

Performance Evaluation of Machine Learning Algorithms for the Purpose of Building Thermal Load Prediction Using Exergy Analysis

Leila Gheisary, Hamed Peyrovedin*

M.Sc. Student in Chemical Engineering-Energy, School of Chemical and Petroleum Engineering, Shiraz University, Shiraz, Iran

Chemical Engineering Department, School of Chemical and Petroleum Engineering, Shiraz University, Shiraz, Iran

**Hamedpeyro@saadi.shirazu.ac.ir*

Abstract

With the increasing energy consumption of buildings, accurate prediction of heating and cooling loads has become essential for improving HVAC system performance and supporting efficient energy management. In this study, the performance of eight machine learning algorithms, including Random Forest, XGBoost, Gradient Boosting, Decision Tree, K-Nearest Neighbors (KNN), Support Vector Regression (SVR), Artificial Neural Network (ANN), and Linear Regression, was comprehensively evaluated using three independent datasets comprising a residential building (768 samples) and two office buildings located in the United States (64,792 samples) and Denmark (55,029 samples). The first dataset was employed to validate the proposed framework against a reference study. In addition, energy and exergy analyses were conducted to assess the building's energy performance. The results demonstrated that Random Forest consistently achieved the best performance, with coefficients of determination (R^2) exceeding 0.95 across all datasets. Feature importance analysis identified indoor temperature, outdoor temperature, relative humidity, and solar radiation as the most influential variables affecting thermal loads. The proposed framework provides a practical approach for thermal load prediction, building energy performance assessment, and intelligent energy management.

Keywords: Energy and Exergy Analyses, Air Conditioning, Random Forest, Dataset, Machine Learning's Algorithm

بررسی عملکرد الگوریتم‌های یادگیری ماشینی به منظور پیش‌بینی بار حرارتی ساختمان‌ها با تحلیل اکسرژی

لیلا قیصری، حامد پیرودین*

- ۱- دانشجوی کارشناسی ارشد مهندسی شیمی گرایش انرژی، دانشکده مهندسی شیمی، نفت و گاز، دانشگاه شیراز، شیراز، ایران.
- ۲- استادیار بخش مهندسی گاز، دانشکده مهندسی شیمی، نفت و گاز، دانشگاه شیراز، شیراز، ایران.

* Hamedpeyro@saadi.shirazu.ac.ir

چکیده

با افزایش مصرف انرژی ساختمان‌ها، پیش‌بینی دقیق بارهای گرمایشی و سرمایشی نقش کلیدی در بهبود عملکرد سیستم‌های تهویه مطبوع (HVAC) و مدیریت مصرف انرژی ایفا می‌کند. در این پژوهش، عملکرد هشت الگوریتم یادگیری ماشینی شامل ANN، SVR، KNN، Decision Tree، Gradient Boosting، XGBoost، Random Forest و Linear Regression برای پیش‌بینی بارهای گرمایشی و سرمایشی بر روی سه مجموعه داده مستقل شامل ساختمان مسکونی (۷۶۸ نمونه) و دو ساختمان اداری در آمریکا (۶۴۷۹۲ نمونه) و دانمارک (۵۵۰۲۹ نمونه) به طور جامع ارزیابی شد. مجموعه داده اول برای اعتبارسنجی با مقاله مرجع استفاده گردید. همچنین تحلیل انرژی و اکسرژی برای ارزیابی عملکرد انرژی ساختمان انجام شد. نتایج نشان داد الگوریتم Random Forest با ضریب تعیین بالاتر از ۰/۹۵ در تمامی مجموعه داده‌ها، عملکرد برتر و پایدارتری نسبت به سایر مدل‌ها داشت. تحلیل اهمیت ویژگی‌ها نیز دمای داخل، دمای خارج، رطوبت نسبی و تابش خورشیدی را به عنوان مؤثرترین پارامترهای تأثیرگذار بر بارهای حرارتی معرفی کرد. این پژوهش با تلفیق یادگیری ماشینی و تحلیل اکسرژی، چارچوبی کاربردی برای پیش‌بینی بارهای حرارتی، تحلیل عملکرد انرژی و پشتیبانی از تصمیم‌گیری در مدیریت هوشمند انرژی ساختمان‌ها ارائه می‌دهد.

واژگان کلیدی: آنالیز انرژی و اکسرژی، تهویه مطبوع، جنگل تصادفی، مجموعه داده، الگوریتم یادگیری ماشینی

۱- مقدمه

امروزه مدیریت مصرف انرژی به یکی از چالش‌های اساسی جوامع بشری تبدیل شده است [۱]. انرژی در بخش‌های مختلف زندگی بشر نقش مهمی ایفا می‌کند و به همین دلیل مدیریت آن از اهمیت ویژه‌ای برخوردار است [۲]. با رشد تقاضای انرژی، توسعه شهرنشینی و افزایش تعداد ساختمان‌ها، بخش ساختمان به یکی از بزرگ‌ترین مصرف‌کنندگان انرژی در جهان تبدیل شده است. به طوری که نزدیک به ۴۰ درصد از کل تقاضای انرژی را به خود اختصاص می‌دهد. در این میان، سیستم‌های تهویه مطبوع (HVAC) سهم عمده‌ای از مصرف انرژی ساختمان را تشکیل می‌دهند و به دلیل عملکرد مداوم، نقش تعیین‌کننده‌ای در بازده انرژی و هزینه‌های بهره‌برداری دارند [۱]. از این رو، بهبود عملکرد انرژی این سیستم‌ها، کاهش اتلاف انرژی و افزایش بهره‌وری ساختمان‌ها به یکی از محورهای اصلی پژوهش‌های سال‌های اخیر تبدیل شده است.

در سال‌های اخیر، استفاده از الگوریتم‌های یادگیری ماشین برای پیش‌بینی بارهای حرارتی و مدیریت انرژی ساختمان‌ها به طور چشمگیری گسترش یافته است [۳]. در همین راستا، اسرافیلیان نجف‌آبادی و همکاران در سال ۲۰۲۲ عملکرد الگوریتم‌های درخت تصمیم و KNN را برای پیش‌بینی مصرف انرژی بررسی کردند و نشان دادند که عملکرد این مدل‌ها به کیفیت داده‌ها و ویژگی‌های ورودی وابستگی زیادی دارد [۴]. همچنین، آلاراج و همکاران در سال ۲۰۲۳ با به کارگیری الگوریتم‌های یادگیری ماشین در ساختمان‌های تجاری، امکان کاهش حدود ۳۳ درصدی مصرف انرژی را گزارش کردند [۵]. در مطالعه دیگری، وانگ و همکاران در سال ۲۰۲۴ مدلی مبتنی بر یادگیری ماشین برای کنترل عملکرد سیستم‌های تهویه مطبوع (HVAC) ارائه کردند که موجب بهبود عملکرد انرژی سیستم و افزایش دقت پیش‌بینی بارهای حرارتی شد [۶]. همچنین، ارشاد و همکاران در سال ۲۰۲۴ با مقایسه چندین الگوریتم یادگیری ماشین نشان دادند که عملکرد مدل‌ها در شرایط مختلف عملیاتی یکسان نبوده و انتخاب الگوریتم مناسب نقش تعیین‌کننده‌ای در دقت پیش‌بینی دارد [۷]. در مطالعه‌ای دیگر، لی و همکاران در سال ۲۰۲۴ با استفاده از مدل‌های یادگیری ماشین برای کنترل مصرف انرژی سیستم HVAC در یک ساختمان مسکونی بدون سکنه پژوهشی انجام دادند و ضمن کاهش مصرف انرژی، بر لزوم انجام مطالعات بیشتر در شرایط واقعی تأکید نمودند [۸]. همچنین، کونگ و همکاران در سال ۲۰۲۵ کاربرد الگوریتم‌های یادگیری ماشین را در محیط‌های صنعتی بررسی کرده و بهبود دقت پیش‌بینی و کاهش مصرف انرژی را گزارش کردند [۹].

با وجود پیشرفت‌های اخیر، بیشتر مطالعات تنها بر یک ساختمان یا یک مجموعه داده متمرکز بوده‌اند و مقایسه جامع الگوریتم‌ها در اقلیم‌های متفاوت کمتر انجام شده است. علاوه بر این، در اغلب پژوهش‌ها نتایج مدل‌های یادگیری ماشین صرفاً بر اساس شاخص‌های آماری ارزیابی شده و ارتباط آن‌ها با تحلیل‌های ترمودینامیکی، به ویژه تحلیل انرژی و اکسرژی، به طور منسجم بررسی نشده است. از این رو، در پژوهش حاضر عملکرد هشت الگوریتم یادگیری ماشین بر روی سه مجموعه داده واقعی و مستقل از ساختمان‌های مسکونی و اداری در دو اقلیم متفاوت ارزیابی شده و نتایج حاصل با تحلیل انرژی و اکسرژی تفسیر شده است.

۲- روش پژوهش

در این پژوهش، به منظور ارزیابی عملکرد الگوریتم‌های یادگیری ماشین در پیش‌بینی بارهای گرمایشی و سرمایشی ساختمان، یک چارچوب پژوهشی شامل انتخاب مجموعه داده‌ها، پیش‌پردازش داده‌ها، آموزش و ارزیابی مدل‌ها و تحلیل انرژی و اکسرژی تدوین شد. در این راستا، سه مجموعه داده مستقل از ساختمان‌های مسکونی و اداری با ویژگی‌ها و شرایط اقلیمی متفاوت مورد استفاده قرار گرفت تا ضمن اعتبارسنجی روش پیشنهادی، عملکرد و پایداری مدل‌ها در شرایط مختلف نیز بررسی شود. ابتدا، مجموعه داده‌های مورد استفاده در مطالعه معرفی شده و سپس توضیحاتی در مورد معیارهای خطا و الگوریتم‌های انتخاب ویژگی و مدل‌های یادگیری ماشین ارائه شده است.

۱. مجموعه داده اول

مجموعه داده اول از مخزن UCI Machine Learning Repository با عنوان Energy Efficiency استخراج شده و شامل ۷۶۸ نمونه داده و ۸ ویژگی ورودی که مربوط به ۱۲ نوع مختلف ساختمان مسکونی است. متغیرهای هدف نیز شامل بار گرمایشی و بار سرمایشی (بر حسب کیلووات ساعت بر متر مربع) می‌باشند [۱۰]. این مجموعه داده به منظور اعتبارسنجی (Validation) روش کار با مقاله مرجع مورد استفاده قرار گرفته است [۱۱] مشخصات ویژگی‌های این مجموعه داده در جدول ۱ ارائه شده است.

