردبررسی قرار گرفته و نشان داده شده است که این روش‌ها می‌توانند به تصمیم‌گیری سریع‌تر و مؤثرتر در شرایط بحرانی کمک کنند (Satterlee, 2024). همچنین، رویکردی ترکیبی مبتنی بر تئوری گزینه‌های واقعی و یادگیری تقویتی عمیق برای تحلیل عوامل سازمانی، فناوری و محیطی مؤثر بر پیاده‌سازی پایدار و انعطاف‌پذیر فناوری‌های هوشمند در بنگاه‌های کوچک و متوسط تحت شرایط بحرانی ارائه شده است (Namdarzadegan & Bozorgi Amiri, 2025).

در مطالعه‌ای دیگر، با ترکیب الگوریتم XGBoost و روش انتخاب ویژگی MRMR، دقت پیش‌بینی مدت زمان قطعی برق به ۹۸/۴ درصد افزایش یافته است؛ سطحی از دقت

که برای برنامه‌ریزی منابع و اطلاع‌رسانی مؤثر به مشترکان اهمیت بالایی دارد (Ghasemkhani et al., 2024). همچنین توسعه یک مدل MLP ترکیبی که داده‌های جمعیتی، اقلیمی و شبکه‌ای را ادغام می‌کند، امکان تحلیل بلادرنگ احتمال وقوع قطعی برق و ارائه بینش‌های راهبردی متنوع برای پشتیبانی از تصمیم‌گیری اپراتورها را فراهم کرده است (Wang et al., 2024).

این رویکردها به‌ویژه با مدل‌های شرطی در یادگیری عمیق هم‌راستا هستند؛ به‌گونه‌ای که ویژگی‌های مکانی به‌عنوان یک شرط برای تعدیل پیش‌بینی‌ها عمل می‌کنند و امکان در نظر گرفتن شرایط خاص منطقه‌ای را فراهم می‌سازند، حتی در مواقعی که داده‌های دقیق مربوط به تأسیسات شبکه به‌طور مستقیم در دسترس نیستند.

علاوه بر این، مطالعات متعددی در سال‌های اخیر برای پیش‌بینی و افزایش تاب‌آوری شبکه‌های برق انجام شده است. در یکی از این مطالعات، یک مدل ترکیبی مبتنی بر هوش مصنوعی و مدل‌سازی ساختاری برای پیش‌بینی خرابی خطوط انتقال در اثر شرایط جوی شدید پیشنهادشده است که با تعریف شاخصی برای طوفان، شدت رویدادهای آب‌وهوایی و اثر آن‌ها بر شبکه را با دقت بیشتری بازتاب می‌دهد (Nyame et al., 2024).

در پژوهشی دیگر(Sharma et al., 2023)، یک مدل رگرسیون لجستیک وزن‌دار برای پیش‌بینی کوتاه‌مدت قطعی‌های برق ارائه شده است که تمرکز اصلی آن بر مدیریت مشکل عدم‌توازن داده‌ها بوده و از طریق وزن‌دهی مناسب به نمونه‌ها، بهبود معناداری در دقت پیش‌بینی‌ها حاصل شده است. همچنین، معماری عمیقی مبتنی بر CNN–LSTM برای پیش‌بینی نواحی جغرافیایی تحت تأثیر قطعی برق معرفی شده است که با ترکیب ویژگی‌های مکانی و زمانی استخراج‌شده از داده‌های هواشناسی، توانسته است دقت پیش‌بینی‌های جغرافیایی را به‌طور قابل توجهی افزایش دهد (Huang et al., 2024).

باوجود پیشرفت‌های چشمگیر، همچنان خلأهای مهمی در مطالعات موجود مشاهده می‌شود که پژوهش حاضر به دنبال پر کردن آن‌هاست. وابستگی زیاد به ویژگی‌های جغرافیایی: بسیاری از مدل‌ها برای پیش‌بینی به‌شدت به ویژگی‌های مکانی خاص وابسته هستند که در صورت عدم وجود داده‌های مکانی مشابه برای مناطق دیگر، منجر به کاهش تعمیم‌پذیری آن‌ها می‌شود.

محدودیت داده‌ها: بخش قابل‌توجهی از مطالعات پیشین بر پایه داده‌های محدود و عمدتاً در مقیاس کشوری یا شهری خاص انجام شده‌اند. این محدودیت، به دلیل وابستگی به تجهیزات و منابع محلی، موجب کاهش قابلیت تعمیم نتایج به سایر مناطق می‌شود.

تمرکز بر رویدادهای مستقل: کمتر مطالعه‌ای به مدل‌سازی و پیش‌بینی تأثیر رویدادهای ترکیبی اقلیمی (مانند موج گرما همراه با طوفان) که اثرات تشدیدکننده‌ای دارند و از داده‌های آب و هوایی قابل‌استخراج هستند، پرداخته است.

پروژه حاضر باهدف رفع این خلأها و ارائه یک چارچوب جامع برای افزایش تاب‌آوری شبکه برق در مواجهه با تغییرات اقلیمی طراحی شده است. متمایزات اصلی و اهداف کلیدی این پژوهش عبارت‌اند از:

۱) تحلیل دقیق عددی تأثیر موج گرما بر تعداد و مدت قطعی‌ها: با استفاده از داده‌های دقیق اقلیمی از منابع معتبری چون NOAA و DWD و همچنین داده‌های مصرف انرژی EIA، به تحلیل عمیق و کمی اثر موج گرما بر ابعاد مختلف قطعی‌های برق خواهیم پرداخت.

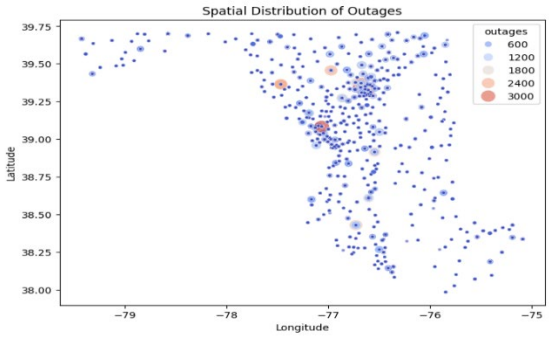
۲) طراحی مدل پیش‌بینی با دقت بالا بدون وابستگی به ویژگی‌های مکانی: با بهره‌گیری از ترکیب مدل‌های یادگیری عمیق، روش پیشنهادی قابلیت پیش‌بینی بالایی دارد و درعین‌حال، قابلیت تعمیم‌پذیری آن به مناطق جغرافیایی مختلف، بدون نیاز شدید به داده‌های مکانی خاص آن منطقه، تضمین می‌شود. این مدل همچنین عدم توازن داده‌ها را از طریق تکنیک‌هایی مانند تابع زیان آنتروپی متقابل وزن‌دار یا زیان نمایی مدیریت خواهد کرد.

این پژوهش نه‌تنها به درک عمیق‌تر پدیده قطعی‌های برق ناشی از تغییرات اقلیمی کمک می‌کند، بلکه با ارائه ابزارهای تحلیلی پیشرفته، به تقویت تاب‌آوری شبکه‌های برق یاری خواهد رساند.

۳- داده‌ها و تجزیه تحلیل توصیفی

با توجه به مطالعات انجام‌شده در ادبیات، اثرات متغیرهای مختلف در حوزه‌های آب‌وهوای اقلیمی، نیاز به انرژی و

محدوده جغرافیایی که مجموعا ۲۳ شاخص و ۶۷۳۲۱ سطر داده را برای مجموعه داده اولیه را شامل می‌شوند، بر احتمال قطعی برق تأثیرگذار هستند. داده‌های موردبررسی شامل اطلاعات آب و هوایی مربوط به موقعیت هر رخداد قطعی برق، میزان نیاز به انرژی در بازه زمانی از ۲۹ مارس تا ۱۹ آوریل می‌باشند. داده‌های مربوط به موقعیت مکانی و میزان قطعی هر منطقه از وب‌سایت Kaggle.com استخراج شده‌اند که یک مرجع معتبر برای مجموعه داده در حوزه‌های مختلف است. این مجموعه داده اولیه تهیه‌شده از وب‌سایت Kaggle شامل اطلاعاتی از موقعیت مکانی قطعی برق و زمان رخداد بود (شکل **۱**).



شکل ۱. توزیع قطعی برق در ایالت مریلند آمریکا.

Figure 1. Power outages in Maryland, USA.

در شکل **۲**، میزان پراکندگی خاموشی در ساعات مختلف روز در مجموعه داده مورداستفاده نمایش داده شده است.

همچنین به کمک اطلاعات موقعیت مکانی مانند طول و عرض جغرافیایی، دسترسی به اطلاعات آب و هوایی هر موقعیت مکانی در زمان مشخص‌شده در مجموعه داده از طریق کتابخانه Meteostat در پایتون امکان‌پذیر بود که دسترسی ساده‌ای به داده‌های تاریخی آب‌وهوا را فراهم کرده و از منابع مختلفی ازجمله NOAA و DWD اطلاعات خود را استخراج می‌کند. NOAA، اداره ملی اقیانوسی و جوی ایالات‌متحده، داده‌های اقلیمی و هواشناسی را ثبت می‌کند که در تحلیل روندهای مصرف انرژی و تأثیر تغییرات جوی بر شبکه برق مفید است.

داده‌های هواشناسی شامل اطلاعاتی مانند دمای روزانه، میزان بارش، سرعت باد، فشار هوا و میزان برف ثبت‌شده است که هرکدام تأثیر مستقیمی بر مصرف برق دارند. علاوه بر اطلاعات آب‌وهوایی، داده‌های مربوط به نیاز انرژی در مناطق مختلف نیز موردبررسی قرار گرفتند. این داده‌ها از مجموعه‌های اطلاعاتی اداره اطلاعات انرژی ایالات‌متحده (EIA) استخراج شده‌اند. EIA داده‌هایی جامع درزمینه مصرف برق در بخش‌های مختلف ازجمله مسکونی، صنعتی و تجاری ارائه می‌دهد.

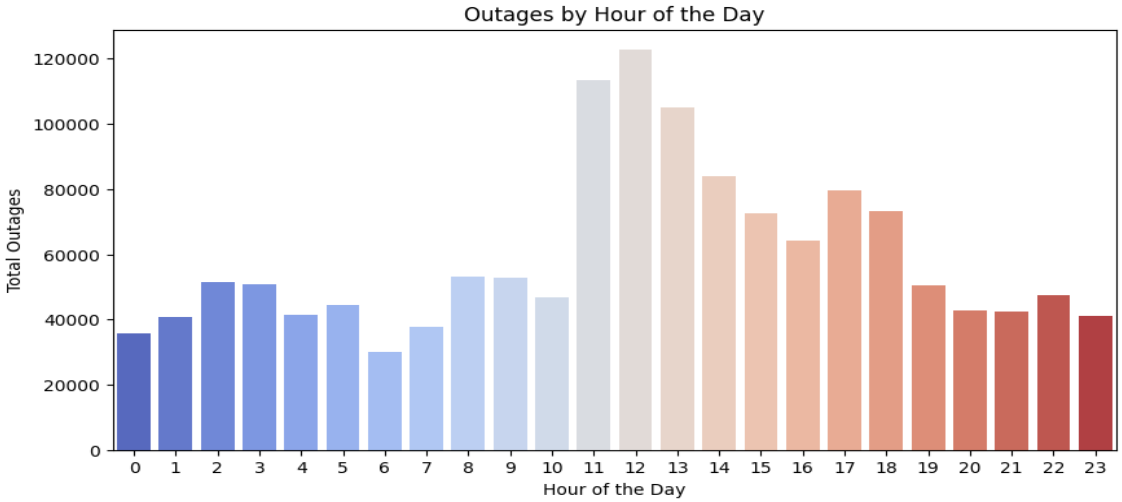
وزارت انرژی ایالات‌متحده نیز منابعی برای تحلیل میزان مصرف برق و رشد تقاضای انرژی در سطح ایالت‌های مختلف ارائه داده که می‌تواند در پیش‌بینی الگوهای قطعی برق و مدیریت بهینه منابع انرژی مؤثر باشد. با ترکیب داده‌های حاصل از EIA و NOAA، امکان بررسی دقیق‌تر ارتباط میان شرایط آب‌وهوایی و مصرف انرژی فراهم شده است. تمامی متغیرهای ذکرشده در جدول **۱** با توضیحات مربوطه قابل‌مشاهده هستند. با درک متغیرهای موردبررسی، لازم است کاوش دقیقی از داده‌های انتخاب شده داشته باشیم؛ بنابراین تحلیل داده‌های آماری برای متغیرهای منتخب در این بخش انجام شده است.

تحلیل آماری اولیه بر پایه سه شاخص اصلی انجام شده است: میانگین (average)، میانه (median) و انحراف معیار (standard deviation).

در ابتدا، متغیر outages که بیانگر تعداد قطعی‌های برق در هر ناحیه است، دارای میانگینی برابر با ۲۲ و میانه‌ای برابر با ۴ می‌باشد. اختلاف قابل‌توجه بین میانگین و میانه نشان‌دهنده وجود توزیعی به‌شدت دچار چولگی است، به‌گونه‌ای که اغلب داده‌ها مربوط به قطعی‌های کم هستند، اما نمونه‌هایی با قطعی‌های بسیار بالا نیز در دیتاست وجود دارد.