جدول ۱. مشخصات متغیرهای ورودی و خروجی مجموعه داده اول (ساختمان مسکونی)

Table 1. Input and Output variables of the first dataset (Residential building)

Parameters	Feature	Description	Unit
X_1	Relative Compactness	Relative compactness, which is the ratio of the building volume to the external surface area	-
X_2	Surface Area	The total external surface area of the building	(m ²)
X_3	Wall Area	The total area of the external walls of the building	(m ²)
X_4	Roof Area	The area of the roof of the building	(m ²)
X_5	Overall Height	The total height of the building from ground level to the highest point of the roof	(m ²)
X_6	Orientation	The orientation of the building relative to the geographical directions	-
X_7	Glazing Area	The total area of the windows of the building	(m ²)
X_8	Glazing Area Distribution	The distribution of the area of the windows on the different walls of the building	-
Y_1	Heating Load	Heating Load	($\frac{kWh}{m^2}$)
Y_2	Cooling Load	Cooling Load	($\frac{kWh}{m^2}$)

این مجموعه داده شامل هشت ویژگی هندسی و حرارتی ساختمان است که مشخصات فیزیکی مؤثر بر رفتار انرژی ساختمان را توصیف می‌کند. این ویژگی‌ها به عنوان متغیرهای ورودی مدل و دو متغیر هدف شامل بار گرمایشی (Heating Load) و بار سرمایشی (Cooling Load) به کار گرفته شد. به دلیل استفاده گسترده از این مجموعه داده در مطالعات پیشین، از آن در پژوهش حاضر به عنوان مجموعه داده مرجع برای اعتبارسنجی و پیاده‌سازی مدل‌ها استفاده شد.

۲. مجموعه داده دوم

این مجموعه داده، مربوط به یک ساختمان اداری دو طبقه در آمریکا است که در سال ۲۰۲۱ جمع‌آوری شده است و شامل ۶۴,۷۹۲ نمونه با فواصل زمانی یک دقیقه‌ای می‌باشد [۱۲]. این مجموعه داده شامل متغیرهای محیطی و شرایط عملکرد سیستم تهویه مطبوع بوده و برای پیش‌بینی هم‌زمان بارهای گرمایشی و سرمایشی مورد استفاده قرار گرفت. مشخصات متغیرهای ورودی و خروجی این مجموعه داده در جدول ۲ ارائه شده است.

جدول ۲. مشخصات متغیرهای ورودی و خروجی مجموعه داده دوم (ساختمان‌های اداری)

Table 2. Input and output variables of the second dataset (Office buildings)

Parameters	Unit
Relative Humidity: رطوبت نسبی داخل آپارتمان	(%)
External Relative Humidity: رطوبت نسبی بیرون آپارتمان	(%)
Temperature: دمای داخل آپارتمان	(°C)
Out-Temperature: دمای بیرون آپارتمان	(°C)
Dir-Solar: تابش مستقیم خورشیدی	(W/m ²)
Heating Load: بار گرمایش	(W)
Cooling Load: بار سرمایش	(W)

۳. مجموعه داده سوم

مجموعه داده سوم از یک ساختمان اداری در کشور دانمارک که در سال ۲۰۲۲ گردآوری شده است و دارای ۵۵,۰۲۹ نمونه می‌باشد [۱۳]. این مجموعه داده تنها شامل بار سرمایشی به عنوان متغیر هدف بوده و به منظور ارزیابی عملکرد مدل‌ها در شرایط اقلیمی متفاوت مورد استفاده قرار گرفت. مشخصات متغیرهای ورودی و خروجی در جدول ۳ ارائه شده است.

جدول ۳. مشخصات متغیرهای ورودی و خروجی مجموعه داده سوم (ساختمان‌های اداری)

Table 3. Input and output variables of the third dataset (Office buildings)

Parameters	Unit
Relative Humidity: رطوبت نسبی داخل آپارتمان	(%)
External Relative Humidity: رطوبت نسبی بیرون آپارتمان	(%)
Temperature: دمای داخل آپارتمان	(°C)
Out-Temperature: دمای بیرون آپارتمان	(°C)
Cooling Load: بار سرمایش	(kW)

۲-۱- پیش پردازش داده‌ها

پس از بارگذاری مجموعه داده‌ها در محیط Python، عملیات پیش‌پردازش با استفاده از کتابخانه‌های Pandas، NumPy و Scikit-learn انجام شد. در مجموعه داده‌های دوم و سوم، ابتدا داده‌های نامعتبر و مقادیر گمشده بررسی و حذف شدند. سپس به منظور یکسان‌سازی مقیاس ویژگی‌های ورودی و جلوگیری از غلبه متغیرهای دارای دامنه بزرگ‌تر، تمامی ویژگی‌های ورودی با استفاده از روش StandardScaler استانداردسازی شدند. در مجموعه داده اول، به منظور حفظ قابلیت مقایسه با مطالعه مرجع، هیچ‌گونه نرمال‌سازی یا استانداردسازی بر روی داده‌ها اعمال نشد.

۲-۱-۱- تقسیم داده‌ها

برای مجموعه داده اول، مطابق با مطالعه مرجع، داده‌ها به صورت تصادفی با نسبت ۸۰٪ برای آموزش و ۲۰٪ برای آزمون تقسیم شدند تا امکان اعتبارسنجی و مقایسه مستقیم نتایج با مقاله مرجع فراهم شود [۱۱]. در مقابل، مجموعه داده‌های دوم و سوم شامل اندازه‌گیری‌های پیوسته زمانی از عملکرد سیستم HVAC هستند. از این رو، به منظور حفظ وابستگی زمانی داده‌ها و جلوگیری از نشت اطلاعات (Data Leakage)، تقسیم داده‌ها بدون درهم‌ریزی و براساس ترتیب زمانی انجام شد. به طوری که ۷۵٪ ابتدایی داده‌ها برای آموزش و ۲۵٪ انتهایی برای آزمون در نظر گرفته شد. این رویکرد، شرایط واقعی پیش‌بینی بارهای حرارتی را شبیه‌سازی کرده و ارزیابی واقع‌بینانه‌تر و قابل‌اعتمادتری از عملکرد مدل‌ها ارائه می‌دهد.

۲-۱-۲- الگوریتم‌های یادگیری ماشین و تنظیم هایپرپارامترها

به منظور پیش‌بینی بارهای گرمایشی و سرمایشی، از هشت الگوریتم یادگیری ماشین شامل Linear Regression، Decision Tree، Random Forest، Gradient Boosting، XGBoost، KNN، SVR و ANN استفاده شد [۱۴]. این الگوریتم‌ها با هدف مقایسه روش‌های خطی، مبتنی بر درخت، فاصله، کرنل و شبکه‌های عصبی انتخاب شدند تا عملکرد مدل‌ها در ساختارهای مختلف مورد ارزیابی قرار گیرد. به منظور افزایش شفافیت روش تحقیق و قابلیت بازتولید نتایج، مهم‌ترین هایپرپارامترهای مورد استفاده برای هر الگوریتم در جدول ۴ ارائه شده است.

جدول ۴. الگوریتم‌های یادگیری ماشین و هایپرپارامترهای مورد استفاده در مجموعه داده‌های دوم و سوم

Table 4. Machine learning algorithms and hyperparameters used for the second and third datasets

Algorithms	Algorithm Type	Hyperparameters dataset (2)	Hyperparameters dataset (3)
Linear Regression	Linear	fit_intercept = True	fit_intercept = True
Random Forest	Ensemble	n_estimators = 100	max_depth = 12 n_estimators = 300
Gradient Boosting	Boosting	n_estimators = 100	n_estimators = 300 max_depth = 7 learning_rate = 0.05
Decision Tree	Tree-based	criterion = squared_error	max_depth = criterion = squared_error 12

SVR	Kernel-based	C=1.0 ,Kernel=RBF epsilon=0.1 ,gamma=scale	kernel = RBF epsilon = 0.1 , C = 1000 ,
XGBoost	Gradient Boosting Ensemble	n_estimators = 100	n_estimators = 300 learning_rate = 0.05 max_depth = 7
KNN	Distance-based	k=5	n_neighbors = 5 weights = distance
ANN	ANN Neural Network	Hidden Layers = (100,50) Activation = ReLU Output=Linear Optimizer=Adam Batch Size=64 ,Epoch=200 Dropout=0.2 Early Stopping (Patience=10)	hidden_layer_sizes = (200,100,50) max_iter = learning_rate_init = 0.001 3000

هایپرپارامترهای گزارش شده، مقادیر نهایی مورد استفاده در مدل سازی هستند. از آن جا که سه مجموعه داده مورد استفاده از نظر حجم، نوع ساختمان، ویژگی های ورودی، کیفیت داده، میزان نویز و رفتار زمانی بارهای حرارتی با یکدیگر تفاوت دارند، تنظیم هایپرپارامترها برای هر مجموعه داده به صورت مستقل انجام شد. برای هر الگوریتم، چندین ترکیب از پارامترهای اصلی به صورت آزمایشی مورد بررسی قرار گرفت و در نهایت مجموعه ای از پارامترها انتخاب شد که بهترین عملکرد را بر اساس معیارهای RMSE، MSE، MAE و R^2 بر روی داده های آزمون فراهم می کرد. از این رو، اختلاف مقادیر هایپرپارامترها بین مجموعه داده ها ناشی از فرآیند تنظیم مدل متناسب با ویژگی های هر مجموعه داده است و تفاوت در ساختار یا نحوه پیاده سازی الگوریتمها نیست.

۲-۱-۳- معیارهای ارزیابی عملکرد مدل ها

برای سنجش دقت و عملکرد الگوریتمها، از چهار معیار خطا به شرح زیر استفاده شد:

- ۱) میانگین خطای مربع (MSE): میانگین تفاوت بین مقادیر واقعی و پیش بینی شده را محاسبه می کند و هر چه به صفر نزدیک تر باشد، دقت مدل بالاتر است و با کمک معادله ۱ محاسبه می شود.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - Y_t)^2 \quad \text{معادله ۱}$$

n: تعداد پیش بینی ها

Y_i : مقادیر مشاهده شده

Y_t : مقادیر پیش بینی شده

- ۲) ریشه میانگین خطای مربع (RMSE): جذر MSE بوده و خطا را در مقیاس همان متغیر هدف بیان می کند. معادله ۲ نشان دهنده این خطا می باشد.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - Y_t)^2} \quad \text{معادله ۲}$$

- ۳) میانگین خطای مطلق (MAE): میانگین قدر مطلق تفاوت بین مقادیر واقعی و پیش بینی شده است و نسبت به MSE به مقادیر پرت کمتر حساس است. این خطا به کمک معادله ۳ محاسبه می شود.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - Y_t| \quad \text{معادله ۳}$$

۴) ضریب تعیین (R^2): معیاری برای نشان دادن میزان تطابق مدل با داده‌ها است. این مقدار بین ۰ و ۱ قرار دارد و هر چه به ۱ نزدیک‌تر باشد، مدل بهتر عمل کرده است.

$$R^2 = 1 - \frac{SSR}{TSS} \quad \text{معادله ۴}$$

در این معادله SSR و TSS به ترتیب مجموع مربعات باقیمانده‌ها و مجموع مربعات کل هستند. علاوه بر معیارهای فوق، برای درک عمیق‌تر رفتار ترمودینامیکی ساختمان مورد مطالعه، آنالیز انرژی و اکسرژی نیز بر روی یکی از ساختمان‌های اداری انجام شده است [۱۵].

۲-۲- آنالیز انرژی و اکسرژی

آنالیز انرژی مبتنی بر قانون اول ترمودینامیک است و به کمیت انرژی توجه دارد [۱۶]. در حالی که آنالیز اکسرژی بر پایه قوانین اول و دوم ترمودینامیک، هم کمیت و هم کیفیت انرژی را بررسی کرده و منابع برگشت‌ناپذیری (تخریب اکسرژی) را شناسایی می‌کند [۱۷]. در این پژوهش، تأثیر استفاده از پنجره‌های تک‌جداره و دوجداره بر مصرف انرژی و تخریب اکسرژی یک ساختمان اداری دو طبقه در فصول تابستان و زمستان مورد ارزیابی قرار گرفته است [۱۸]. در ادامه معادلات مورد استفاده ارائه می‌شود:

الف) آنالیز انرژی:

موازنه انرژی در فصل تابستان به صورت معادله ۵ محاسبه می‌شود.

$$Q_{wall} + Q_{window} + Q_{lamp} + Q_{people} + Q_{solar} = Q_{HVAC} \quad \text{معادله ۵}$$

که در این معادله:

Q_{wall} : انتقال حرارت دیوار (W)

Q_{window} : انتقال حرارت پنجره (W)

Q_{lamp} : انرژی ناشی از لامپ (W)

Q_{people} : انرژی ناشی از افراد حاضر در ساختمان (W)

Q_{solar} : مقدار انرژی جذب شده از تابش خورشید (W)

Q_{HVAC} : مقدار بار حرارتی سیستم تهویه مطبوع (W)

نرخ انتقال حرارت هدایتی از دیوار و پنجره در فصل تابستان از معادلات ۶ و ۷ محاسبه می‌شود:

$$Q_{wall} = U_{wall} A_{wall} (T_o - T_i) \quad \text{معادله ۶}$$

$$Q_{window} = U_{window} A_{window} (T_o - T_i) \quad \text{معادله ۷}$$

که در این معادله:

U_{wall} : ضریب انتقال حرارت کلی دیوار ($\frac{W}{m^2K}$)

U_{window} : ضریب انتقال حرارت کلی پنجره ($\frac{W}{m^2K}$)

A_{wall} : مساحت دیوار (m^2)

A_{window} : مساحت پنجره (m^2)

T_o : دمای بیرون (K)

T_i : دمای داخل (K)

میزان انرژی جذب شده از خورشید نیز از معادله ۸ محاسبه می‌شود:

$$Q_{solar} = I * A * f_a \quad \text{معادله ۸}$$

I : شدت تابش خورشید ($\frac{W}{m^2}$)

f_a : ضریب جذب تابش خورشید

A : مساحت سطح (m^2)

انرژی مربوط به افراد حاضر در ساختمان و لامپ‌های مورد استفاده در تابستان و زمستان نیز به ترتیب از حاصل ضرب تعداد در میزان انرژی متوسط هر فرد و هر لامپ در واحد وات محاسبه می‌شود.

در فصل زمستان، به دلیل پایین‌تر بودن دمای بیرون نسبت به داخل، جهت انتقال حرارت معکوس شده و انرژی از داخل ساختمان به محیط خارج هدر می‌رود. بنابراین، موازنه انرژی در زمستان به صورت معادله ۹ نوشته می‌شود:

$$Q_{HVAC} + Q_{lamp} + Q_{people} + Q_{solar} = Q_{wall} + Q_{window} \quad \text{معادله ۹}$$

نحوه محاسبه هر یک از روابط این معادله مشابه فصل تابستان است، با این تفاوت که در روابط مربوط به انتقال حرارت هدایتی دیوار و پنجره، به دلیل پایین‌تر بودن دمای بیرون، اختلاف دما به صورت $(T_i - T_o)$ می‌باشد [۱۹].

ب) آنالیز اکسرژی:

تخریب اکسرژی در فصل تابستان (Ex_{d_summer}) به کمک معادله ۱۰ محاسبه می‌شود.

$$Ex_{d_summer} = Q_{solar} \left| 1 - \frac{T'}{T_s} \right| + Q_{wall+window} \left| 1 - \frac{T'}{T_o} \right| + Q_{people} \left| 1 - \frac{T'}{T_{people}} \right| + Q_{lamp} \left| 1 - \frac{T'}{T_{lamp}} \right| - Q_{HVAC} \left| 1 - \frac{T'}{T_i - 10} \right| \quad \text{معادله ۱۰}$$

T' : دمای مرجع

T_s : دمای خورشید

سایر مؤلفه‌های اکسرژی متناظر با ترم‌های انرژی معرفی شده در بخش قبل بوده است. در سمت راست این معادله، جمله اول اکسرژی خورشید، جمله دوم اکسرژی انتقال حرارت هدایتی از دیوار و پنجره، جملات سوم و چهارم به ترتیب اکسرژی ناشی از افراد و لامپ‌ها و جمله آخر اکسرژی سیستم تهویه مطبوع می‌باشد.

همچنین تخریب اکسرژی در فصل زمستان (Ex_{d_winter}) از معادله ۱۱ محاسبه می‌شود [۱۶]:

$$Ex_{d_winter} = Q_{solar} \left| 1 - \frac{T'}{T_s} \right| + Q_{HVAC} \left| 1 - \frac{T'}{T_i + 10} \right| + Q_{people} \left| 1 - \frac{T'}{T_{people}} \right| + Q_{lamp} \left| 1 - \frac{T'}{T_{lamp}} \right| - Q_{wall+window} \left| 1 - \frac{T'}{T_o} \right| \quad \text{معادله ۱۱}$$

در این معادلات دمای مرجع $T' = 298 \text{ K}$ و دمای خورشید $T_s = 5670 \text{ K}$ در نظر گرفته شده است.

۳. نتایج و بحث

در این پژوهش، عملکرد الگوریتم‌های یادگیری ماشین برای پیش‌بینی بارهای گرمایشی و سرمایشی ساختمان با استفاده از سه مجموعه داده مستقل مورد ارزیابی قرار گرفت. در گام نخست، به منظور اطمینان از صحت پیاده‌سازی مدل‌ها و بررسی قابلیت بازتولید نتایج، فرآیند اعتبارسنجی با استفاده از مجموعه داده مرجع انجام شد [۱۱]. در این مرحله، تمامی شرایط مطالعه مرجع به طور کامل رعایت گردید. به طوری که هیچ‌گونه نرمال‌سازی یا استانداردسازی بر روی داده‌ها اعمال نشد و داده‌ها به صورت تصادفی با نسبت ۸۰٪ برای آموزش و ۲۰٪ برای آزمون ($random_state=42$) تقسیم شدند. عملکرد مدل‌ها نیز تنها بر اساس داده‌های آزمون و با استفاده از معیارهای RMSE، MAE و R^2 ارزیابی شد. نتایج اعتبارسنجی مربوط به پیش‌بینی بار گرمایشی در جدول ۵ ارائه شده است.

جدول ۵. نتایج اعتبارسنجی بار گرمایشی مجموعه داده اول با مقاله مرجع

Table 5. Results of validation of the heating load of the first dataset with the reference article

Algorithm	MAE, this work ($\frac{kWh}{m^2}$)	RMSE, this work ($\frac{kWh}{m^2}$)	R^2 this work	MAE ($\frac{kWh}{m^2}$) [11]	RMSE ($\frac{kWh}{m^2}$) [11]	R^2 [11]
Decision Tree	0.42	0.62	0.99	0.42	0.62	0.99
Random Forest	0.35	0.49	0.99	0.35	0.49	0.99
XGBoost	0.26	0.40	0.99	0.26	0.40	0.99
Linear Regression	2.18	2.92	0.91	2.20	3.04	0.91
SVR	4.23	5.65	0.69	4.23	5.65	0.69

بنابراین، نتایج حاصل از پیاده‌سازی الگوریتم‌ها در این پژوهش با مقادیر گزارش شده در مطالعه مرجع تطابق بسیار بالایی دارد. این همخوانی نشان می‌دهد که فرآیند آماده‌سازی داده‌ها، آموزش مدل‌ها و محاسبه معیارهای ارزیابی به درستی انجام شده است.

به منظور بررسی عملکرد مدل‌ها در پیش‌بینی بار سرمایشی نیز، نتایج اعتبارسنجی در جدول ۶ ارائه شده است.

جدول ۶. نتایج اعتبارسنجی بار سرمایشی مجموعه داده اول با مقاله مرجع

Table 6. Results of validation of the cooling load of the first dataset with the reference article

Algorithm	MAE, this work ($\frac{kWh}{m^2}$)	RMSE, this work ($\frac{kWh}{m^2}$)	R^2 this work	MAE ($\frac{kWh}{m^2}$) [11]	RMSE ($\frac{kWh}{m^2}$) [11]	R^2 [11]
Decision Tree	1.15	2.01	0.95	1.15	2.01	0.95
Random Forest	1.06	1.71	0.96	1.06	1.71	0.96
XGBoost	0.44	0.85	0.99	0.44	0.85	0.99
Linear Regression	2.19	3.14	0.89	2.19	3.14	0.89
SVR	3.96	5.35	0.69	3.96	5.35	0.69

با توجه به نتایج که روندی مشابه با بار گرمایشی را نشان می‌دهد و بیانگر تطابق بسیار مناسب نتایج این پژوهش با مقادیر گزارش شده در مطالعه مرجع است. در مجموع، تطابق بسیار بالای نتایج با مطالعه مرجع، اعتبار چارچوب پیاده‌سازی و صحت مراحل مدل‌سازی در پژوهش حاضر را تأیید می‌کند و نشان می‌دهد که روش پیشنهادی از قابلیت اطمینان کافی برای تحلیل مجموعه داده‌های مستقل بعدی برخوردار است.