در متغیرهای دمایی (tavg, tmin, tmax) میانگین و میانه نزدیک به یکدیگر هستند که نشان‌دهنده توزیع متعادل‌تر و بدون چولگی شدید است؛ اما در مقابل، متغیر prcp (بارش) دارای میانگینی برابر با ۵/۱۱ و میانه‌ای برابر با ۰/۵ است که به‌وضوح توزیع بسیار نامتقارنی دارد. به‌عبارت‌دیگر، در بیشتر روزها بارشی رخ نداده، اما در برخی نمونه‌ها بارش شدیدی ثبت شده است. در ادامه تحلیل، شاخص‌هایی مانند چولگی (skewness)، میانه ۵۰ درصدی (IQR) و کشیدگی (range) موردبررسی قرار گرفته‌اند. در زمینه کشیدگی، بیشترین مقادیر متعلق به demand_mw و outages است. متغیر demand_mw دارای کشیدگی بیش از ۱۵ هزار است که نشان‌دهنده اختلاف زیاد بین حداقل و حداکثر مصرف برق در مناطق مختلف یا در شرایط مختلف زمانی و جوی است. این میزان تنوع می‌تواند ویژگی بسیار مهمی برای پیش‌بینی

قطعی‌های برق باشد. با بررسی ماتریس همبستگی (شکل **۳**) که میزان ارتباط خطی بین ویژگی‌های مختلف را نشان می‌دهد، مشاهده می‌شود که اغلب ویژگی‌ها همبستگی پایینی با یکدیگر دارند. این موضوع بیانگر آن است که داده مورداستفاده بیشتر ماهیت جدولی (Tabular Data) دارد؛ بدین معنا که برخلاف داده‌های ساختاریافته، ارتباط و همبستگی قوی میان ویژگی‌ها وجود ندارد.



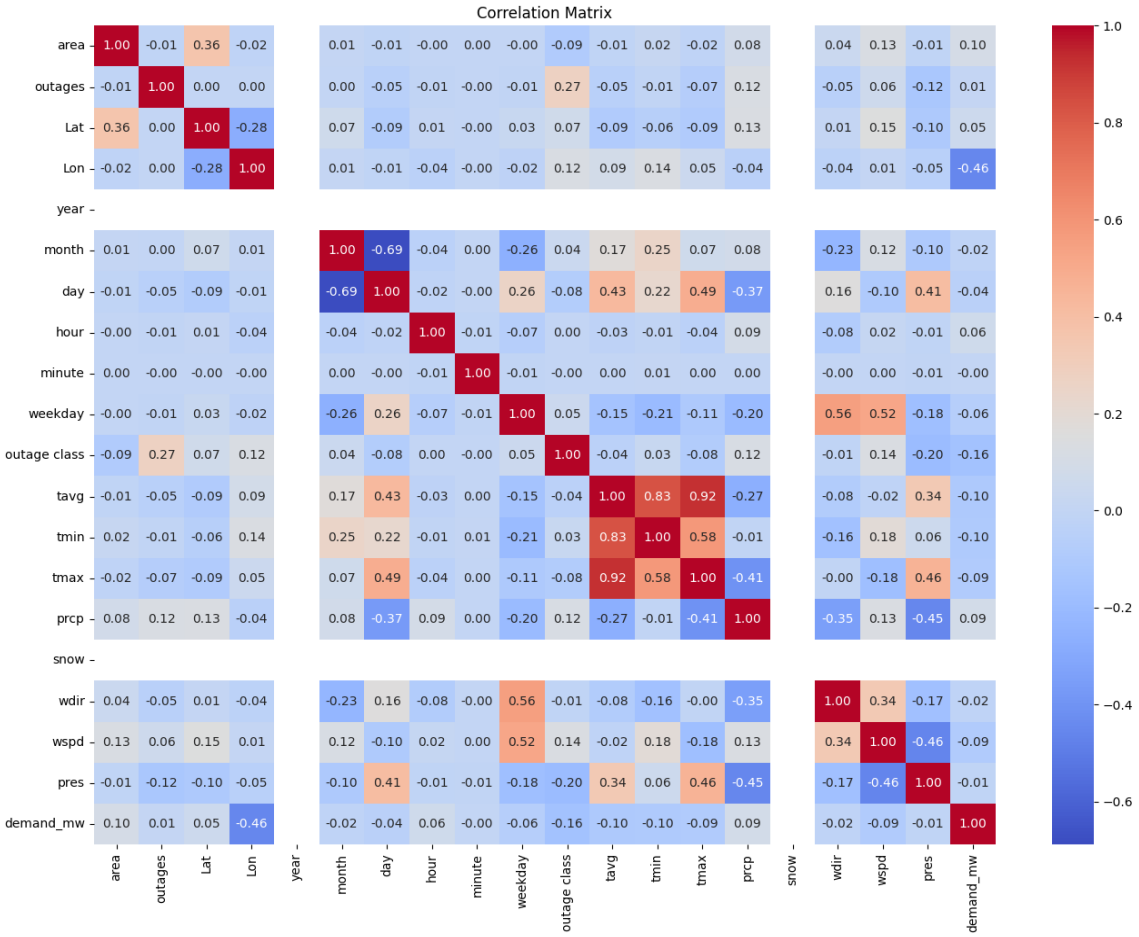
شکل ۲. توزیع تعداد قطعی برق در ساعات مختلف روز.

Figure 2. Distribution of electricity outages at different times of the day.

جدول ۱. متغیرهای موردبررسی در مجموعه داده اولیه

Table 1. Variables analyzed in the initial dataset

ویژگی	توصیف ویژگی
uid	شناسه یکتای هر نمونه داده که برای شناسایی ردیف‌های مختلف استفاده می‌شود.
Area	کد منطقه‌ای که نشان‌دهنده موقعیت مکانی داده‌ها است.
outages	تعداد قطعی‌های برق ثبت‌شده در هر منطقه در یک بازه زمانی مشخص.
dt_stamp	زمان و تاریخ ثبت داده به‌صورت رشته متنی.
Lat	موقعیت جغرافیایی منطقه (طول جغرافیایی).
Lon	موقعیت جغرافیایی منطقه (عرض جغرافیایی).
Year,month,day,hour, Minute, Weekday	اطلاعات زمانی رخداد
outage class	دسته‌بندی نوع قطعی برق بر اساس شدت یا عوامل ایجادکننده آن. (این ستون در مراحل پیش‌پردازش و با دسته‌بندی مقادیر ستون outages اضافه خواهد شد و در مجموعه داده اولیه موجود نیست.)
tavg	متوسط دمای ثبت‌شده در روز. این متغیرها تأثیر زیادی بر مصرف برق دارند.
tmin	کمترین دمای ثبت‌شده در روز. این متغیرها تأثیر زیادی بر مصرف برق دارند.
tmax	بیشترین دمای ثبت‌شده در روز. این متغیرها تأثیر زیادی بر مصرف برق دارند.
prcp	میزان بارش ثبت‌شده در روز که می‌تواند بر مصرف انرژی تأثیرگذار باشد.
snow	مقدار برف گزارش‌شده که ممکن است باعث افزایش مصرف برق برای گرمایش شود.
Wdir	جهت باد که بر انتقال انرژی و تأثیرات آب‌وهوایی منطقه تأثیر دارد.
wspd	سرعت باد که ممکن است بر عملکرد شبکه برق تأثیرگذار باشد.
pres	فشار هوا که در تحلیل شرایط جوی و تأثیرات آن بر شبکه برق مهم است.
zone	نام یا دسته‌بندی منطقه‌ای که داده به آن تعلق دارد.
demand_mw	میزان مصرف برق (تقاضا) در هر منطقه به مگاوات که نشان‌دهنده نیاز انرژی منطقه است.



شکل ۳. ماتریس همبستگی مجموعه داده اول (اصلی).

Figure 3. Correlation matrix of the first (primary) dataset.

داده‌های جدولی به دلیل پراکندگی و استقلال نسبی ویژگی‌ها، الگوبرداری و استخراج روابط پنهان را دشوارتر می‌سازند. در چنین شرایطی، مدل‌های یادگیری ماشین و یادگیری عمیق برای دستیابی به عملکرد مناسب نیازمند طراحی دقیق‌تر هستند در غیر این صورت، احتمال وقوع پدیده بیش‌برازش افزایش می‌یابد، چراکه مدل ممکن است به جای یادگیری الگوهای کلی، صرفاً روابط تصادفی موجود در داده‌های آموزشی را حفظ کند. به‌بیان‌دیگر، ماهیت جدولی داده‌ها چالشی مضاعف در فرآیند مدل‌سازی ایجاد می‌کند و لزوم بهره‌گیری از روش‌های مناسب برای کاهش پیچیدگی و افزایش تعمیم‌پذیری مدل را برجسته می‌سازد.

علاوه بر این، نتایج حاصل از ارزیابی مدل‌های پایه یادگیری ماشین نشان داد که در صورت استفاده از ویژگی‌های جغرافیایی نظیر طول جغرافیایی، عرض جغرافیایی و منطقه (Area) در فرآیند مدل‌سازی، این ویژگی‌ها به دلیل ماهیت جدولی داده و همچنین قابلیت بالای تفکیک‌پذیری که فراهم می‌کنند، نقش بسیار پررنگی در تصمیم‌گیری مدل ایفا می‌کنند. بررسی مقادیر اهمیت ویژگی‌ها (Feature Importance) نشان داد که سهم این سه متغیر به‌مراتب بیشتر از سایر ویژگی‌ها است و در عمل، مدل عمدتاً بر پایه آن‌ها تصمیم‌گیری می‌کند.

هرچند این موضوع منجر به افزایش ظاهری عملکرد مدل در محیط داده‌های موجود می‌شود، اما یک محدودیت اساسی نیز به همراه دارد. تکیه بیش‌ازحد مدل بر ویژگی‌های جغرافیایی، عمومیت‌پذیری آن را کاهش داده و مانع از کاربرد مؤثر مدل در سایر مناطق با شرایط مکانی متفاوت می‌گردد. افزون بر این، هدف اصلی ما تمرکز بر شاخص‌های آب‌وهوایی و الگوی مصرف انرژی به‌عنوان عوامل بنیادین در بروز و پیش‌بینی خاموشی‌ها است، نه اتکای صرف به موقعیت مکانی داده‌ها.

بنابراین، به‌منظور افزایش قابلیت تعمیم مدل و فراهم آوردن امکان استفاده از آن در مناطق گوناگون، تصمیم گرفته شد که ویژگی‌های جغرافیایی (طول، عرض و

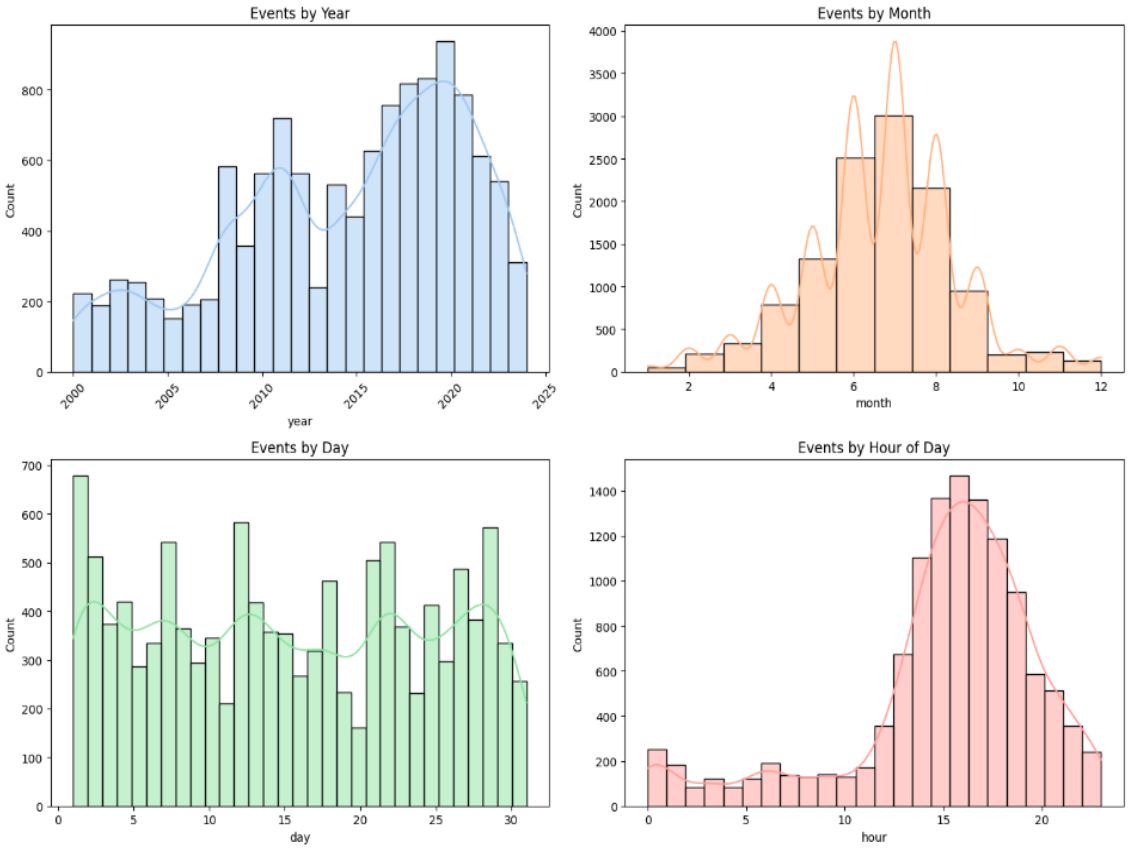
منطقه) از فرآیند مدل‌سازی کنار گذاشته شوند. در ادامه، مدل‌ها صرفاً بر پایه سایر متغیرها (ازجمله متغیرهای آب‌وهوایی و شاخص‌های مرتبط با تقاضای انرژی) آموزش داده می‌شوند تا علاوه بر کاهش خطر بیش‌برازش، وابستگی مدل به شرایط مکانی خاص نیز به حداقل برسد. در ادامه برای بهبود عملکرد مدل، مجموعه داده ثانویه نیز از مراجع مختلف که شامل اطلاعات آب و هوایی و جغرافیایی قطعی برق در سال‌های ۲۰۰۰ تا ۲۰۲۴ بوده، جمع‌آوری شد. این مجموعه داده از تعداد داده و ویژگی‌های کمتری به نسبت مجموعه داده اولیه برخوردار بوده، اما ازلحاظ تنوع زمانی غنی‌تر است. همچنین مجموعه داده دوم برخلاف مجموعه داده اولیه، شامل اطلاعات برچسب خورده نیست و به همین دلیل نمی‌توان از آن برای مدل‌های با نظارت استفاده کرد. اطلاعات جغرافیایی مربوط به قطعی‌ها از سال ۲۰۰۰ تا ۲۰۲۴ از EIA و اطلاعات آب و هوایی این مجموعه داده به کمک API تهیه‌شده از NASA جمع‌آوری شده‌اند. این مجموعه داده شامل ۱۱۹۰۵ سطر داده و ۲۴ ستون است. هرچند به دلیل عدم تطابق تعدادی از ستون‌های دو مجموعه داده، نمی‌توان از ستون‌هایی که در مجموعه دوم بوده و در مجموعه اول نیست، استفاده کرد. در ادامه در شکل **۴**، توزیع خاموشی‌ها در مجموعه داده دوم به‌صورت سالیانه، ماهانه، هفتگی و روزانه نشان داده‌شده است.

۴- روش حل

۴-۱- پاک‌سازی داده‌ها

پیش‌بینی خاموشی برق یکی از مؤلفه‌های حیاتی در سامانه‌های نوین انرژی است؛ زیرا وقوع ناگهانی خاموشی‌ها نه‌تنها موجب خسارات اقتصادی می‌شود، بلکه با ایجاد اختلال در توزیع بار، افزایش فشار بر تجهیزات و نیاز به تعمیرات پیش‌بینی‌نشده، کارایی کلی سیستم را کاهش می‌دهد.

در پروژه حاضر که هدف آن پیش‌بینی خاموشی‌های برق با استفاده از داده‌های متنوع اقلیمی، مکانی و انرژی است، کیفیت و انسجام داده‌ها نقش کلیدی در موفقیت مدل‌های یادگیری ماشین دارد؛ بنابراین، یکی از مهم‌ترین مراحل قبل از طراحی مدل، فاز پیش‌پردازش داده‌هاست.



شکل ۴. تعداد قطعی در طول زمان در مجموعه داده دوم (فرعی).

Figure 4. Analysis of outages over time in the second (secondary) dataset.

مجموعه داده اولیه مورداستفاده در این پروژه شامل ۶۷,۳۲۱ ردیف و ۲۳ ویژگی بوده و از منابعی معتبر مانند NOAA، EIA و وب‌سایت Kaggle استخراج شده است. این داده‌ها بازه زمانی بین ۲۹ مارس تا ۱۹ آوریل را پوشش داده و متشکل از اطلاعات خاموشی، شرایط آب‌وهوایی، تقاضای انرژی و ویژگی‌های زمانی و مکانی مرتبط با هر رخداد است. مراحل مختلف پیش‌پردازش برای آماده‌سازی این داده‌ها جهت آموزش مدل‌ها در ادامه آورده شده‌اند.

یکی از چالش‌های مهم در تحلیل داده‌های دنیای واقعی، وجود مقادیر گمشده است. در این پروژه، پیش از انجام هرگونه تحلیل یا مدل‌سازی، تمامی ستون‌ها ازنظر وجود مقادیر ناقص بررسی شدند. برای مدیریت این مقادیر:

ابتدا میزان داده‌های گمشده برای هر ستون محاسبه شد. قاعده‌ای در نظر گرفته شد که اگر درصد داده‌های گمشده در یک ویژگی بیش از ۲۰٪ باشد، آن ستون حذف شود.

برای پر کردن مقادیر گمشده از روش KNN Imputation استفاده شد. این روش بر اساس ویژگی‌های مشابه در نمونه‌های همسایه، مقدار گمشده را با میانگین نمونه‌های نزدیک تخمین می‌زند و در مقایسه با روش‌هایی مانند میانگین یا میانه، دقت بیشتری دارد. وجود داده‌های تکراری می‌تواند باعث بایاس در یادگیری مدل شده و صحت نتایج را کاهش دهد؛ بنابراین، مجموعه داده‌ها ازنظر تکرار بررسی شدند. ردیف‌هایی که تمامی ویژگی‌هایشان مشابه بودند به‌عنوان داده تکراری شناسایی و حذف شدند. این مرحله باعث کاهش حجم داده بدون از دست دادن اطلاعات معنادار شد.

در مجموعه داده موردبررسی، ویژگی‌هایی مانند دمای هوا، سرعت باد، فشار جو، میزان بارش و حتی متغیرهایی مانند تقاضای برق دارای مقیاس‌های متفاوتی بودند. به همین منظور، برای هماهنگ‌سازی این مقیاس‌ها، از تکنیک StandardScaler استفاده شد که مقادیر را به‌صورت نرمال با میانگین صفر و انحراف معیار یک تبدیل می‌کند.

ضرورت این مرحله به‌ویژه زمانی بیشتر می‌شود که از مدل‌هایی مانند شبکه عصبی چندلایه (MLP) یا خودرمزگذار استفاده شود. این مدل‌ها به مقیاس داده‌ها بسیار حساس هستند، چراکه وزن‌ها در فرآیند بهینه‌سازی (مثل Gradient Descent) به‌طور مستقیم با مقدار ویژگی‌ها در ارتباط‌اند. بدون نرمال‌سازی، مدل ممکن است در یادگیری ویژگی‌هایی با مقادیر بزرگ‌تر تمرکز کند و از ویژگی‌های دیگر غافل شود. به همین دلیل، استانداردسازی شرط لازم برای پایداری، سرعت آموزش و دقت نهایی این مدل‌ها محسوب می‌شود.

تحلیل داده‌های پرت یا غیرعادی نقش مهمی در درک بهتر توزیع و رفتار داده‌ها دارد. در این پروژه، توجه ویژه‌ای به ویژگی outages (تعداد خاموشی‌ها) شد. این ویژگی دارای چولگی بسیار بالایی بود (بیش از ۱۴)، به‌گونه‌ای که اکثر نمونه‌ها دارای مقدار ۱ یا ۲ بودند، اما مواردی با بیش از ۱۰۰ خاموشی نیز ثبت شده بود.

برای تحلیل دقیق‌تر، از تحلیل صدک‌ها (Quantile-Based Outlier Detection) استفاده شد. با بررسی صدک‌های بالا و پایین، محدوده‌های مشکوک به داده پرت شناسایی شدند. بااین‌حال، برخلاف بسیاری از پروژه‌ها که داده‌های پرت حذف می‌شوند، در این پروژه تصمیم به عدم حذف داده‌های پرت گرفته شد. علت این تصمیم آن بود که:

۱) داده‌های پرت در ویژگی outages به احتمال زیاد، منعکس‌کننده رخدادهای واقعی بحرانی هستند.
۲) مطابق با توصیف ارائه‌شده توسط مرجع داده (Kaggle)، داده‌های مربوط به خاموشی‌های بزرگ، مربوط به بلایای طبیعی یا شرایط جوی شدید بوده‌اند.

حذف آن‌ها می‌توانست منجر به کاهش دقت مدل در پیش‌بینی شرایط بحرانی واقعی شود که خود هدف اصلی پروژه است.

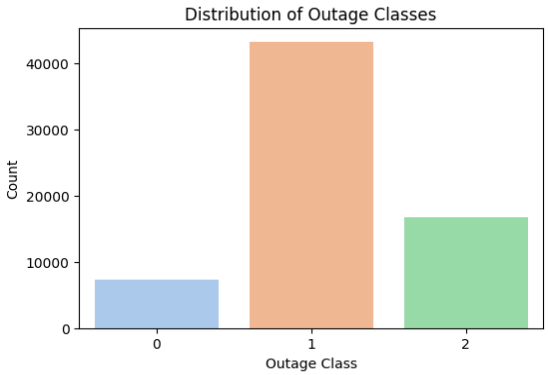
برای بررسی هم‌بستگی میان ویژگی‌های عددی، از نقشه حرارتی (Heatmap) ماتریس هم‌بستگی استفاده شد. نتایج نشان داد که بیشتر ویژگی‌ها ازنظر آماری، مستقل از یکدیگر هستند. این ویژگی اهمیت بالایی دارد، چراکه در صورت وجود هم‌بستگی بالا بین ویژگی‌ها، مدل دچار مشکل هم‌خطی (Multicollinearity) می‌شود و قدرت تفکیک آن کاهش می‌یابد.

همچنین وجود تنوع و استقلال میان ویژگی‌ها به الگوریتم یادگیری اجازه می‌دهد تا اطلاعات غیرتکراری و غنی‌تری برای پیش‌بینی خروجی‌ها داشته باشد.

با توجه به چولگی بسیار زیاد متغیر outages (که در تحلیل داده‌های پرت موردبررسی قرار گرفت)، استفاده از آن به‌صورت مستقیم در مدل‌سازی طبقه‌بندی باعث عدم

تعادل شدید کلاس‌ها (Imbalanced Classes) می‌شد. عدم تعادل کلاس، وضعیتی است که در آن تعداد نمونه‌ها در یک کلاس (کلاس اکثریت) به‌مراتب بیشتر از تعداد نمونه‌ها در کلاس‌های دیگر (کلاس اقلیت) است.

برای مدیریت این چالش، متغیر پیوسته outages به سه کلاس طبقه‌بندی گسسته تبدیل شد: ریسک کم (Low Risk)، ریسک متوسط (Medium Risk) و ریسک بالا (High Risk). توزیع واقعی کلاس‌ها پس‌ازاین دسته‌بندی در شکل **۵** نمایش داده شده است.



شکل ۵. توزیع داده‌ها بین کلاس‌های مختلف.

Figure 5. Distribution of data among different classes.

همان‌طور که مشاهده می‌شود، کلاس ریسک بالا با تنها ۷,۳۰۲ نمونه، به‌شدت کلاس اقلیت محسوب می‌شود. در مقابل، کلاس ریسک کم با بیش از ۶۴٪ نمونه‌ها، کلاس اکثریت است. این عدم تعادل، چالش‌های جدی برای آموزش مدل‌های یادگیری عمیق (مانند MLP و خودرمزگذار در بخش‌های بعدی) ایجاد می‌کند. برای مثال، می‌توان به موارد زیر اشاره کرد:

۱) بایاس به سمت کلاس اکثریت: مدل در طول آموزش تمایل پیدا می‌کند که برای به حداقل رساندن خطای کلی، تمامی نمونه‌ها را در کلاس اکثریت پیش‌بینی کند. درنتیجه، مدل در پیش‌بینی رخدادهای حیاتی اما نادر (High Risk) که اتفاقاً بیشترین اهمیت را برای سیستم انرژی دارند، عملکرد ضعیفی از خود نشان می‌دهد.

۲) کاهش قابلیت تعمیم: توانایی مدل در تعمیم آموخته‌ها به نمونه‌های جدید از کلاس اقلیت (رخدادهای بحرانی) به‌شدت کاهش می‌یابد.

برای رفع این مشکل در مرحله پیاده‌سازی، اگرچه دسته‌بندی بر اساس صدک‌ها ساختار مسئله را به یک طبقه‌بندی سه‌گانه تبدیل کرد، اما در مرحله آموزش مدل (MLP) نیاز به تدابیر بیشتری بود. این تدابیر شامل استفاده از تابع زیان آنتروپی متقابل وزن‌دار (Weighted Cross-Entropy Loss) یا سایر تکنیک‌های مدیریت عدم تعادل در سطح الگوریتم است تا به کلاس اقلیت وزن بیشتری در محاسبه خطا داده شود و مدل را وادار به یادگیری دقیق‌تر این موارد حیاتی نماید.

به‌منظور ارزیابی منصفانه عملکرد مدل‌ها، معیار Weighted-average F1-score به‌عنوان شاخص اصلی ارزیابی در تمامی آزمایش‌ها انتخاب شده است.

در مدل‌های پایه شامل MLP ساده، MLP همراه با مکانیزم توجه و MLP دارای اتصالات باقیمانده، از تابع خطای Categorical Cross-Entropy استاندارد استفاده شده است تا مقایسه معماری‌ها بدون تأثیرگذاری عامل تابع خطا انجام شود. در مقابل، برای مدل پیشنهادی نهایی، از Focal Loss استفاده شده است. این تابع خطا با کاهش وزن نمونه‌های به‌درستی طبقه‌بندی‌شده و تمرکز بیشتر بر نمونه‌های دشوار و کلاس‌های اقلیت، نقش مؤثری در مدیریت عدم‌توازن کلاس‌ها ایفا می‌کند. انتخاب این تابع خطا منجر به بهبود پایداری یادگیری و افزایش عملکرد مدل در کلاس‌های کم‌نمونه شده است.