در ادامه، به منظور ارزیابی عملکرد مدل‌های یادگیری ماشین در شرایط واقعی، الگوریتم‌ها بر روی مجموعه داده دوم مربوط به یک ساختمان اداری در ایالات متحده آمریکا اجرا شدند. این مجموعه داده شامل هر دو بار گرمایشی و سرمایشی است و امکان مقایسه عملکرد مدل‌ها را در دو حالت مختلف بهره‌برداری ساختمان فراهم می‌کند. نتایج حاصل از پیش‌بینی بار گرمایشی توسط هشت الگوریتم مورد مطالعه در جدول ۷ ارائه شده است.

جدول ۷. نتایج معیارهای خطا برای پیش‌بینی بار گرمایشی مجموعه داده دوم

Table 7. Error metrics for heating load prediction of the second dataset

Algorithms	MSE (W^2)	MAE (W)	RMSE (W)	R^2
Linear Regression	256.28	13.22	16.01	0.26
Random Forest	5.85	1.36	2.42	0.98
Gradient Boosting	71.90	5.94	8.45	0.79
Decision Tree	7.14	0.79	2.67	0.97
SVR	150.22	6.91	12.29	0.56
XGBoost	17.40	2.74	4.17	0.95
KNN	15.44	2.40	3.93	0.95
ANN	78.88	5.14	8.88	0.77

براساس نتایج، الگوریتم Random Forest بهترین عملکرد را در پیش‌بینی بار گرمایشی از خود نشان داد و با دستیابی به بیشترین مقدار ضریب تعیین (R^2) و کمترین مقادیر RMSE و MAE، نسبت به سایر مدل‌ها دقت بالاتری داشت. این نتایج نشان می‌دهد که مدل از دقت و پایداری مناسبی در پیش‌بینی این مجموعه داده برخوردار است.

به منظور بررسی عملکرد مدل‌ها در پیش‌بینی بار سرمایشی همین ساختمان، نتایج حاصل از اجرای الگوریتم‌ها برای این بخش نیز در جدول ۸ ارائه شده است.

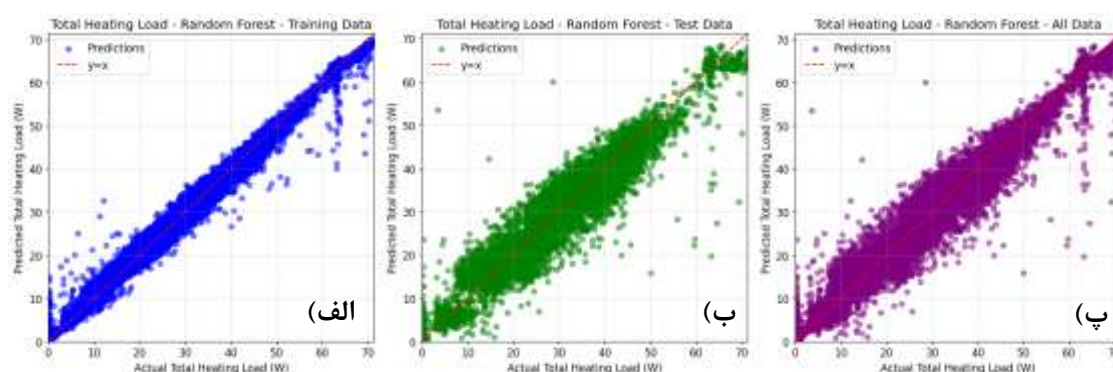
جدول ۸. نتایج معیارهای خطا برای پیش‌بینی بار سرمایشی مجموعه داده دوم

Table 8. Error metrics for cooling load prediction of the second dataset

Algorithms	MSE (W^2)	MAE (W)	RMSE (W)	R ²
Linear Regression	56.27	5.42	7.50	0.62
Random Forest	4.27	0.99	2.06	0.97
Gradient Boosting	23.32	3.17	4.83	0.84
Decision Tree	5.18	0.55	2.28	0.96
SVR	32.25	3.45	5.68	0.78
XGBoost	11.12	1.92	3.34	0.92
KNN	10.70	1.74	3.27	0.92
ANN	21.50	95.2	4.64	0.85

با توجه به نتایج، الگوریتم Random Forest در پیش‌بینی بار سرمایشی نیز عملکرد برتری نسبت به سایر الگوریتم‌ها داشته است. این نتایج نشان می‌دهد که این مدل علاوه بر پیش‌بینی بار گرمایشی، در مدل‌سازی رفتار بار سرمایشی نیز از دقت و پایداری مناسبی برخوردار است. عملکرد مطلوب Random Forest در هر دو بار بیانگر توانایی این الگوریتم در استخراج روابط پیچیده میان شرایط محیطی و بار انرژی ساختمان و همچنین مقاومت مناسب آن در برابر نویز موجود در داده‌ها می‌باشد.

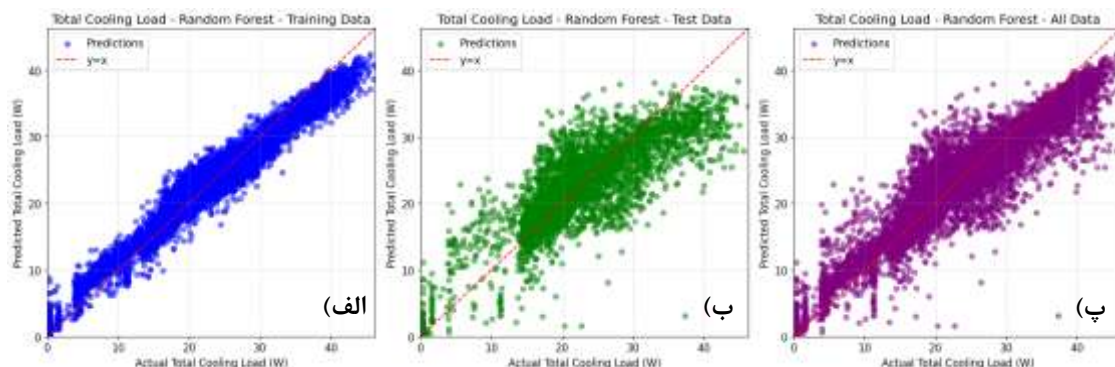
بر همین اساس، نمودار پراکندگی این الگوریتم در شکل‌های ۱ و ۲ نشان داده شده است. این نمودارها مقادیر پیش‌بینی شده براساس مقادیر واقعی را برای بارهای گرمایش و سرمایش نشان می‌دهد.



شکل ۱. نمودار پراکندگی بار گرمایش برای (الف) داده آموزشی، (ب) داده آزمایشی، (پ) کل داده‌ها براساس الگوریتم Random Forest

Figure 1. Scatter plot of heating load for (a) training, (b) test, and (c) all data using Random Forest algorithm

با توجه به نمودار، بخش عمده‌ای از نقاط مربوط به داده‌های آموزش و آزمون در مجاورت خط مرجع $y=x$ قرار گرفته‌اند که بیانگر انطباق مطلوب مقادیر پیش‌بینی شده با مقادیر واقعی است. همچنین، توزیع یکنواخت نقاط در اطراف خط مرجع و عدم مشاهده الگوی مشخص در میزان خطاها نشان می‌دهد که مدل Random Forest توانسته است بار گرمایشی را در سطوح مختلف بهره‌برداری ساختمان با دقت مناسبی پیش‌بینی کند. این نتایج بیانگر پایداری عملکرد مدل و نبود سوگیری قابل توجه در فرآیند پیش‌بینی است. در ادامه، به منظور بررسی عملکرد این الگوریتم در پیش‌بینی بار سرمایشی، نمودار پراکندگی آن در شکل ۲ ارائه شده است.



شکل ۲. نمودار پراکندگی بار سرمایش برای (الف) داده آموزشی، (ب) داده آزمایشی، (پ) کل داده‌ها براساس الگوریتم

Random Forest

Figure 2. Scatter plot of cooling load for (a) training, (b) test, and (c) all data using Random Forest algorithm

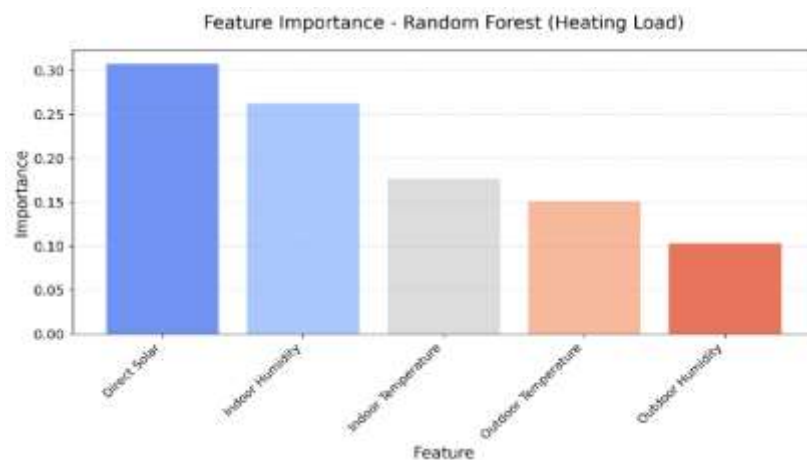
با توجه به نتایج، تمرکز مناسب نقاط در اطراف خط مرجع و نبود پراکندگی شدید، بیانگر توانایی مدل در برآورد بار سرمایشی در سطوح مختلف بهره‌برداری ساختمان است. این موضوع نشان می‌دهد که مدل تنها روند کلی داده‌ها را یاد نگرفته، بلکه در پیش‌بینی مقادیر واقعی در بازه‌های مختلف بار نیز عملکرد قابل قبولی ارائه کرده است.