در این مطالعه، مجموعه داده مریلند به‌طور دقیق به سه زیرمجموعه مجزا تقسیم شد تا آموزش مدل و ارزیابی آن به‌صورت قابل‌اعتماد انجام شود: ۷۰٪ از داده‌ها برای آموزش، ۱۵٪ برای اعتبارسنجی و ۱۵٪ باقی‌مانده برای تست اختصاص یافت. مجموعه تست کاملاً جدا نگه داشته شد و در طول آموزش مدل و تنظیم hyperparameters استفاده نشد و تنها یک‌بار برای ارزیابی نهایی مدل آموزش‌دیده مورداستفاده قرار گرفت.

برای جلوگیری از نشت اطلاعات و حفظ توزیع یکسان ویژگی‌ها بین زیرمجموعه‌ها، تمامی مراحل پیش‌پردازش شامل مقیاس‌بندی (scaling) و جایگزینی مقادیر گمشده (imputation) تنها روی داده‌های آموزش انجام شد. پارامترهای به‌دست‌آمده از آموزش، برای مقیاس‌بندی و جایگزینی مقادیر گمشده در مجموعه‌های اعتبارسنجی و تست اعمال شدند. این روش تضمین می‌کند که مدل هیچ‌گونه اطلاعات قبلی از داده‌های دیده‌نشده کسب نکند

و توزیع ویژگی‌های ورودی بین زیرمجموعه‌ها یکسان باقی بماند که درنتیجه ارزیابی عملکرد مدل به‌صورت واقعی و قابل‌اعتماد انجام شود. با رعایت دقیق این پروتکل، خطر بیش‌برازش کاهش‌یافته و معیارهای ارزیابی به‌دست‌آمده از مجموعه‌های اعتبارسنجی و تست، عملکرد واقعی مدل را روی داده‌های کاملاً دیده‌نشده منعکس می‌کنند.

درنهایت مراحل پیش‌پردازش انجام‌شده در این پروژه نه‌تنها باعث افزایش کیفیت داده‌ها شدند، بلکه نقش حیاتی در کاهش بایاس، افزایش قابلیت تعمیم مدل و دقت در پیش‌بینی شرایط واقعی خاموشی برق ایفا کردند. بهره‌گیری از تکنیک‌هایی چون KNN Imputation، حذف داده‌های تکراری، نرمال‌سازی داده‌ها با StandardScaler، تحلیل چولگی و کشیدگی، بررسی هم‌بستگی‌ها و تعریف کلاس‌بندی هدفمند برای برچسب خروجی ازجمله اقدامات کلیدی در این مرحله بوده‌اند. با این آمادگی، داده‌ها به مرحله بعدی یعنی طراحی مدل‌های یادگیری عمیق و ارزیابی عملکرد آن‌ها وارد شدند.

۴-۲- انگیزه و دید کلی

پیش‌بینی قابل‌اعتماد خاموشی برق در شرایط متغیر محیطی، یک مسئله پیچیده و چالش‌برانگیز است، به‌ویژه زمانی که هدف، تعمیم‌پذیری مدل در بعد زمان و مکان باشد. در بسیاری از روش‌های سنتی، متغیرهای جغرافیایی مانند طول و عرض جغرافیایی یا نواحی منطقه‌ای، نقش کلیدی در استخراج الگوها دارند؛ اما این وابستگی جغرافیایی، کارایی مدل را برای استفاده در مناطق جدید (که داده‌های دقیق مکانی ممکن است موجود نباشد یا به‌درستی تنظیم نشده باشد) محدود می‌کند. برای رفع این چالش و افزایش بهره‌وری سیستم‌های انرژی، یک چارچوب دومرحله‌ای مبتنی بر یادگیری عمیق پیشنهاد می‌شود که وابستگی مدل به مکان را از بین می‌برد و با استفاده از یادگیری بدون نظارت، ویژگی‌های عمیق و معناداری را از داده‌ها استخراج می‌کند. هسته این ایده بر ترکیب دو مجموعه داده با ساختار متفاوت استوار است:

مجموعه داده A: شامل بیش از ۶۷۰۰۰ نمونه از یک ماه خاص می‌باشد که هرکدام به یکی از سه کلاس برچسب‌گذاری شده‌اند: بدون خاموشی، خاموشی جزئی و خاموشی شدید. این مجموعه دارای برچسب‌های دقیق است، ولی افق زمانی محدود آن، دید مناسبی نسبت به رفتارهای بلندمدت سیستم ارائه نمی‌دهد.

مجموعه داده B: شامل حدود ۱۱۹۰۰ رکورد ساعتی از داده‌های آب‌وهوایی و تقاضای انرژی در بازه بلندمدت ۲۴ ساله (۲۰۰۰ تا ۲۰۲۴) می‌باشد. این مجموعه داده بدون برچسب است ولی ازنظر تنوع فصلی و شرایط محیطی غنی است. این تفاوت ساختاری میان دو مجموعه‌داده، امکان پیاده‌سازی رویکرد یادگیری انتقالی (Transfer Learning) را فراهم کرد؛ بدین‌صورت که دانش تعمیم‌یافته حاصل از داده‌های بلندمدت در قالب ویژگی‌های فشرده‌شده، به مجموعه‌داده هدف که شامل رویدادهای واقعی خاموشی بود، منتقل شد.

برای استخراج این ویژگی‌های تعمیم‌پذیر از Dataset B، از خودرمزگذار استفاده شد. خودرمزگذار یک معماری غیرنظارتی است که بدون نیاز به برچسب، قادر به یادگیری نمایش‌های سطح بالا از داده‌هاست. این معماری با فشرده‌سازی ورودی به یک فضای میان‌بُعدی (Bottleneck) و بازسازی مجدد آن، مدلی می‌سازد که اطلاعات مهم را حفظ کرده و نویزها را حذف می‌کند. در این پروژه، خودرمزگذار آموزش‌دیده بر روی ویژگی‌های مشترک و نرمال‌شده Dataset B، بردارهای فشرده‌ای از هر نمونه را تولید کرد که پس از آموزش مدل، به‌عنوان embedding برای داده‌های Dataset A نیز استفاده شد. این کار با فرض اشتراک ساختاری داده‌ها و یکسان بودن اسامی و مقیاس‌های ویژگی‌ها ممکن شد. سپس این embedding ها به بردارهای ویژگی‌های Dataset A الصاق شده و مجموعه‌ای غنی از اطلاعات تاریخی و فعلی برای آموزش مدل فراهم شد. این ترکیب در عمل باعث شد مدل ما نه‌تنها اطلاعات لحظه‌ای، بلکه نوعی از حافظه تاریخی از رفتار داده‌ها طی بیست سال گذشته را نیز در اختیار داشته باشد.

معماری اصلی شبکه عصبی طراحی‌شده بر مبنای MLP تقویت‌شده با مکانیزم توجه و اتصالات باقیمانده است. طراحی این مدل با هدف غلبه بر دو چالش اصلی انجام شد:

۱) شناسایی ویژگی‌های کلیدی در میان تعداد زیاد ویژگی‌های عددی

۲) جلوگیری از ناپدید شدن گرادیان‌ها و بیش‌برازش در شبکه‌های عمیق. لایه توجه با استفاده از توابع softmax و tanh به‌صورت تطبیقی وزن‌هایی به ویژگی‌ها اختصاص داد تا مدل بتواند با تمرکز بیشتر بر مهم‌ترین آن‌ها، عملکرد طبقه‌بندی را بهبود دهد.

از سوی دیگر، استفاده از بلوک‌های باقیمانده باعث شد که اطلاعات اولیه به‌صورت مستقیم در طول مسیر شبکه حفظ شود و یادگیری بهتر انجام گیرد. این معماری درنهایت توانست با استفاده از داده‌های ترکیبی f1-score کلی را به ۸۱ درصد برساند که نسبت به مدل پایه افزایش قابل‌توجهی محسوب می‌شود. جالب‌تر آنکه دقت در کلاس دشوار Major Outage نیز به‌طور محسوس بهبود یافت.

این افزایش عملکرد، نشان از آن دارد که مدل نهایی توانسته است با استفاده از دانش تاریخی منتقل‌شده از طریق خودرمزگذار، ساختارهای پنهان و روابط پیچیده‌تری را در داده‌ها تشخیص دهد. به‌عبارتی‌دیگر، این پروژه با موفقیت نشان داد که استفاده از خودرمزگذار به‌عنوان ابزار یادگیری انتقالی، به همراه طراحی مدل ترکیبی عمیق، می‌تواند منجر به ساخت سیستمی دقیق، تعمیم‌پذیر و قدرتمند برای پیش‌بینی خاموشی شود.

درحالی‌که مدل‌های سنتی عمدتاً بر ویژگی‌های جغرافیایی مانند طول و عرض جغرافیایی تکیه دارند تا الگوهای خاموشی را شناسایی کنند، این وابستگی باعث کاهش قابلیت تعمیم‌پذیری آن‌ها به مناطق و بازه‌های زمانی جدید می‌شود. نتایج حاصل نشان می‌دهند که حتی با حذف کامل ویژگی‌های جغرافیایی، مدل پیشنهادی عملکرد بسیار خوبی از خود نشان می‌دهد. این امر باعث می‌شود چارچوب ارائه‌شده قابلیت تطبیق و گسترش‌پذیری بالایی داشته باشد. درنتیجه، این روش نه‌تنها وابستگی به مکان را کاهش می‌دهد، بلکه با بهبود دقت پیش‌بینی، موجب افزایش بهره‌وری سامانه‌های انرژی از طریق آمادگی بهتر، توزیع هوشمند بار و مدیریت پیشگیرانه منابع می‌شود.

۴-۳- معماری روش پیشنهادی

روش پیشنهادی بر اساس یک ساختار دومرحله‌ای ایجادشده است. انتخاب این معماری دومرحله‌ای مستقیماً به‌منظور غلبه بر دو چالش حیاتی در مدل‌سازی قطعی برق انجام شده است:

۱) محدودیت تعمیم‌پذیری ناشی از وابستگی به ویژگی‌های مکانی

۲) عدم توازن و کمبود داده‌های برچسب‌خورده در مقایسه با داده‌های تاریخی غنی بدون برچسب. دو مرحله روش پیشنهادی در ادامه با جزئیات لازم معرفی می‌شود.

الف) مرحله اول: خودرمزگذار برای یادگیری ویژگی‌های تعمیم‌پذیر

خودرمزگذار یک شبکه عصبی خودنظارتی است که هدف آن، بازسازی ورودی از طریق فشرده‌سازی به فضای نهفته است. به‌طور خاص برای فشرده‌سازی اطلاعات در یک فضای نهفته (Latent Space) با ابعاد پایین‌تر طراحی شده است (Ghosh, 2025). برخلاف مدل‌های طبقه‌بندی یا رگرسیون، خودرمزگذار تلاش می‌کند اطلاعات حیاتی و غیرتکراری داده را در یک‌لایه فشرده (Bottleneck) حفظ کند.

معماری شامل:

رمزگذار (Encoder): تابعی که ویژگی‌های اصلی داده را به برداری با ابعاد پایین‌تر فشرده می‌کند.

رمزگشا (Decoder): تابعی که ورودی اصلی را از بردار فشرده بازسازی می‌کند.

شبکه با استفاده از خطای میانگین مربعات (MSE) بین ورودی و خروجی بازسازی‌شده آموزش داده می‌شود:

$$L_{recon} = \frac{1}{N}\sum_{i=1}^{N} ||x^{(i)} - \hat{x}^{(i)}||_2^2 \qquad (۱)$$

بدون نیاز به برچسب، خودرمزگذار، انتخابی ایده‌آل برای تحلیل داده‌های تاریخی بزرگ مانند Dataset B است.

در گام نخست روش پیشنهادی، از یک خودرمزگذار برای استخراج ویژگی و یادگیری انتقالی استفاده شد که سه هدف اصلی را در این پژوهش دنبال می‌کند:

۱) استخراج ویژگی‌های مکان-مستقل و افزایش تعمیم‌پذیری: هدف این پژوهش، ساخت مدلی است که وابستگی به ویژگی‌های جغرافیایی ایستا (مانند مختصات) را حذف کند. با آموزش خودرمزگذار بر روی یک مجموعه‌داده بلندمدت ۲۴ ساله و فاقد برچسب (که شامل روندهای آب‌وهوایی و تقاضای انرژی است)، خودرمزگذار

الگوهای زیربنایی پیچیده و مشترکی را که بر رفتار سیستم انرژی حاکم است، فرا می‌گیرد. این نگاه مشابه پژوهش‌هایی است که کاربرد خودرمزگذار (نوع تعمیم‌یافته‌اش) را برای تشخیص ناهنجاری در سیستم‌های قدرت و انرژی بررسی کرده‌اند؛ به‌عنوان مثال، در مقاله‌ای نشان داده شده است که مدل مبتنی بر خودرمزگذار می‌تواند در سیستم‌های توزیع برق با DERs پیچیده، رفتار طبیعی شبکه را یاد بگیرد و ناهنجاری‌ها یا وقایع غیرعادی را با دقت بالا تشخیص دهد (Zideh & Solanki, 2024).

۲) انتقال دانش (Knowledge Transfer): با استفاده از بخش رمزگذار آموزش‌دیده، اطلاعات نهفته و فشرده‌شده از مجموعه داده بزرگ (منبع) استخراج و به‌عنوان ویژگی‌های جدید به مجموعه داده کوچک برچسب‌خورده (هدف) الصاق می‌شود.

۳) مدیریت داده‌های با ابعاد بالا و نویز: خودرمزگذار با فشرده‌سازی داده‌های ورودی با ابعاد بالا، مهم‌ترین ویژگی‌ها را حفظ کرده و در عین حال نویز را کاهش می‌دهد و به‌طور موثرتری داده‌های جدولی انرژی را پردازش می‌کند (Azzalini et al., 2025).

ب) مرحله دوم: شبکه پرسپترون چندلایه (MLP) بهبودیافته

در گام دوم، یک شبکه MLP برای وظیفه نهایی طبقه‌بندی (پیش‌بینی احتمال قطعی) استفاده شد. این شبکه ورودی خود را از ترکیب ویژگی‌های خام لحظه‌ای و بردارهای نهفته استخراج‌شده توسط خودرمزگذار دریافت می‌کند.

انتخاب MLP برای طبقه‌بندی نهایی پس از استخراج ویژگی، با چشم‌انداز بهره‌برداری از فضای پنهان یاد گرفته‌شده توسط خودرمزگذار برای پیش‌بینی احتمالی خاموشی ایجاد شده است. چنین ترکیبی در مقالات جدید موردتوجه قرار گرفته است. برای نمونه در یک مقاله از ایده ترکیب برای تشخیص خطا بهره گرفته شده است (Arifeen et al., 2024). یا در مقاله دیگر از مدل‌های عمیق با قابلیت یادگیری غیرخطی برای تشخیص ناهنجاری‌ها در میکروگریدهای مبتنی بر اینورتر، بدون نیاز به مدل ریاضی دقیق سیستم استفاده شده است (Tabassum et al., 2024).

۴-۴- جزییات پیاده‌سازی

معماری مورداستفاده شامل یک خودرمزگذار متقارن با سه لایه رمزگذار و سه لایه رمزگشا است. اندازه فضای نهفته به‌صورت تجربی برابر با ۸ انتخاب شد که تعادلی میان دقت بازسازی و فشرده‌سازی ایجاد می‌کند. تمام ورودی‌ها با استفاده از نرمال‌سازی استاندارد (Standard Scaling) پیش‌پردازش شدند. از تکنیک‌های Dropout و Batch Normalization برای جلوگیری از بیش‌برازش به شرح زیر استفاده می‌شود:

Dropout یک تکنیک ساده و درعین‌حال قدرتمند برای کاهش بیش‌برازش است. ایده اصلی این روش، غیرفعال‌سازی تصادفی برخی نرون‌ها در زمان آموزش است تا مدل به ترکیبات خاصی از نرون‌ها وابسته نشود و ساختار شبکه در هر بار عبور داده‌ها تغییر کند.

در عمل، برای هر نرون از لایه، با استفاده از یک توزیع برنولی با احتمال نگه‌داشتن p، تصمیم گرفته می‌شود که آیا نرون فعال بماند یا خیر. فرمول ریاضی این فرآیند به‌صورت زیر است:

$$\tilde{h}_i = m_i . h_i, \qquad m_i \sim Bernoulli(p) \qquad (۲)$$

در این رابطه:

h_i خروجی نرون i پیش از اعمال *dropout* است، m_i متغیر برنولی است که مقدار ۰ یا ۱ می‌گیرد، $\tilde{h}_i$ خروجی نهایی نرون پس از اعمال *dropout* است. برای جبران تأثیر حذف نرون‌ها، در زمان inference (پیش‌بینی)، از مقیاس‌دهی استفاده می‌شود.

تکنیک Batch Normalization یک روش نرمال‌سازی برای شبکه‌های عصبی عمیق است که با تثبیت توزیع آماری ورودی لایه‌های میانی، فرآیند آموزش را سریع‌تر، پایدارتر و قابل تعمیم‌تر می‌سازد. این روش با نرمال‌سازی خروجی هر لایه در سطح هر دسته آموزشی، از وقوع پدیده موسوم به تغییر هم‌توزیعی داخلی جلوگیری می‌کند؛ پدیده‌ای که در آن تغییرات مکرر در توزیع ورودی لایه‌ها موجب کندی و ناپایداری یادگیری می‌گردد.

Batch Normalization با ایجاد توزیعی یکنواخت در مقیاس استاندارد (میانگین صفر و واریانس یک) برای هر ویژگی، سبب می‌شود که مدل نسبت به مقداردهی اولیه پارامترها حساسیت کمتری داشته و با نرخ یادگیری بالاتر، به همگرایی برسد. پس از نرمال‌سازی، مقادیر به‌دست‌آمده با استفاده از دو پارامتر قابل یادگیری یعنی γ (ضریب مقیاس‌دهی) و β (ضریب انتقال)، مجدداً تنظیم می‌شوند (مطابق رابطه ۶). این بازتنظیم اجازه می‌دهد که مدل در صورت نیاز، توزیع نرمال‌شده را به مقیاس بهینه‌ای که برای یادگیری مؤثرتر است، بازگرداند.

مراحل اعمال نرمال‌سازی برای ورودی‌های مدل پیشنهادی معادل $x = \{x_1, x_2, \dots, x_m\}$ در یک batch به‌صورت زیر است:

۱. محاسبه میانگین batch:

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \qquad (۳)$$

که در آن m تعداد نمونه‌ها است.

۲. محاسبه واریانس batch:

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_B)^2 \qquad (۴)$$

۳. نرمال‌سازی:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \qquad (۵)$$

۴. مقیاس‌دهی و انتقال با پارامترهای قابل یادگیری:

$$y_i = \gamma . \hat{x}_i + \beta \qquad (۶)$$

که در آن β و γ پارامترهایی قابل یادگیری هستند که به مدل اجازه می‌دهند در صورت نیاز توزیع نرمال‌شده را تغییر دهد. و ε یک مقدار کوچک برای جلوگیری از تقسیم بر صفر است. در مرحله تست، به‌جای محاسبه میانگین و واریانس روی‌داده batch، از میانگین و واریانس میانگین‌گیری‌شده در طول آموزش استفاده می‌شود.

۴-۵- انتقال ویژگی و غنی‌سازی

برای هر نمونه در مجموعه داده A، بردار نهفته تولید و به ویژگی‌های اولیه افزوده شد:

$$X_{aug} = [X_A, Z_A] \in R^{d+k} \qquad (۷)$$

این فرآیند، یک فضای ویژگی ترکیبی ایجاد کرد که هم شامل اطلاعات زمانی-مکانی جاری و هم الگوهای بلندمدت محیطی است. این کار باعث افزایش درک زمینه‌ای مدل نسبت به شرایط سیستم در طول زمان می‌شود. مهم‌تر اینکه، هیچ ویژگی مکانی به کار گرفته نشد که هدف اصلی این معماری یعنی ایجاد مدلی مستقل از مکان و قابل‌تعمیم را تضمین می‌کند.

۴-۶- طبقه بند MLP با مکانیزم توجه

برای انجام‌وظیفه طبقه‌بندی بر روی بردارهای ویژگی غنی‌شده حاصل از فاز انتقال دانش، یک معماری چندلایه پرسپترون طراحی شد که در آن چندین تکنیک نوین برای افزایش دقت، پایداری و تعمیم‌پذیری مدل به‌کار گرفته شده‌اند. این معماری شامل اجزای کلیدی زیر است:

۱) اتصالات باقیمانده

یکی از مشکلات متداول در شبکه‌های عمیق، پدیده محو شدن گرادیان (Vanishing Gradients) است؛ حالتی که در آن طی فرایند Backpropagation، مقدار گرادیان در لایه‌های اولیه بسیار کوچک‌شده و عملاً یادگیری متوقف می‌شود.

برای حل این مشکل، از بلوک‌های باقیمانده استفاده شد که نخستین بار در شبکه‌های ResNet معرفی شده‌اند. در این ساختار، خروجی لایه غیرخطی به‌صورت زیر تعریف می‌شود:

$$Y = F(X) + X \qquad (۸)$$

که در آن x ورودی اولیه به بلوک است، $F(x)$ تابعی شامل لایه‌های غیرخطی مانند Dense و BatchNorm است و y خروجی نهایی بلوک است.

با این طراحی، حتی در صورت ناتوانی لایه‌های میانی در یادگیری بهینه، اطلاعات اولیه ورودی می‌تواند مستقیماً به لایه‌های بالاتر منتقل شود. این ویژگی باعث تسهیل یادگیری عمیق‌تر، افزایش سرعت همگرایی و جلوگیری از تخریب اطلاعات مفید اولیه می‌شود.

۲) مکانیزم توجه

در بسیاری از مسائل با تعداد ویژگی‌های زیاد و روابط غیرخطی پیچیده، استفاده از یک مکانیزم تطبیقی برای شناسایی و تمرکز بر ویژگی‌های مهم‌تر ضروری است. مکانیزم توجه با الهام از ساختارهای مغزی انسان، به مدل

اجازه می‌دهد تا در هر نمونه ورودی، به ویژگی‌هایی که برای پیش‌بینی مهم‌تر هستند، وزن بیشتری اختصاص دهد. در این پروژه، از مکانیزم توجه در سطح ویژگی استفاده‌شده است. روابط ریاضی آن به‌صورت زیر قابل‌تعریف است:

مرحله ۱: محاسبه وزن توجه برای هر ویژگی

$$e_i = \tanh\left(w_{x_i} + b\right) \quad (۹)$$

$$\alpha_i = \frac{\exp\left(e_i\right)}{\sum_{j=1}^{d} \exp\left(e_j\right)} \quad (۱۰)$$

که در آن x_i مقدار ویژگی i در بردار ورودی است، W و b پارامترهای قابل یادگیری هستند، e_i نمره خام توجه برای ویژگی i است و α_i وزن توجه نهایی پس از نرمال‌سازی softmax است.

مرحله ۲: وزن‌دهی تطبیقی ویژگی‌ها

$$z = \sum_{i=1}^{d} \alpha_i x_i \quad (۱۱)$$

که در آن z بردار نهایی است که ویژگی‌ها در آن با وزن توجه خود ترکیب شده‌اند. این بردار، ورودی به لایه‌های بعدی شبکه خواهد بود.

مزیت اصلی این مکانیزم آن است که برخلاف MLP های کلاسیک که به همه ویژگی‌ها وزن مساوی می‌دهند، مدل می‌تواند بسته به هر نمونه، تمرکز خود را به‌طور پویا بر روی ویژگی‌های حساس‌تر قرار دهد. این موضوع به‌ویژه در داده‌های پیچیده مانند خاموشی برق که به شرایط اقلیمی و غیرخطی‌های متعدد وابسته‌اند، بسیار مؤثر است.

۳) تابع فعال‌سازی Swish

تابع فعال‌سازی انتخاب‌شده برای لایه‌های میانی، Swish است که توسط تیم Google Brain پیشنهادشده و نسبت به توابع کلاسیکی مانند ReLU، رفتار نرم‌تر و گرادیان پایدارتری دارد. فرم ریاضی آن به‌صورت زیر است:

$$Swish(x) = x.\sigma(x) \quad (۱۲)$$

که در آن $\sigma(x) = \frac{1}{1+e^{-x}}$، تابع سیگموئید است.

Swish در نواحی منفی غیر صفر باقی می‌ماند (برخلاف ReLU)، بنابراین جریان گرادیان قطع نمی‌شود. همچنین مشتق‌پذیر و پیوسته است که به بهبود همگرایی کمک می‌کند. در مسائل پیچیده و noisy مانند پیش‌بینی قطعی برق، طبق نتایج تجربی در این پروژه عملکرد بهتری نسبت به ReLU دارد.

۴-۷- نتایج

در این بخش، نتایج حاصل از اجرای چارچوب پیشنهادی مبتنی بر یادگیری انتقالی و شبکه‌های عصبی پیشرفته برای پیش‌بینی سطح خاموشی برق ارائه و تحلیل می‌گردد. همان‌طور که در بخش‌های قبل بیان شد، رویکرد اصلی این پژوهش بر استخراج ویژگی‌های عمیق از داده‌های بدون برچسب (Dataset B) با استفاده از خودرمزگذار و انتقال آن‌ها به داده‌های برچسب‌دار هدف (Dataset A) جهت طبقه‌بندی دقیق سطوح خاموشی متمرکز بود. برای ارزیابی دقیق عملکرد مدل‌ها و کاهش تأثیر تصادفی بودن تقسیم‌بندی داده‌ها، از روش K-Fold Cross Validation با مقدار k=5 استفاده شد. در این روش داده‌ها به ۵ بخش (Fold) تقسیم می‌شوند؛ در هر تکرار، یکی از بخش‌ها به‌عنوان داده آزمون (Test) در نظر گرفته‌شده و چهار بخش دیگر برای آموزش (Train) استفاده می‌شوند؛ این فرآیند ۵ بار تکرار شده و میانگین نتایج حاصل به‌عنوان معیار نهایی گزارش می‌شود. هفت مدل مختلف در سه رویکرد اصلی برای طبقه‌بندی نهایی مورد ارزیابی قرار گرفتند:

۴-۷-۱- رویکرد اول: یادگیری انتقالی با خودرمزگذار و طبقه‌بندهای پیشرفته

در این رویکرد، از خودرمزگذار به‌عنوان ابزاری برای یادگیری نمایش‌های معنایی از داده‌های اقلیمی بلندمدت و بدون برچسب استفاده شد. این نمایش‌ها، پس از آموزش خودرمزگذار روی مجموعه Dataset B (داده‌های ۲۰۰۰ تا ۲۰۲۴)، به داده‌های برچسب‌دار Dataset A منتقل شدند و به‌عنوان ورودی مدل‌های طبقه‌بند استفاده شدند.