به منظور بررسی میزان تأثیر هر یک از متغیرهای ورودی بر پیش‌بینی بارهای گرمایشی و سرمایشی، تحلیل اهمیت ویژگی‌ها انجام شد. با توجه به این که در این پژوهش از چندین الگوریتم یادگیری ماشین با ساختارهای متفاوت استفاده شده است، برای ایجاد مبنایی یکسان جهت مقایسه نتایج، روش‌های زیر به کار گرفته شد. برای تمامی مدل‌های یادگیری ماشین به جز رگرسیون خطی، اهمیت ویژگی‌ها با استفاده از روش Permutation Importance محاسبه شد. در این روش، مقادیر هر ویژگی به صورت تصادفی جابه‌جا شده و سپس کاهش عملکرد مدل در اثر این جابه‌جایی اندازه‌گیری می‌شود [۲۰]. هر چه جابه‌جایی یک ویژگی موجب افزایش بیشتر خطای مدل شود، آن ویژگی از اهمیت بالاتری در فرآیند پیش‌بینی برخوردار است. در این پژوهش، برای مدل‌های درختی سه تکرار و برای مدل‌های ANN و SVR دو تکرار در محاسبه Permutation Importance در نظر گرفته شد تا ضمن حفظ دقت نتایج، هزینه محاسباتی نیز کنترل شود.

در مدل رگرسیون خطی، اهمیت ویژگی‌ها بر اساس قدر مطلق ضرایب رگرسیون پس از استانداردسازی متغیرهای ورودی محاسبه شد؛ زیرا در این مدل، ضرایب رگرسیون بیانگر میزان تأثیر نسبی هر متغیر بر متغیر هدف هستند.

روش Permutation Importance از مزیت استقلال از نوع مدل برخوردار بوده و امکان مقایسه اهمیت ویژگی‌ها در میان الگوریتم‌های مختلف را فراهم می‌کند. با این حال، در صورت وجود همبستگی زیاد میان متغیرهای ورودی، ممکن است اهمیت برخی ویژگی‌ها کمتر از مقدار واقعی برآورد شود. از این رو، استفاده از روش‌های تبیین‌پذیری پیشرفته مانند SHAP می‌تواند در پژوهش‌های آینده برای تحلیل دقیق‌تر سهم هر ویژگی و تعیین جهت اثرگذاری آن مورد استفاده قرار گیرد.

به منظور بررسی سهم نسبی هر یک از متغیرهای ورودی در پیش‌بینی بار گرمایشی، تحلیل اهمیت ویژگی‌ها برای مدل منتخب انجام شد. شکل ۳ میزان اهمیت نسبی هر یک از پارامترهای ورودی را در فرآیند پیش‌بینی بار گرمایشی نشان می‌دهد.

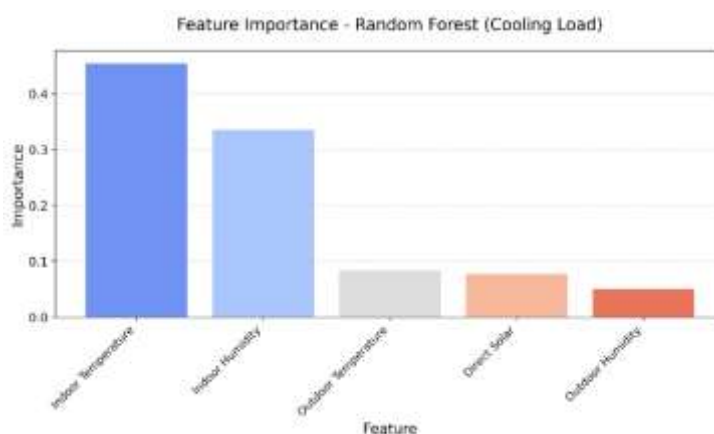


شکل ۳. نمودار اهمیت ویژگی‌های مختلف بر بار گرمایش براساس الگوریتم Random Forest

Figure 3. Feature importance of different variables for heating load using Random Forest algorithm

براساس نمودار، تابش مستقیم خورشید (Direct Solar) بیشترین سهم را در پیش‌بینی بار گرمایشی داشته است. این نتیجه از نظر فیزیکی قابل انتظار است، زیرا افزایش تابش خورشیدی موجب افزایش بهره‌رسانی ساختمانی شده و نیاز به سیستم گرمایشی را کاهش یا تغییر می‌دهد. بنابراین مدل یادگیری ماشین نیز این متغیر را به عنوان مهم‌ترین عامل مؤثر شناسایی کرده است.

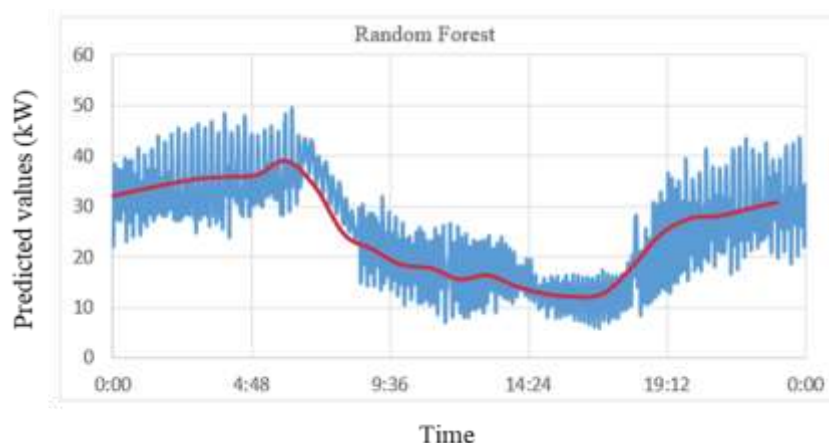
همچنین، رطوبت داخلی (Indoor Humidity) نیز یکی از مهم‌ترین متغیرها بوده است. تغییرات رطوبت بر خواص ترمودینامیکی هوا، از جمله آنتالپی و بار محسوس و نهان، اثر می‌گذارد و در نتیجه بر عملکرد سیستم HVAC و میزان انرژی مورد نیاز برای حفظ شرایط آسایش تأثیرگذار است. اهمیت بالای این دو متغیر نشان می‌دهد که مدل توسعه‌یافته توانسته است عوامل فیزیکی مؤثر بر بار گرمایشی را به درستی شناسایی کند. به منظور شناسایی مهم‌ترین عوامل مؤثر بر پیش‌بینی بار سرمایشی، تحلیل اهمیت ویژگی‌ها برای مدل منتخب انجام شد. شکل ۴ اهمیت نسبی متغیرهای ورودی در پیش‌بینی بار سرمایشی را نشان می‌دهد.



شکل ۴. نمودار اهمیت ویژگی‌های مختلف بر بار سرمایش براساس الگوریتم Random Forest

Figure 4. Feature importance of different variables for cooling load using Random Forest algorithm

مطابق شکل، دمای داخلی بیشترین اهمیت را در پیش‌بینی بار سرمایشی داشته است. این نتیجه از دیدگاه ترمودینامیکی قابل انتظار است، زیرا هرچه دمای هوای داخل از دمای تنظیم شده فاصله بیشتری بگیرد، سیستم تهویه مطبوع برای حذف گرمای اضافی و بازگرداندن شرایط آسایش حرارتی به ظرفیت سرمایشی بیشتری نیاز خواهد داشت. بنابراین، دمای داخلی یکی از مهم‌ترین متغیرهای تعیین‌کننده بار سرمایشی ساختمان محسوب می‌شود. همچنین، رطوبت داخلی نیز دومین پارامتر اثرگذار بوده است. افزایش رطوبت موجب افزایش بار نهان سیستم تهویه مطبوع شده و انرژی بیشتری برای رطوبت‌زدایی مورد نیاز خواهد بود. از این رو، مدل‌های یادگیری ماشین نیز این متغیر را به عنوان یکی از مهم‌ترین عوامل تعیین‌کننده بار سرمایشی شناسایی کرده‌اند. پس از تحلیل اهمیت ویژگی‌های مؤثر، عملکرد مدل در پیش‌بینی سری زمانی بار گرمایشی نیز مورد ارزیابی قرار گرفت. به این منظور، شکل ۵ مقایسه مقادیر واقعی و پیش‌بینی شده بار گرمایشی توسط الگوریتم Random Forest را در یک دوره ۲۴ ساعته نمایش می‌دهد تا میزان تطابق مدل با رفتار واقعی سیستم بررسی شود.



شکل ۵. نمودار مقایسه بار گرمایش واقعی و پیش‌بینی شده توسط الگوریتم Random Forest در طول یک شبانه روز

Figure 5. Comparison of actual and predicted heating load using the Random Forest algorithm over a 24-hour period

همان‌گونه که مشاهده شد، مقادیر واقعی بار گرمایشی (خط آبی) دارای نوسانات لحظه‌ای و پیک‌های متعددی هستند که ناشی از تغییرات شرایط محیطی و عملکرد لحظه‌ای تجهیزات HVAC است. در مقابل، مقادیر پیش‌بینی شده توسط الگوریتم Random Forest (خط قرمز) ضمن هموارسازی نوسانات تصادفی، روند کلی تغییرات بار گرمایشی را با دقت مناسبی دنبال کرده است. بیشترین بار گرمایشی در ساعات شب و ابتدای بامداد مشاهده می‌شود که با کاهش دمای هوای محیط و افزایش نیاز گرمایشی ساختمان مطابقت دارد. همچنین کاهش تدریجی بار در ساعات روز و افزایش مجدد آن در ساعات عصر و شب نیز به خوبی توسط مدل

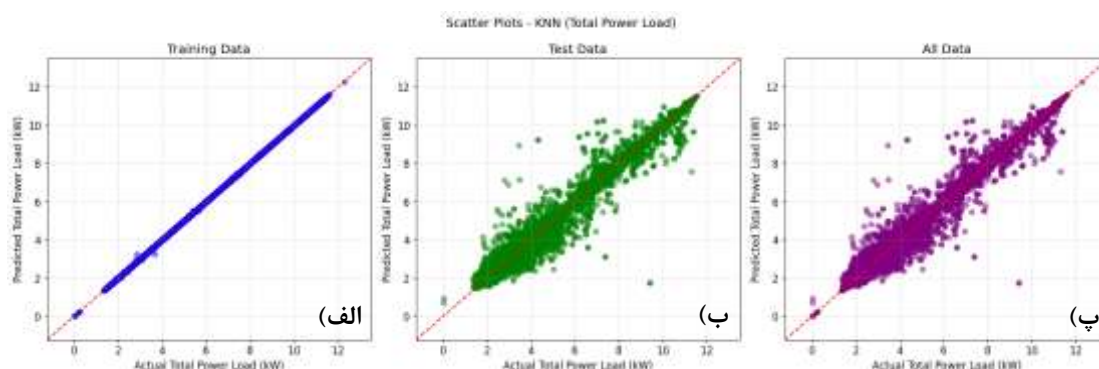
بازتولید شده است. این نتایج نشان می‌دهد که الگوریتم Random Forest علاوه بر دستیابی به دقت آماری مناسب، قادر است رفتار زمانی سیستم HVAC را نیز با قابلیت اطمینان مطلوب شبیه‌سازی کند. در ادامه، به منظور بررسی عملکرد الگوریتم‌های پیشنهادی در شرایط و مجموعه داده‌ای متفاوت، عملکرد هشت الگوریتم یادگیری ماشین با استفاده از مجموعه داده سوم، مربوط به یک ساختمان اداری در دانمارک، ارزیابی شد. شایان ذکر است که این مجموعه داده تنها شامل بار سرمایشی می‌باشد. جدول ۹ نتایج پیش‌بینی بار سرمایشی را برای هشت الگوریتم نشان می‌دهد.