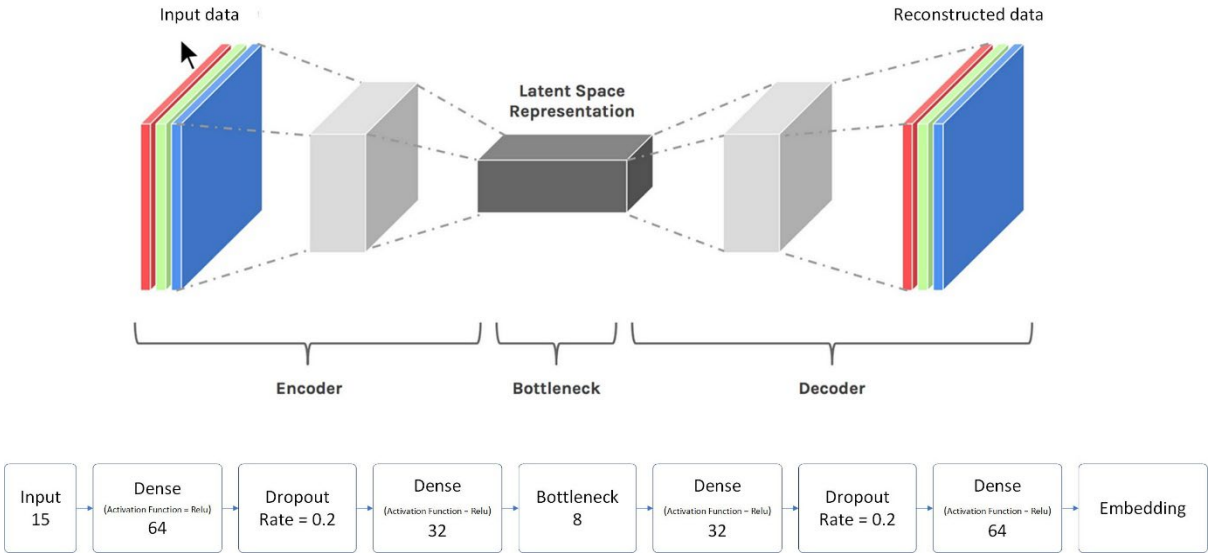
در گام نخست، برای استخراج نمایش‌های فشرده و معنادار از داده‌ها، از یک خودرمزگذار عمیق استفاده شد. در بخش رمزگذار (Encoder) این شبکه، داده‌های ورودی که شامل ویژگی‌های هواشناسی، متغیرهای زمانی و شاخص‌های مکانی بودند، ابتدا وارد یک لایه کاملاً متصل با ۶۴ نرون و تابع فعال‌ساز ReLU شدند. به‌منظور

جلوگیری از بیش‌برازش، یک لایه Dropout با نرخ ۰/۲ به آن افزوده شد. سپس یک لایه ۳۲ نرونی با ReLU اعمال شد تا ابعاد داده بیش‌تر کاهش یابد. درنهایت، داده‌ها به یک لایه میانی (Bottleneck) با ۸ نرون منتقل شدند. این لایه نقش نمایش فشرده داده‌ها را ایفا می‌کند و اساساً یک بردار کم‌بعد از اطلاعات اصلی را در خود جای می‌دهد.

در بخش رمزگشا (Decoder)، مسیر معکوس برای بازسازی داده‌ها طی شد. این بخش شامل لایه‌هایی با ۳۲ و ۶۴ نرون و فعال‌ساز ReLU به همراه Dropout بود و در انتها یک لایه خروجی با ابعاد برابر داده‌های اولیه و فعال‌ساز خطی قرار گرفت. برای آموزش خودرمزگذار از تابع هزینه خطای میانگین مربعات (MSE) و بهینه‌ساز Adam با نرخ یادگیری ۰/۰۰۱ استفاده شد. همچنین، برای جلوگیری از بیش‌برازش، مکانیزم توقف زودهنگام (Early Stopping) با معیار خطای اعتبارسنجی و صبر ده دوره به‌کار گرفته شد. بدین ترتیب، خروجی رمزگذار یک بردار هشت‌بعدی از هر نمونه داده بود که در مراحل بعدی به‌عنوان ویژگی نهفته (Embedding) مورداستفاده قرار گرفت.

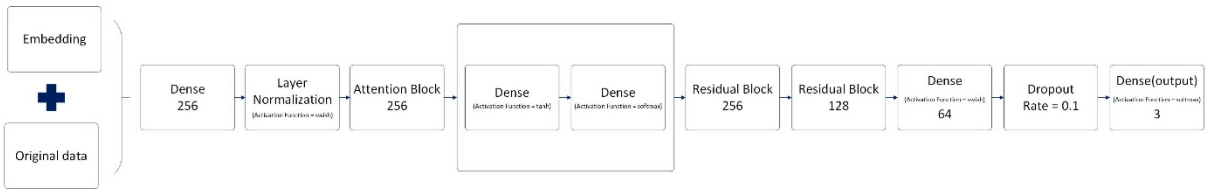
پس‌ازاین مرحله، داده‌های مجموعه اصلی با استفاده از مقیاس‌گذار استاندارد (StandardScaler) نرمال‌سازی شدند و سپس با بردارهای نهفته استخراج‌شده ترکیب گردیدند. بدین ترتیب، بردار ورودی نهایی مدل شامل دو بخش بود: بخش نخست ویژگی‌های اصلی مجموعه داده و بخش دوم نمایش‌های فشرده‌شده حاصل از خودرمزگذار.

شبکه اصلی منطبق بر شکل **۶** و **۷** که برای طبقه‌بندی خاموشی‌ها طراحی شد، بر پایه یک شبکه عصبی چندلایه با بلوک‌های باقی‌مانده و مکانیزم توجه بنا گردید. داده‌های ترکیبی ابتدا وارد یک لایه کاملاً متصل با ۲۵۶ نرون شدند. این لایه پس از نرمال‌سازی و عبور از تابع فعال‌ساز Swish، به یک بلوک توجه هدایت گردید. مکانیزم توجه ویژگی‌محور با استفاده از دو لایه Dense پی‌درپی و فعال‌سازهای tanh و softmax پیاده‌سازی شد. این مکانیزم امکان می‌دهد شبکه بر ویژگی‌های مهم‌تر تمرکز کرده و وزن بیشتری به آن‌ها اختصاص دهد.



شکل ۶. ساختار خودرمزگذار استفاده شده برای استخراج ویژگی های تاریخی.

Figure 6. Structure of the Autoencoder Used for Extracting Historical Features.



شکل ۷. ساختار کلی مدل استفاده‌شده برای آموزش و طبقه‌بندی داده‌ها.

Figure 7. Structure of the model used for training and classifying data.

در ادامه، دو بلوک باقیمانده به‌ترتیب با ۲۵۶ و ۱۲۸ نرون اعمال شدند. هر بلوک شامل چندین لایه Dense، نرمال‌سازی، فعال‌ساز Swish و Dropout بود و یک اتصال میان‌بر (Skip Connection) نیز داشت که باعث می‌شد اطلاعات اولیه مستقیماً به خروجی بلوک اضافه شود. این معماری باقی‌مانده به جلوگیری از مشکل ناپدید شدن گرادیان کمک کرده و پایداری آموزش شبکه را افزایش داد. پس‌ازاین بلوک‌ها، یک لایه Dense با ۶۴ نرون و فعال‌ساز Swish قرار گرفت که به همراه Dropout با نرخ ۰/۱ نقش فشرده‌سازی نهایی را ایفا کرد. لایه خروجی شبکه نیز یک لایه کاملاً متصل با سه نرون و تابع فعال‌ساز Softmax بود که وظیفه طبقه‌بندی داده‌ها در سه دسته بدون خاموشی، خاموشی جزئی و خاموشی گسترده را بر عهده داشت.

برای آموزش این مدل، از بهینه‌ساز Adam با نرخ یادگیری اولیه ۰/۰۰۱ استفاده شد. تابع هزینه به‌کاررفته آنتروپی متقاطع دسته‌ای (Categorical Crossentropy) است که به‌طور خاص برای مسائل طبقه‌بندی چندکلاسه مناسب است. به‌منظور جلوگیری از بیش‌برازش، از توقف زودهنگام با صبر پانزده دوره استفاده گردید و برای تنظیم تطبیقی نرخ یادگیری، مکانیزم ReduceLROnPlateau به‌کار گرفته شد که در صورت عدم بهبود عملکرد مدل طی پنج دوره، نرخ یادگیری را نصف کرده و آن را تا حداقل 10^{-5} کاهش می‌داد. مدل با دسته‌های آموزشی ۲۵۶ تایی و حداکثر ۳۰۰ دوره تمرین داده شد که در عمل، با توجه به توقف زودهنگام، تعداد واقعی دوره‌های آموزش کمتر از این مقدار بود.

علاوه بر شبکه عصبی غنی‌شده پیشنهادی، از مدل TabNet نیز به‌عنوان یک رویکرد مقایسه‌ای بهره گرفته شد. TabNet یک معماری عمیق مبتنی بر مکانیزم توجه ترتیبی است که به‌طور ویژه برای داده‌های جدولی طراحی شده است. در این مدل، فرآیند انتخاب ویژگی‌ها به‌صورت مرحله‌ای انجام می‌شود و در هر گام، توجه شبکه بر مجموعه‌ای از ویژگی‌های مهم متمرکز می‌گردد. بدین ترتیب، شبکه نه‌تنها قادر به یادگیری نمایش‌های غیرخطی از داده‌هاست، بلکه تفسیرپذیری بالاتری نسبت به شبکه‌های کلاسیک نیز دارد. برای پیاده‌سازی، پارامترهای اصلی مدل شامل تعداد واحدهای تصمیم و توجه (۳۲)، تعداد گام‌های تصمیم‌گیری (۵)، ضریب تنظیم پراکندگی ($\lambda_{\text{sparse}} = 10^{-4}$) و نرخ یادگیری اولیه ۰/۰۲ در نظر گرفته شد. آموزش مدل با استفاده از بهینه‌ساز Adam و زمان‌بندی‌کننده StepLR به مدت حداکثر ۲۰۰ دوره انجام گرفت و برای جلوگیری از بیش‌برازش، از مکانیزم توقف زودهنگام با patience بیست دوره استفاده شد. Ensemble ترکیبی، با وزن‌دهی به پیش‌بینی‌های دو مدل، پیش‌بینی نهایی را انجام می‌دهد. نتایج مدل‌های این رویکرد در جدول **۲** قابل‌مشاهده است.

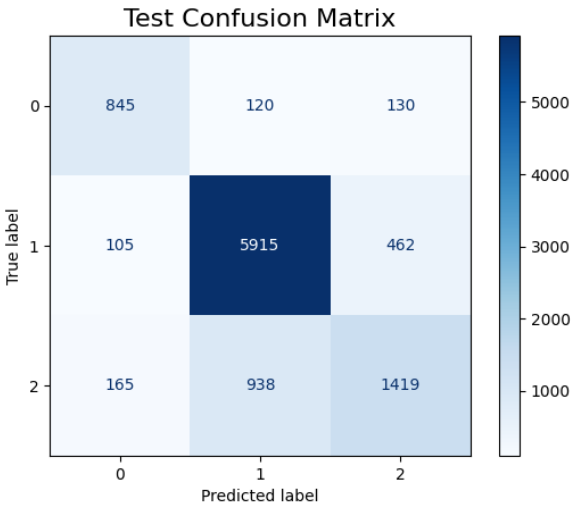
به‌منظور ارزیابی پایداری و اطمینان‌پذیری مدل پیشنهادی، صرفاً به گزارش مقدار میانگین امتیاز F1 اکتفا نشده و میزان پراکندگی نتایج نیز موردبررسی قرار گرفته است.

گزارش انحراف معیار (Std ±) این امکان را فراهم می‌سازد که میزان حساسیت مدل نسبت به تقسیم‌بندی داده‌ها مشخص شود. انحراف معیار پایین نشان‌دهنده پایداری مدل و عدم وابستگی شدید عملکرد آن به یک تقسیم خاص از داده‌هاست. به‌صورت کلی مدل MLP با ساختار توجه-باقیمانده، باوجود حذف ویژگی‌های مکانی، توانست وابستگی زمانی، اقلیمی و الگوهای منطقه‌ای را از طریق داده‌های بلندمدت رمزگذاری‌شده توسط خودرمزگذار بهتر یاد بگیرد. نتایج عملکرد این مدل از طریق ماتریس آشفتگی و نمودار AUC-ROC که در شکل **۸** و شکل **۹** قابل‌مشاهده است.

جدول ۲. بررسی نتایج مدل‌های Tabnet، MLP، MlP+Tabnet با خودرمزگذار

Table 2. Results of Tabnet, MLP, MLP+Tabnet models with autoencoder

مدل	F1(بدون خاموشی)	F1(جزئی)	F1(کلی)	F1کل
MLP + Residual + Attention	۰/۸۰	۰/۸۵	۰/۷۵	۰/۸۱±۰/۰۰۸
TabNet	۰/۷۸	۰/۸۴	۰/۷۰	۰/۷۹±۰/۰۰۵
Ensemble (MLP + TabNet)	۰/۷۵	۰/۸۷	۰/۶۰	۰/۸۰±۰/۰۰۴



شکل ۸. ماتریس آشفتگی کلاس تست.

Figure 8. Test confusion matrix.

۴-۷-۲ رویکرد دوم: CNN + ML با ترکیب رأی‌گیری وزن‌دار

در این رویکرد، CNN به‌عنوان استخراج‌کننده ویژگی عمل کرده و خروجی‌های آن به مدل‌های سنتی ماشین یادگیری منتقل شدند. سپس خروجی مدل‌ها در یک سیستم رأی‌گیری نرم وزن‌دار (Hybrid Soft Voting) ادغام شدند.

مراحل کلی این رویکرد به شرح زیر بوده است:

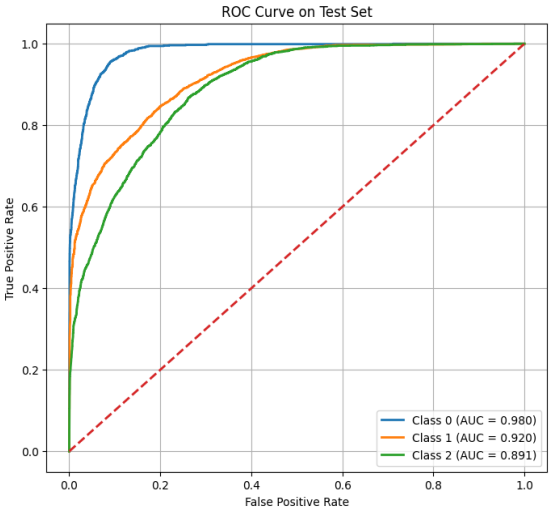
۱) استخراج ویژگی از CNN: خروجی لایه‌های میانی CNN به‌عنوان ورودی به مدل‌های XGBoost، RF، LR و Naive Bayes داده شد.

۲) استفاده از soft outputs: احتمالات نرم CNN نیز در کنار سایر مدل‌ها برای تصمیم نهایی استفاده شدند.

۳) رأی‌گیری وزن‌دار بهینه: وزن مدل‌ها از طریق بهینه‌سازی بیزی (Bayesian Optimization) تنظیم شد.

۴) هرس مدل‌ها (Ensemble Pruning): ترکیب فقط سه مدل برتر CNN + برای کاهش پیچیدگی و افزایش دقت.

بر طبق جدول **۳**، استفاده از CNN به‌عنوان استخراج‌کننده ویژگی باعث شد که مدل‌های کلاسیک ML بتوانند الگوهای پیچیده را راحت‌تر درک کنند. همچنین استفاده از رأی‌گیری نرم وزن‌دار باعث شد تا



شکل ۹. نمودار ROC روی داده تست.

Figure 9. ROC Curve on test set.

نتایج مدل‌هایی با خطای جزئی نیز در پیش‌بینی نهایی سهم مناسب داشته باشند.

جدول ۳. نتایج مدل های مبتنی بر CNN

Table 3. Results of CNN-based models

مدل	F1 بدون خاموشی	جزئی	کلی	کل
CNN + ML Hybrid	۰/۷۲	۰/۸۶	۰/۵۹	۰/۷۸±۰/۰۰۳
CNN + MLP	۰/۶۸	۰/۸۵	۰/۶۰	۰/۷۵±۰/۰۰۶

۴-۷-۳ رویکرد سوم: عملکرد روش پیشنهادی در مقایسه با روش‌های پایه‌ای

به‌منظور تحلیل دقیق اثر هر یک از مؤلفه‌های معماری پیشنهادی و جلوگیری از تفسیر نادرست بهبود عملکرد مدل نهایی، مجموعه‌ای از آزمایش‌های فرسایش بر روی ساختارهای پایه انجام شد (جدول **۴**). در این آزمایش‌ها، عملکرد یک MLP ساده به‌عنوان مدل مرجع با نسخه‌های توسعه‌یافته مقایسه می‌شوند. نتایج حاصل از اعتبارسنجی متقاطع پنج‌بخشی نشان داد که مدل MLP پایه بدون هیچ‌گونه مکانیزم کمکی، به امتیاز F1=۰/۷۱۸ ± ۰/۰۰۲ دست یافته است. این مقدار به‌عنوان سطح عملکرد پایه‌ای در نظر گرفته شد که بیانگر توان محدود MLP کلاسیک در مدل‌سازی روابط غیرخطی پیچیده و وابستگی‌های بلندمدت موجود در داده‌های خاموشی برق است. در گام بعد، مکانیزم توجه به‌صورت مستقل به ساختار MLP

افزوده شد. برخلاف انتظار اولیه، نتایج نشان داد که استفاده صرف از مکانیزم توجه بدون سایر اجزای کمکی، منجر به کاهش عملکرد کلی مدل شده است. این کاهش عملکرد بیانگر آن است که مکانیزم توجه، در غیاب ساختارهای پایدارکننده‌ای مانند اتصالات باقی‌مانده یا نمایش‌های نهفته غنی، می‌تواند باعث بیش‌تمرکز یا نوسان در وزن‌دهی ویژگی‌ها شده و درنتیجه یادگیری را ناپایدار کند. افزایش انحراف معیار نیز مؤید حساسیت بالاتر این ساختار نسبت به تقسیم‌بندی داده‌هاست. افزودن اتصالات باقیمانده به MLP پایه، بدون استفاده از خودرمزگذار یا مکانیزم توجه، منجر به بهبود عملکرد مدل می‌شود. این نتیجه نشان می‌دهد اتصالات باقیمانده نقش کلیدی در تسهیل جریان گرادیان، حفظ اطلاعات اولیه و پایدارسازی فرآیند یادگیری در شبکه‌های عمیق ایفا می‌کنند و حتی به‌تنهایی می‌توانند توان مدل را در استخراج الگوهای مفید افزایش دهند.

جدول ۴. نتایج آزمون فرسایش

Table 4. Ablation study results

مدل	F1
MLP پایه	۰/۷۱۸±۰/۰۰۲
MLP با ساختار توجه	۰/۷۰۳±۰/۰۰۷
MLP با اتصالات باقیمانده	۰/۷۴۷±۰/۰۰۳
مدل پیشنهادی بدون ویژگی‌های جغرافیایی	۰/۸۱۳±۰/۰۰۸
مدل پیشنهادی با ویژگی‌های جغرافیایی	۰/۹۰۱±۰/۰۰۷

بررسی این نتایج نشان می‌دهد که تأثیر مؤلفه‌های معماری به‌شدت وابسته به نحوه ترکیب آن‌هاست و استفاده مجزای برخی اجزا، مانند مکانیزم توجه، لزوماً منجر به بهبود عملکرد نمی‌شود. این یافته‌ها اهمیت طراحی یکپارچه و هماهنگ معماری پیشنهادی را برجسته می‌سازد و توجیه می‌کند که چرا بهترین عملکرد نهایی تنها در صورت ترکیب هم‌زمان خودرمزگذار، اتصالات باقی‌مانده و مکانیزم توجه حاصل شده است. مدل پیشنهادی نهایی که از ترکیب خودرمزگذار برای استخراج نمایش‌های فشرده و مستقل از مکان، معماری MLP دارای اتصالات باقیمانده، مکانیزم توجه بهره می‌برد، به دقت مناسبی (۰/۸۱۳ ±۰/۰۰۸) دست یافته است. این بهبود قابل‌توجه نسبت به تمامی مدل‌های پایه نشان می‌دهد که هم‌افزایی بین یادگیری نمایش، معماری پایدار و مدیریت عدم‌توازن کلاس‌ها نقش تعیین‌کننده‌ای در ارتقای دقت و پایداری پیش‌بینی داشته است. علاوه بر این، مقدار انحراف معیار محدود در مدل پیشنهادی بیانگر پایداری مناسب عملکرد در فولدهای مختلف و قابلیت تعمیم‌پذیری بالاتر مدل در مقایسه با سایر معماری‌ها است. در کل و بر طبق جدول **۴**، استفاده از خودرمزگذار و انتقال دانش نقش مهمی در کارکرد روش پیشنهادی دارد به صورتی که عدم استفاده از آن منجر به کاهش نتایج می‌گردد. از طرف دیگر ویژگی جغرافیایی نیز در نتایج مدل تأثیرگذار است به‌نحوی‌که حذف آن دقت مدل را در حدود ۹ درصد کاهش می‌دهد، ولی روش پیشنهادی حتی بدون استفاده از این ویژگی نیز دقت قابل قبولی دارد که کارکرد آن برای بسیاری از مجموعه داده‌ها و شرایطی که اطلاعات جغرافیایی دقیقی از محل رخداد قطعی وجود ندارد ساده می‌نماید.

در ادامه در شکل **۱۰** به بررسی عملکرد مدل‌های یادگیری عمیق آزمایش‌شده در روند توسعه این پروژه با روش پیشنهادی می‌پردازیم. نتایج تجربی نشان داد که رویکرد اول، متشکل از خودرمزگذار و طبقه بند MLP نسبت به سایر مدل‌های کلاسیک یا ترکیبی، از دقت، انسجام و تعمیم‌پذیری بسیار بالاتری برخوردار است. در چارچوب پیشنهادی، ابتدا یک خودرمزگذار بر روی داده‌های اقلیمی بلندمدت (۲۰۰۰ تا ۲۰۲۴) آموزش داده شد. این مدل، بدون نیاز به برچسب، موفق شد نمایش‌های معنایی فشرده‌ای از الگوهای بلندمدت را از متغیرهای آب‌وهوایی استخراج کند. سپس این نمایش‌ها به داده‌های برچسب‌دار هدف (Dataset A) منتقل شدند و نقش مؤثری در غنی‌سازی اطلاعات ورودی مدل‌های طبقه‌بند نهایی ایفا کردند.

مزایای روش پیشنهادی به‌صورت خلاصه به شرح زیر است:

- کاهش وابستگی جغرافیایی: با حذف ویژگی‌های lat, lon و area، مدل کاملاً از وابستگی به منطقه خاص رها شد.
- توجه به روابط بلندمدت اقلیمی: استفاده از خودرمزگذار، الگوهای پنهان در داده‌های تاریخی را استخراج کرده و آن‌ها را در اختیار طبقه‌بندها قرار داد.
- افزایش دقت طبقه‌بندی، به‌ویژه در کلاس‌های اقلیت: برخلاف مدل‌های متداول که در

تشخیص خاموشی‌های شدید دقت پایینی داشتند، این روش در طبقه‌بندی این کلاس حساس، عملکرد قابل‌توجهی داشت.

- تعمیم‌پذیری بالا به سایر مناطق: با توجه به اینکه مدل نهایی تنها از اطلاعات اقلیمی و زمانی استفاده می‌کند و به مختصات جغرافیایی وابسته نیست، قابلیت استفاده مستقیم برای پیش‌بینی خاموشی در مناطق جدید و ناشناخته را دارد.

در مقایسه، سایر رویکردهای مبتنی بر ترکیب CNN با مدل‌های ماشین یادگیری، یا مدل‌های منفرد مانند TabNet و MLP خام، با وجود ساختارهای بهینه و تنظیمات دقیق، حداکثر تا F1-score = ۰/۷۹ عملکرد داشتند که در هر سه کلاس از مدل نهایی پیشنهادی ضعیف‌تر ظاهر شدند.

۴-۸-۳- نقش در افزایش کارایی سیستم

رویکرد ارائه‌شده در این پژوهش شامل مراحل زیر بود:

۱) پاک‌سازی و نرمال‌سازی داده‌ها با استفاده از تکنیک‌هایی مانند KNN Imputation، StandardScaler، حذف داده‌های تکراری و تحلیل صدکی داده‌های پرت

۲) یادگیری بدون نظارت با خودرمزگذار روی Dataset B جهت استخراج ویژگی‌های فشرده و غنی از داده‌های اقلیمی ۲۴ ساله

۳) انتقال این بردارهای فشرده به Dataset A و الحاق آن‌ها به ویژگی‌های اصلی برای غنی‌سازی فضای ورودی

۴) حذف کامل ویژگی‌های مکانی (مانند طول و عرض جغرافیایی) جهت افزایش قابلیت تعمیم به سایر مناطق

۵) طراحی معماری طبقه‌بندی MLP پیشرفته شامل مکانیزم توجه، اتصالات باقیمانده، swish و focal loss جهت مقابله با چالش‌های داده‌های پیچیده و نامتوازن

۶) ارزیابی مدل‌ها با K-Fold Cross Validation جهت دستیابی به نتایج پایدار و قابل‌تعمیم

ترکیب خودرمزگذار برای یادگیری نمایش‌های معنایی عمیق از داده‌های بدون برچسب، با معماری MLP تطبیقی و مدرن، چارچوبی قدرتمند و قابل تعمیم برای پیش‌بینی سطوح خاموشی برق ایجاد کرد. این چارچوب حتی در نبود اطلاعات مکانی نیز عملکرد قابل قبولی ارائه داد که آن را برای استقرار در مناطق جدید و داده‌های ناقص، ایده‌آل می‌سازد. همچنین، عملکرد مطلوب در شناسایی خاموشی‌های شدید (که معمولاً در اقلیت هستند)، نشان از توانمندی مدل در تحلیل شرایط بحرانی واقعی دارد.

این روش نه‌تنها امکان استقرار سریع در مناطق جدید بدون نیاز به آموزش مجدد را فراهم می‌کند، بلکه به درک بهتر تعاملات نادر بین آب‌وهوا و خاموشی‌ها کمک کرده و بار عملیاتی را در مواقع بحرانی کاهش می‌دهد. از این منظر، می‌توان نتیجه‌گیری کرد که رویکرد یادگیری انتقالی مبتنی بر خودرمزگذار به همراه طبقه‌بند پیشرفته MLP، به‌عنوان روش نهایی و پیشنهادی این پژوهش معرفی می‌شود. همچنین یک تصمیم استراتژیک در طراحی سیستم‌های مقاوم، هوشمند و مقیاس‌پذیر انرژی آینده به‌شمار می‌آید.