جدول ۹. نتایج معیارهای خطا برای پیش‌بینی بار سرمایشی مجموعه داده سوم

Table 9. Error metrics for cooling load prediction of the third dataset

Algorithms	MSE (kW^2)	MAE (kW)	RMSE (kW)	R ²
Linear Regression	2.12	1.13	1.45	0.70
Random Forest	0.28	0.35	0.53	0.96
Gradient Boosting	0.29	0.37	0.54	0.95
Decision Tree	0.38	0.37	0.61	0.94
SVR	0.84	0.61	0.91	0.88
XGBoost	0.30	0.38	0.55	0.95
KNN	0.03	0.03	0.18	0.99
ANN	0.32	0.40	0.57	0.95

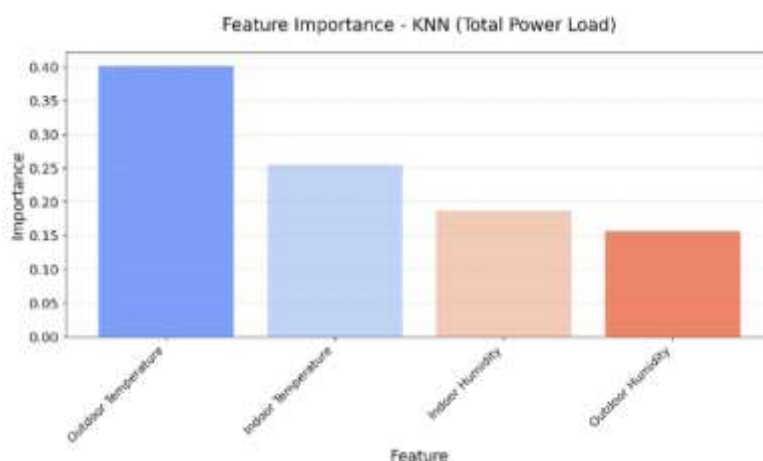
بر اساس نتایج، در مجموعه داده سوم، الگوریتم KNN با ضریب تعیین ۰/۹۹ و کمترین مقادیر MSE، RMSE و MAE، بهترین عملکرد را در پیش‌بینی بار سرمایشی ارائه داد. این نتیجه با نتایج مجموعه داده قبلی که در آن الگوریتم Random Forest بهترین عملکرد را داشت، متفاوت است و نشان می‌دهد که کارایی الگوریتم‌های یادگیری ماشین به ویژگی‌های مجموعه داده وابسته است. در این مجموعه داده، شباهت زیاد بین نمونه‌های مجاور و الگوهای تکرارشونده بار سرمایشی موجب شده است که KNN، پیش‌بینی و عملکرد بسیار مطلوبی داشته باشد. در مقابل، در مجموعه داده قبلی، وجود روابط غیرخطی پیچیده‌تر میان متغیرهای ورودی و بار حرارتی، سبب برتری مدل Random Forest شده بود. بنابراین، نتایج این پژوهش نشان می‌دهد که انتخاب مدل بهینه باید متناسب با ویژگی‌های داده و رفتار دینامیکی ساختمان انجام شود و نمی‌توان یک الگوریتم را به عنوان بهترین گزینه برای تمامی مجموعه داده‌ها معرفی کرد. نمودارهای پراکندگی مربوط به الگوریتم KNN در شکل ۶ ارائه شده‌اند.



شکل ۶. نمودار پراکندگی بار سرمایش برای (الف) داده آموزشی، (ب) داده آزمایشی، (پ) کل داده‌ها براساس الگوریتم KNN

Figure 6. Scatter plot of cooling load for (a) training, (b) test, and (c) all data using KNN algorithm

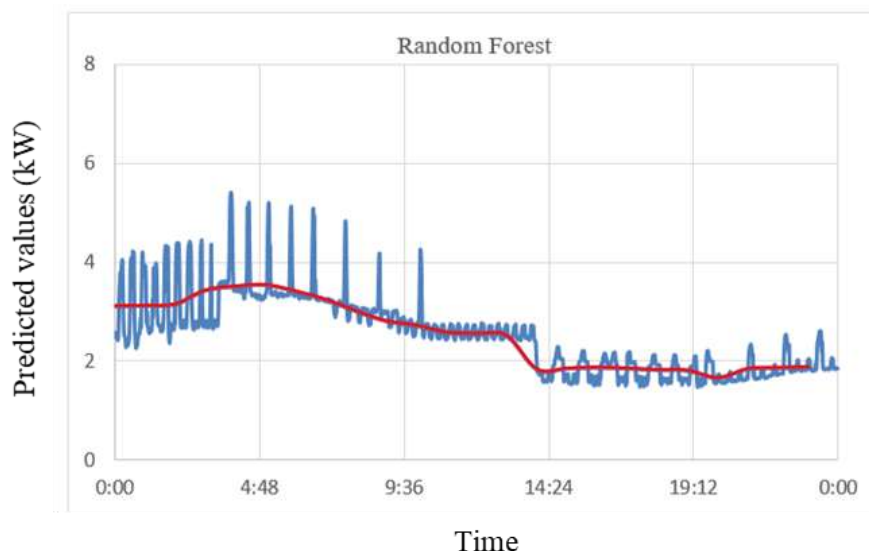
همانطور در نمودار مشاهده شد، اغلب نقاط در مجاورت خط ایده آل $y=x$ قرار گرفته‌اند که نشان‌دهنده انطباق مناسب مقادیر پیش‌بینی شده با مقادیر واقعی و دقت بالای الگوریتم KNN در برآورد بار سرمایشی است. عملکرد مطلوب این الگوریتم را می‌توان به ویژگی‌های مجموعه داده سوم نسبت داد. این مجموعه داده مربوط به یک ساختمان اداری است که الگوی بهره‌برداری نسبتاً منظم و شرایط عملکردی مشابهی در ساعات کاری دارد. در نتیجه، تغییرات بار سرمایشی عمدتاً تحت تأثیر متغیرهای محیطی نظیر دمای داخلی، رطوبت و تابش خورشیدی قرار گرفته و نمونه‌های مشابه متعددی در فضای ویژگی‌ها ایجاد شده است. از آن جا که الگوریتم KNN پیش‌بینی را بر اساس نزدیک‌ترین نمونه‌های آموزشی انجام می‌دهد، وجود این شباهت‌ها موجب افزایش دقت پیش‌بینی شده است. با این حال، عملکرد بسیار بالای KNN در این مجموعه داده لزوماً به معنای برتری مطلق آن در سایر ساختمان‌ها یا شرایط بهره‌برداری متفاوت نیست؛ زیرا در مجموعه داده‌هایی با تغییرات شدید بار، رفتار غیرخطی یا الگوهای کمتر تکرارشونده، عملکرد الگوریتم‌های مبتنی بر درخت تصمیم، مانند Random Forest، می‌تواند پایدارتر باشد. به منظور بررسی عوامل مؤثر بر پیش‌بینی بار سرمایشی، نمودار اهمیت ویژگی‌ها در شکل ۷ ارائه شده است.



شکل ۷. نمودار اهمیت ویژگی‌های مختلف بر بار سرمایش براساس الگوریتم KNN

Figure 7. Feature importance of different variables for cooling load using KNN algorithm

براساس نمودار اهمیت ویژگی‌ها، دمای بیرون و دمای داخل به ترتیب بیشترین تأثیر را بر بار سرمایشی این ساختمان دارند. این نتیجه با اصول فیزیکی انتقال حرارت در ساختمان‌ها سازگار است؛ زیرا افزایش دمای محیط و فضای داخلی موجب افزایش بار سرمایشی مورد نیاز سیستم تهویه مطبوع می‌شود. پس از شناسایی پارامترهای مؤثر، توانایی مدل در دنبال کردن تغییرات زمانی بار سرمایشی نیز مورد بررسی قرار گرفت. به این منظور، نمودار مقایسه مقادیر واقعی و پیش‌بینی شده بار سرمایشی توسط الگوریتم Random Forest طی یک شبانه روز در شکل ۸ ارائه شده است. اگرچه الگوریتم KNN در این مجموعه داده بالاترین دقت را کسب کرده است، اما با توجه به عملکرد پایدار و دقت مناسب الگوریتم Random Forest در هر سه مجموعه داده مورد بررسی، نمودار مربوط به این الگوریتم به عنوان نماینده عملکرد کلی مدل‌ها ارائه شده است.



شکل ۸. نمودار مقایسه بار سرمایش واقعی و پیش‌بینی شده توسط الگوریتم Random Forest در طول یک شبانه روز

Figure 8. Comparison of actual and predicted cooling load using the Random Forest algorithm over a 24-hour period

براساس نمودار ارائه شده، مقادیر واقعی (خط آبی) و مقادیر پیش‌بینی شده (خط قرمز) بار سرمایشی در طول یک شبانه روز تطابق مناسبی با یکدیگر دارند. با وجود نوسانات نسبتاً زیاد در داده‌های واقعی، الگوریتم Random Forest توانسته است روند زمانی تغییرات بار را با دقت مطلوب دنبال کند و بخش قابل توجهی از نویز موجود در داده‌ها را حذف نماید. همچنین، در بازه‌هایی که بار سرمایشی افزایش یا کاهش محسوسی دارد، مدل قادر است تغییرات اصلی را بازتولید کند و صرفاً میانگین بار را پیش‌بینی نکرده است. هرچند در برخی پیک‌های لحظه‌ای اختلاف اندکی بین مقادیر واقعی و پیش‌بینی شده مشاهده می‌شود، اما این رفتار به دلیل ماهیت تصادفی تغییرات بار و محدود بودن اطلاعات ورودی، طبیعی است. در مجموع، نتایج نشان می‌دهد که مدل علاوه بر پیش‌بینی روند کلی بار سرمایشی، توانایی مناسبی در دنبال کردن تغییرات زمانی و نوسانات اصلی سیستم نیز دارد که بیانگر قابلیت کاربرد آن در پیش‌بینی بار ساختمان است.

۳-۱- نتایج آنالیز انرژی و اکسرژی

به منظور تفسیر نتایج حاصل از پیش‌بینی بارهای حرارتی توسط مدل‌های یادگیری ماشین، تحلیل انرژی و اکسرژی بر روی ساختمان اداری مورد مطالعه انجام شد. از آن جا که بارهای گرمایشی و سرمایشی تابعی از ویژگی‌های حرارتی پوسته ساختمان، شرایط محیطی و دمای داخل هستند، این تحلیل امکان بررسی تأثیر نوع پنجره بر انتقال حرارت، بار سیستم HVAC و تخریب اکسرژی را فراهم می‌کند. پارامترها، نمادها، واحدها و منابع مورد استفاده در محاسبات در جدول ۱۰ ارائه شده‌اند.

جدول ۱۰. پارامترهای مورد استفاده در آنالیز انرژی و اکسرژی

Table 10. Parameters used in energy and exergy analyses

Parameters	Value	Unit	Reference
U_{wall}	1.5	$(\frac{W}{m^2K})$	[18]
$U_{\text{Window-single}}$	5.55	$(\frac{W}{m^2K})$	[21]
$U_{\text{Window-double}}$	1.95	$(\frac{W}{m^2K})$	[21]
A_{window}	111.30	(m^2)	[12]
A_{wall}	445.21	(m^2)	[10]
$f_a - \text{single}$	0.79	-	[22]
$f_a - \text{double}$	0.58	-	[22]
$T_i(\text{Summer})$	295	(K)	[12]
$T_o(\text{Summer})$	300	(K)	[12]
$T_i(\text{Winter})$	294	(K)	[12]
$T_o(\text{Winter})$	283	(K)	[12]
T_s	5670	(k)	[12]
$I(\text{Summer})$	240	$(\frac{W}{m^2K})$	[12]
$I(\text{Winter})$	149	$(\frac{W}{m^2K})$	[12]
(N_{people})	10	-	[12]
(Q_{people})	75	(W)	[12]
(N_{lamp})	20	-	[12]
(P_{lamp})	100	(W)	[12]

پارامترهای ارائه شده شامل مشخصات حرارتی پوسته ساختمان، شرایط مرزی، بارهای داخلی و شرایط بهره‌برداری مورد استفاده در محاسبات انرژی و اکسرژی هستند. مقادیر هندسی و ضرایب انتقال حرارت از منابع معتبر استخراج شده و سایر شرایط عملیاتی مطابق مشخصات ساختمان مورد مطالعه و فرضیات پژوهش تعیین شده‌اند. بر این اساس، نتایج تحلیل انرژی برای فصول تابستان و زمستان در جدول ۱۱ ارائه شده است.

جدول ۱۱. نتایج آنالیز انرژی

Table 11. Energy analysis results

Season	$T_i(K)$	Window type	$Q_{\text{conduction}}^1 (W)$	$Q_{\text{HVAC}} (W)$
Summer	295	Single-glazed	6427.66	19728.90
Summer	295	Double-glazed	4424.26	14920.74
Winter	294	Single-glazed	14140.86	4840.30
Winter	294	Double-glazed	9733.38	2174.11

براساس نتایج، استفاده از پنجره‌های دوجداره در هر دو فصل موجب کاهش انتقال حرارت از پوسته ساختمان و در نتیجه کاهش بار مورد نیاز سیستم HVAC شده است. این کاهش در فصل زمستان محسوس‌تر است. زیرا اختلاف دمای داخل و خارج ساختمان بیشتر بوده و سهم انتقال حرارت در بار گرمایشی افزایش می‌یابد. در مقابل، در فصل تابستان نیز اگرچه اختلاف دما کمتر است، اما کاهش ضریب انتقال حرارت پنجره سبب کاهش

^۱ مجموع ضریب انتقال حرارت هدایتی از دیوار و پنجره

بار سرمایشی مورد نیاز سیستم می‌شود. این نتایج نشان می‌دهد که بهبود عملکرد حرارتی پوسته ساختمان می‌تواند مستقیماً موجب کاهش تقاضای انرژی سیستم تهویه مطبوع شود.

از سوی دیگر، نتایج بخش یادگیری ماشین نیز نشان داد که دمای داخل، دمای محیط و تابش خورشیدی از مهم‌ترین متغیرهای مؤثر بر پیش‌بینی بارهای گرمایشی و سرمایشی هستند. از آن جا که همین پارامترها مستقیماً در محاسبات انرژی ساختمان نیز نقش دارند، نتایج دو بخش از نظر فیزیکی یکدیگر را تأیید می‌کنند.

به عبارت دیگر، مدل‌های یادگیری ماشین تنها قادر به پیش‌بینی دقیق بارهای حرارتی نبوده‌اند، بلکه اهمیت متغیرهای شناسایی شده توسط آن‌ها نیز با رفتار ترمودینامیکی ساختمان در تحلیل انرژی سازگار است.

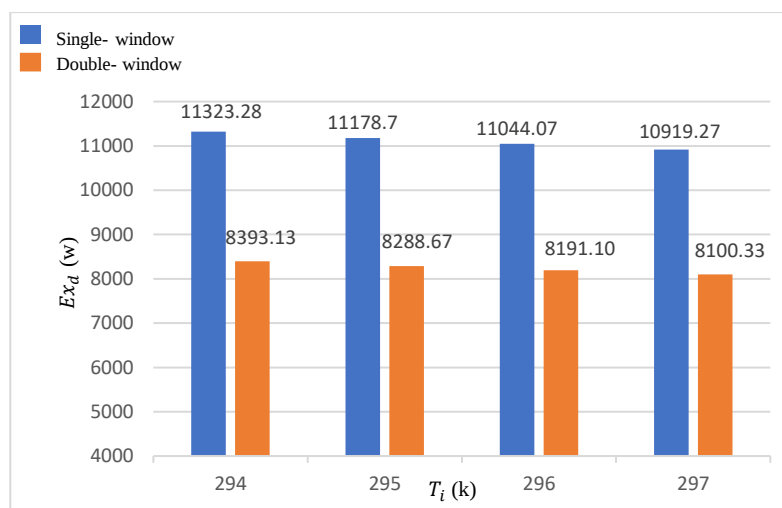
در ادامه، نتایج آنالیز اکسرژی برای ساختمان با پنجره‌های تک‌جداره و دوجداره، در فصول تابستان و زمستان در جدول ۱۲ ارائه شده است.

جدول ۱۲. نتایج آنالیز اکسرژی
Table 12. Exergy analysis results

Season	T_i (K)	Window type	Ex_d (W)
Summer	295	Single-glazed	11178.7
Summer	295	Double-glazed	8288.67
Winter	294	Single-glazed	7290.57
Winter	294	Double-glazed	5354.56

نتایج نشان می‌دهد که استفاده از پنجره‌های دوجداره موجب کاهش تخریب اکسرژی در هر دو فصل شده است. اگرچه کاهش مصرف انرژی اهمیت زیادی دارد، اما کاهش تخریب اکسرژی بیانگر استفاده مؤثرتر از کیفیت انرژی و کاهش برگشت‌ناپذیری‌های ترمودینامیکی در ساختمان است. بنابراین، استفاده از پنجره‌های دوجداره علاوه بر کاهش بار سیستم HVAC، موجب افزایش بهره‌وری ترمودینامیکی ساختمان نیز می‌شود.

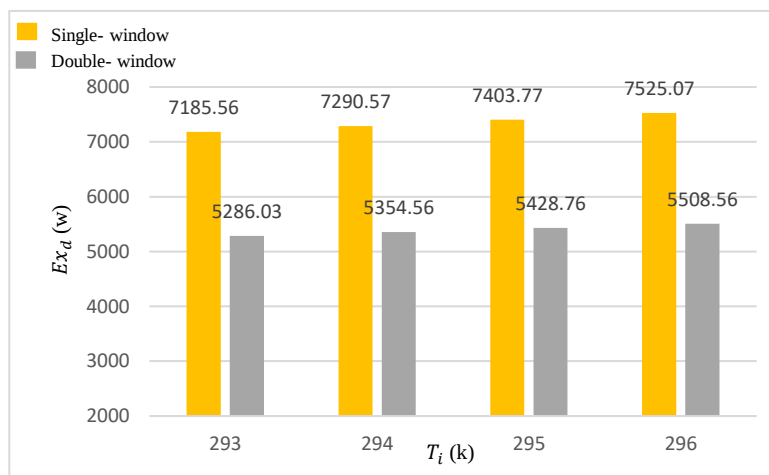
به منظور بررسی تأثیر تغییر دمای داخل ساختمان بر میزان تخریب اکسرژی، تغییرات Ex_d برای پنجره‌های تک‌جداره و دوجداره در بازه دمایی ۲۹۴ تا ۲۹۷ کلین در فصل تابستان در شکل ۹ ارائه شده است.



شکل ۹. تغییرات تخریب اکسرژی بر حسب دمای داخل ساختمان در فصل تابستان برای پنجره‌های تک‌جداره و دوجداره

Figure 9. Variation of exergy destruction with indoor temperature during summer for single- and double-glazed windows

همانطور که مشاهده شد، با افزایش دمای داخل ساختمان از ۲۹۴ به ۲۹۷ کلوین، میزان تخریب اکسرژی در هر دو نوع پنجره روندی کاهشی دارد. این رفتار ناشی از کاهش اختلاف دمای بین محیط داخل و خارج ساختمان است که موجب کاهش انتقال حرارت و در نتیجه کاهش برگشت‌ناپذیری‌های ترمودینامیکی می‌شود. همچنین، در تمامی دماهای بررسی شده، مقادیر تخریب اکسرژی در ساختمان مجهز به پنجره دوجداره کمتر از حالت تک‌جداره است. این موضوع نشان می‌دهد که کاهش ضریب انتقال حرارت کلی پنجره موجب محدود شدن انتقال حرارت ناخواسته، کاهش اتلاف کیفیت انرژی و در نهایت بهبود عملکرد ترمودینامیکی ساختمان می‌شود. به منظور بررسی اثر دمای داخل ساختمان بر تخریب اکسرژی در فصل زمستان، تغییرات Ex_d برای پنجره‌های تک‌جداره و دوجداره در بازه دمایی ۲۹۳ تا ۲۹۶ کلوین در شکل ۱۰ ارائه شده است.