۵- نتیجه‌گیری و تحقیقات آتی

۵-۱- نتیجه‌گیری

هدف این پژوهش، توسعه یک مدل پیش‌بینی احتمال خاموشی برق با قابلیت تعمیم‌پذیری بالا بود تا بر محدودیت‌های مدل‌های سنتی وابسته به ویژگی‌های مکانی غلبه کند. ما یک چارچوب یادگیری عمیق دومرحله‌ای را معرفی کردیم که شامل یک خودرمزگذار برای استخراج ویژگی‌های تعمیم‌پذیر و یک شبکه MLP بهبودیافته با بلوک‌های باقیمانده و مکانیزم توجه برای طبقه‌بندی نهایی ریسک خاموشی بود. نتایج تجربی نشان داد که رویکرد یادگیری انتقالی از داده‌های بلندمدت و بدون برچسب، امکان تولید بردارهای ویژگی غنی و مکان-مستقل را فراهم ساخت. این چارچوب توانست با دقت کلی ۸۱ درصد، احتمال خاموشی‌ها را در سه دسته (ریسک کم، متوسط و بالا) پیش‌بینی کند، درحالی‌که وابستگی به ویژگی‌های جغرافیایی ایستا مانند طول و عرض جغرافیایی کاملاً حذف شده است. این دستاورد مؤید توانایی مدل در حفظ عملکرد بالا و افزایش پتانسیل به‌کارگیری آن در مناطق جغرافیایی جدید بدون نیاز به بازآموزی گسترده است. مدل پیشنهادی نه‌تنها به

پیش‌بینی بلادرنگ کمک می‌کند، بلکه با برجسته کردن اهمیت متغیرهای اقلیمی و تقاضای انرژی، ابزاری ارزشمند برای برنامه‌ریزی تاب‌آوری شبکه برق در مواجهه با تغییرات آب و هوایی فراهم می‌آورد.

۵-۲- پیشنهادات برای تحقیقات آتی

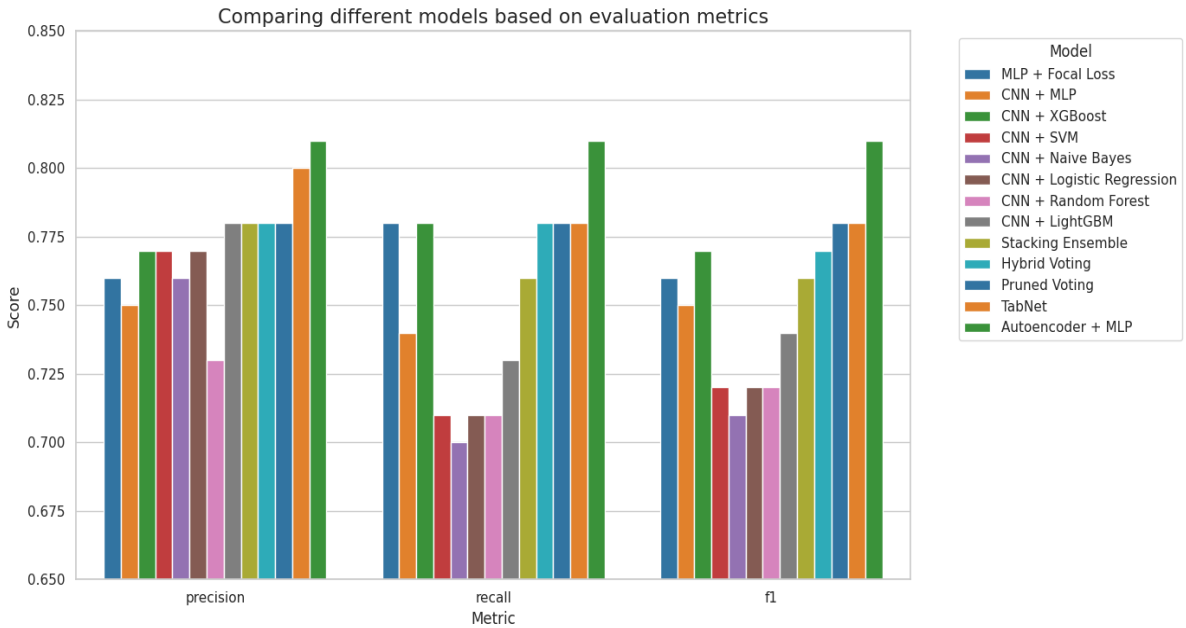
بر اساس نتایج و محدودیت‌های موجود در این پژوهش، مسیرهای زیر برای توسعه و ارتقاء مدل‌های پیش‌بینی خاموشی در آینده پیشنهاد می‌شوند:

۱) مدل‌سازی فضایی-زمانی پیشرفته: باوجود حذف وابستگی به مختصات جغرافیایی ساده، می‌توان از ویژگی‌های استخراج‌شده توسط خودرمزگذار در کنار شبکه‌های عصبی گراف استفاده کرد. این رویکرد می‌تواند روابط توپولوژیکی بین پست‌های برق و خطوط انتقال را مدل‌سازی کرده و دقت پیش‌بینی را در سطح محلی بهبود بخشد، بدون اینکه صرفاً به مختصات وابسته باشد.

۲) ادغام داده‌های بلادرنگ: تحقیق در مورد سازگاری مدل با داده‌های بلادرنگ (Real-Time Data Streams) از حسگرهای پیشرفته، کنتورهای هوشمند و سیستم‌های نظارت بر سلامت تجهیزات ضروری است. همچنین، می‌توان بهینه‌سازی مدل برای استقرار و اجرای سریع آن در دستگاه‌های لبه را جهت فعال‌سازی قابلیت پیش‌بینی‌های لحظه‌ای بررسی کرد.

۳) مدل‌سازی ریسک رخدادهای اقلیمی ترکیبی (Compound Events): تمرکز ویژه بر پیش‌بینی تأثیر رخدادهای اقلیمی ترکیبی که احتمال وقوع هم‌زمان آن‌ها رو به افزایش است (مانند موج گرما همراه با باد شدید یا طوفان). مدل‌سازی این پدیده‌ها باید به‌صورت صریح انجام شود، زیرا تأثیر آن‌ها بر سیستم قدرت معمولاً چندبرابر است نه صرفاً افزایشی.

۴) تحقیقات بر روی تفسیرپذیری و شفافیت مدل (XAI): با توجه به اینکه پیش‌بینی‌های مدل‌های یادگیری عمیق اغلب شبیه به یک جعبه سیاه هستند، استفاده از تکنیک‌های هوش مصنوعی توضیح‌پذیر مانند SHAP یا LIME برای تحلیل خروجی مدل و درک اینکه چرا یک منطقه خاص در خطر بالای خاموشی قرار دارد، می‌تواند اعتماد اپراتورهای سیستم انرژی را به نتایج مدل افزایش داده و توجیه‌پذیری تصمیمات عملیاتی را بهبود بخشد.



شکل ۱۰. نتایج روش‌های آزمایش‌شده یادگیری عمیق.

Figure 10. Results of the deep learning methods.

مشارکت‌های نویسندگان

همه نویسندگان در مقاله نقش و سهم یکسان داشته اند.

تضاد منافع

نویسندگان اعلام می‌کنند که هیچ‌گونه تضاد منافع مرتبط با تحقیق حاضر ندارند و نتایج به‌صورت بی‌طرفانه و بدون دخالت منافع شخصی یا حرفه‌ای به‌دست‌آمده است.

قدردانی

نویسندگان از عوامل اجرایی نشریه مهندسی سیستم و بهره وری و همچنین داوران محترم که کیفیت این مقاله را افزایش دادند، قدردانی مینمایند.

مراجع

Arifeen, M., Petrovski, A., Hasan, M. J., Noman, K., Navid, W. U., & Haruna, A. (2024). Graph-Variational Convolutional Autoencoder-Based Fault Detection and Diagnosis for Photovoltaic Arrays. *Machines*, *12*(12), 894. https://doi.org/10.3390/machines12120894

Atrigna, M., Buonanno, A., Carli, R., Cavone, G., Scarabaggio, P., Valenti, M., ... & Dotoli, M. (2023). A machine learning approach to fault prediction of power distribution grids under heatwaves. *IEEE Transactions on Industry Applications*, *59*(4), 4835-4845. https://doi.org/10.1109/TIA.2023.3262230

Avazpour, M., Zarei, J., & Alinajad, E. (2025). Evaluation and prioritization of electricity generation technologies in Iran using a multi-criteria decision-making approach. *Systems Engineering and Productivity, 5*(3), 179–198 (In Persian). https://doi.org/10.22034/sep.2025.2063697.1333

Azzalini, D., Flammini, B., Emanuele, C. A., Guadagno, A., Ragaini, E., & Amigoni, F. (2025). An empirical evaluation of deep autoencoders for anomaly detection in the electricity consumption of buildings. *Energy and Buildings*, *327*, 115069. https://doi.org/10.1016/j.enbuild.2024.115069

Campbell, R. J., & Lowry, S. (2012). Weather-related power outages and electric system resiliency. Washington, DC: Congressional Research Service, Library of Congress.

Coleman, J. (2022). Climate change made South Asian heatwave 30 times more likely. *Nature*. https://doi.org/10.1038/d41586-022-01444-1

Falabretti, D., Schiavo, L. L., Liotta, S., & Palazzoli, A. (2020). A novel method for evaluating the resilience of distribution networks during heat waves. *International Journal of Electrical and Electronic Engineering & Telecommunications, 9*(2), 73–79. https://doi.org/10.18178/ijeetc.9.2.73-79

Ghasemkhani, B., Kut, R. A., Yilmaz, R., Birant, D., Arıkök, Y. A., Güzelyol, T. E., & Kut, T. (2024). Machine learning model development to predict power outage duration (POD): A case study for electric utilities. *Sensors, 24*(13), 4313. https://doi.org/10.3390/s24134313

Ghosh, H. (2025). *Autoencoders and transfer learning*. Medium.

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

Namdarzadegan, M., & Bozorgi Amiri, A. (2025). Developing a hybrid sustainability–resilience model for implementing appropriate smart technologies in small and medium-sized enterprises under crisis conditions. *Systems Engineering and Productivity* (In Persian). https://doi.org/10.22034/sep.2025.2064025.1337

Nyame, S., Taylor, W. O., Hughes, W., Hong, M., Koukoula, M., Yang, F., ... & Cerrai, D. (2024). Transmission Failure Prediction Using AI and Structural Modeling Informed by Distribution Outages. *IEEE Access*, 13, 42-55. https://doi.org/10.1109/ACCESS.2024.3523415

Rahbaripour, K., Pakdelfard, M., Satarisarbanqoli, H., & Valizadeh, N. (2025). Analysis of barriers to achieving the Fourth Manufacturing Industry using the Interpretive Structural Modeling approach. *Systems Engineering and Productivity* (In Persian). https://doi.org/10.22034/sep.2025.2068090.1374

Razzaghy, S., Ahmadvand, A., Samadi Foroushani, M., (2024). Presenting a Dynamic Model of Iran's Electrical Energy Supply System Based on the Water-Food-Energy-Climate Change Nexus. *Systems Engineering and Productivity, 3*(4), 36-80. https://doi.org/10.22034/msb.2024.2022934.1183

Romitti, Y., & Sue Wing, I. (2022). Heterogeneous climate change impacts on electricity demand in world cities circa mid-century. *Scientific reports*, *12*(1), 4280. https://doi.org/10.1038/s41598-022-07922-w

Satterlee, K. (2024). Combining big data and machine learning to predict power outages and help consumers prepare. *Texas A&M University News*.

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

Stone, B., Jr., Gronlund, C. J., Mallen, E., Hondula, D., O'Neill, M. S., Rajput, M., Grijalva, S., Lanza, K., Harlan, S., Larsen, L., et al. (2023). How blackouts during heat waves amplify mortality and morbidity risk. *Environmental Science & Technology,* *57*(11), 8245–8255. https://doi.org/10.1021/acs.est.2c09588

Stone, B., Mallen, E., Rajput, M., Broadbent, A., Krayenhoff, E. S., Augenbroe, G., & Georgescu, M. (2021). Climate change and infrastructure risk: Indoor heat exposure during a concurrent heat wave and blackout event in Phoenix, Arizona. *Urban Climate,* *36*, 100787. https://doi.org/10.1016/j.uclim.2021.100787

Tabassum, T., Toker, O., & Khalghani, M. R. (2024). Cyber–physical anomaly detection for inverter-based microgrid using autoencoder neural network. *Applied Energy*, *355*, 122283. https://doi.org/10.1016/j.apenergy.2023.122283

Wang, X., Fatehi, N., Wang, C., & Nazari, M. H. (2024). Deep learning-based weather-related power outage prediction with socio-economic and power infrastructure data. *IEEE Power & Energy Society General Meeting (PESGM)* (pp. 1-5). IEEE. https://doi.org/10.1109/PESGM51994.2024.10688596

Zhu, S., Yao, R., Xie, Y., Qiu, F., Qiu, Y., & Wu, X. (2025). Quantifying grid resilience against extreme weather using large-scale customer power outage data. *INFORMS Journal on Data Science*. https://doi.org/10.1287/ijds.2023.0017

Zideh, M. J., & Solanki, S. K. (2024). Multivariate Physics-Informed Convolutional Autoencoder for Anomaly Detection in Power Distribution Systems with High Penetration of DERs. *arXiv preprint arXiv:2406.02927*. https://arxiv.org/abs/2406.02927