شکل ۱۰. تغییرات تخریب اکسرژی بر حسب دمای داخل ساختمان در فصل زمستان برای پنجره‌های تک‌جداره و دوجداره

Figure 10. Variation of exergy destruction with indoor temperature during winter for single- and double-glazed windows

مطابق شکل، با افزایش دمای داخل ساختمان از ۲۹۳ به ۲۹۶ کلوین، میزان تخریب اکسرژی در هر دو نوع پنجره روندی افزایشی دارد. در فصل زمستان، افزایش دمای داخل موجب افزایش اختلاف دمای بین فضای داخل و محیط بیرون شده و در نتیجه انتقال حرارت از پوسته ساختمان افزایش می‌یابد. این امر باعث افزایش برگشت‌ناپذیری‌های ترمودینامیکی و در نهایت افزایش تخریب اکسرژی می‌شود. با این وجود، در تمامی دماهای بررسی شده، ساختمان مجهز به پنجره دوجداره مقادیر کمتری از تخریب اکسرژی را نسبت به پنجره تک‌جداره نشان می‌دهد که بیانگر نقش مؤثر عایق‌بندی حرارتی در حفظ کیفیت انرژی و کاهش اتلاف‌های ترمودینامیکی است.

هم‌راستایی این نتایج با تحلیل اهمیت ویژگی‌ها در بخش یادگیری ماشین، نشان می‌دهد که دمای داخل نه تنها یکی از مهم‌ترین متغیرهای مؤثر در پیش‌بینی بارهای حرارتی است، بلکه از دیدگاه ترمودینامیکی نیز نقش تعیین‌کننده‌ای در میزان مصرف و کیفیت انرژی ساختمان دارد.

۴. نتیجه‌گیری

در این پژوهش، عملکرد هشت الگوریتم یادگیری ماشین برای پیش‌بینی بارهای گرمایشی و سرمایشی ساختمان بر روی سه مجموعه داده مستقل ارزیابی شد. نوآوری این پژوهش در ارائه یک چارچوب تلفیقی برای ارزیابی عملکرد الگوریتم‌های یادگیری ماشین در سه مجموعه داده با شرایط اقلیمی و عملیاتی متفاوت و پیوند نتایج پیش‌بینی با تحلیل انرژی و اکسرژی است که علاوه بر افزایش قابلیت تعمیم مدل‌ها، امکان تفسیر مهندسی نتایج و ارزیابی کیفیت مصرف انرژی ساختمان را فراهم می‌سازد.

نتایج نشان داد که مدل‌های مبتنی بر درخت، به ویژه Random Forest، از دقت و پایداری بالاتری نسبت به سایر الگوریتم‌ها برخوردار بوده و قابلیت مناسبی در پیش‌بینی بارهای حرارتی ساختمان دارد. همچنین، تحلیل اهمیت ویژگی‌ها نشان داد که دمای داخل، دمای خارج، رطوبت نسبی و تابش خورشیدی مهم‌ترین عوامل مؤثر بر تغییرات بارهای حرارتی هستند که با رفتار فیزیکی سیستم‌های تهویه مطبوع و اصول انتقال حرارت مطابقت دارد. نتایج تحلیل انرژی و اکسرژی نیز بیانگر آن بود که استفاده از پنجره‌های دوجداره موجب کاهش انتقال حرارت، کاهش بار حرارتی و کاهش تخریب اکسرژی ساختمان می‌شود. بنابراین، تلفیق مدل‌های یادگیری ماشین با تحلیل‌های ترمودینامیکی می‌تواند ابزاری مناسب برای پیش‌بینی بارهای حرارتی، تحلیل عملکرد انرژی ساختمان و پشتیبانی از تصمیم‌گیری در طراحی و بهره‌برداری از سیستم‌های HVAC فراهم کند. در مطالعات آینده، ارزیابی مدل‌ها بر روی ساختمان‌هایی با کاربری‌ها و اقلیم‌های متنوع‌تر و همچنین استفاده از روش‌های پیشرفته‌تر یادگیری ماشین و یادگیری عمیق و همچنین انجام آنالیز حساسیت، می‌تواند به افزایش قابلیت تعمیم و دقت پیش‌بینی کمک کند.

۵. قدردانی

نویسندگان این پژوهش مراتب قدردانی خود را از دانشگاه شیراز به پاس فراهم‌آوری امکانات نرم‌افزاری و سخت‌افزاری پیشرفته که انجام این مطالعه را در سال ۱۴۰۵ میسر ساخت، ابراز می‌دارند.

۶. منابع

Reference

- [1] Guo, Y., Tan, Z., Chen, H., Li, G., Wang, J., Huang, R., Liu, J., & Ahmad, T. (2018). Deep learning-based fault diagnosis of variable refrigerant flow air-conditioning system for building energy saving. *Applied Energy*, 225, 732–745.
<https://doi.org/https://doi.org/10.1016/j.apenergy.2018.05.075>

- [2] Jafari Nasr, M. R., Chavoshi Gilani Nejad, S. (2018). Thermodynamic and mathematical methods in energy optimization of Tehran refinery atmospheric distillation unit. *Iranian Chemical Engineering Journal*, 17(99), 6-19, In Persian.
- [3] Hsu, P.-C., Gao, L., Hwang, Y., & Radermacher, R. (2025). A review of the state-of-the-art data-driven modeling of building HVAC systems. *Energy and Buildings*, 342, 115881. <https://doi.org/https://doi.org/10.1016/j.enbuild.2025.115881>
- [4] M. Esrafilian-Najafabadi and F. Haghighat. (2022). Impact of occupancy prediction models on building HVAC control system performance: Application of machine learning techniques. *Energy and Buildings*, 257, 111808.
- [5] M. Alaraj, M. Parodi, M. Radi, E. Alsisi, M. F. Abbod, and M. Majdalawieh. (2023). Occupancy-based one-year-ahead heating, ventilation, and air-conditioning electricity consumption optimization using machine learning. *Journal of Building Engineering*, 80, 108051.
- [6] Wang, H., Chen, X., Vital, N., Duffy, E., & Razi, A. (2024). Energy optimization for HVAC systems in multi-VAV open offices: A deep reinforcement learning approach. *Applied Energy*, 356, 122354. <https://doi.org/https://doi.org/10.1016/j.apenergy.2023.122354>
- [7] K. Irshad, A. I. Khan, M. E. Zayed, S. Algarni, and T. Alqahtani. (2024). Performance analysis of air conditioning system integrated with thermal energy storage using enhanced machine learning modelling coupled with fire hawk optimizer. *Journal of Building Engineering*, 98, 111308.
- [8] Li, J., Wang, M., Fang, P., Zhao, Q., Zheng, X., & Gao, H. (2024). Modeling and optimization of a passive building HVAC system based on improved PSO and H-J Algorithm. *Energy Reports*, 11, 2415–2438. <https://doi.org/https://doi.org/10.1016/j.egy.2024.02.011>
- [9] Kong, D., Hong, Y., Yang, Y., Gu, T., Fu, Y., Ye, Y., ... & Zhang, Z. (2025). A parametric, control-integrated and machine learning-enhanced modeling method of demand-side HVAC systems in industrial buildings: A practical validation study. *Applied Energy*, 379, 124971.
- [10] Athanasios Tsanas, & Angeliki Xifara. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49, 560-567. <https://doi.org/10.1016/j.enbuild.2012.03.003>.
- [11] Kaya Keleş, M., Keleş, A. E., Kavak, E., & Górecki, J. (2025). Comparison of Nature-Inspired Optimization Models and Robust Machine-Learning Approaches in Predicting the Sustainable Building Energy Consumption: Case of Multivariate Energy Performance Dataset. *Sustainability*, 17(23), 10718.
- [12] Yoon, Y., Jung, S., Im, P., & Gehl, A. (2022). Datasets of a Multizone Office Building under Different HVAC System Operation Scenarios. *Scientific Data*, 9(1), 775. <https://doi.org/10.1038/s41597-022-01858-6>
- [13] Melgaard, S. P., Johra, H., Nyborg, V. Ø., Marszał-Pomianowska, A., Jensen, R. L., Kantas, C., Larsen, O. K., Hu, Y., Frandsen, K. M., Larsen, T. S., Svidt, K., Andersen, K. H., Leiria, D., Schaffer, M., Frandsen, M., Veit, M., Ussing, L. F., Lindhard, S. M., Pomianowski, M. Z., ... Heiselberg, P. K. (2024). Detailed operational building data for six office rooms in Denmark: Occupancy, indoor environment, heating, ventilation, lighting and room control monitoring with sub-hourly temporal resolution. *Data in Brief*, 54, 110326. <https://doi.org/https://doi.org/10.1016/j.dib.2024.110326>
- [14] Hosamo, H., & Mazzetto, S. (2024). Performance evaluation of machine learning models for predicting energy consumption and occupant dissatisfaction in buildings. *Buildings*, 15(1), 39.
- [15] Tatachar, A. V. (2021). Comparative Assessment of Regression Models Based on Model Evaluation Metrics. *International Research Journal of Engineering and Technology*, 8, 853-860.
- [16] I. Dincer and M. A. Rosen. (2021). "Exergy, energy, environment and sustainable development", Third edition, Elsevier.
- [17] Shoeibi, H., Mehrpooya, M., Assareh, E., Izadi, M., & Pourfayaz, F. (2022). Transient simulation and exergy analysis of heat-pump systems integrated with solar compound parabolic collector. *Iranian Chemical Engineering Journal*, 6(44), 2121-2134, In Persian.
- [18] Rodríguez, V., Melgar, G., Márquez, A., & Manuel, J. (2024). Comparative analysis of the thermal transmittance (U-value) of a multilayer brick cladded façade in hot and cold periods, *proceedings of the REHABEND 2024 Congress*, Gijón, Spain, 1458-1464.

- [19] Tronchin, L., & Fabbri, K. (2008). Analysis of buildings' energy consumption by means of exergy method. *International Journal of Exergy*, 5(5–6), 605–625. <https://doi.org/10.1504/IJEX.2008.020828>
- [20] Altmann, A., Tološi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10), 1340-1347.
- [21] Foroushani, S., Wright, J. L., & Naylor, D. (2016). Sensitivity of the solar heat gain coefficient of complex fenestration systems to the indoor convection coefficient. *Simulation*, 12, 277–285.
- [22] ASHRAE. (2021). "*ASHRAE Handbook – Fundamentals*". American Society of Heating, Refrigerating and Air-Conditioning Engineers.