Review

Machine Learning and Deep Learning Methods for Enhancing Building Energy Efficiency and Indoor Environmental Quality – A Review

Paige Wenbin Tien [*], Shuangyu Wei, Jo Darkwa, Christopher Wood, John Kaiser Calautit

Department of Architecture and Built Environment, University of Nottingham, Nottingham NG7 2RD, UK

HIGHLIGHTS

- Machine and deep learning models for building energy efficiency, thermal comfort and air quality.
- Applications of machine and deep learning models and trends in the built environment sector.
- Challenges of adopting machine and deep learning in the built environment sector.
- Limited studies applied machine and deep learning-based strategies in real buildings and conducted post-occupancy evaluation.
- Limited machine and deep learning studies investigated energy use, thermal comfort, and indoor air quality simultaneously.

GRAPHICAL ABSTRACT
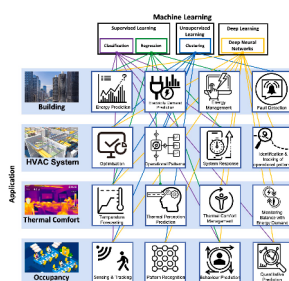
ARTICLE INFO

ABSTRACT

The built environment sector is responsible for almost one-third of the world's final energy consumption. Hence, seeking plausible solutions to minimise building energy demands and mitigate adverse environmental impacts is necessary. Artificial intelligence (AI) techniques such as machine and deep learning have been increasingly and successfully applied to develop solutions for the built environment. This review provided a critical summary of the existing literature on the machine and deep learning methods for the built environment over the past decade, with special reference to holistic approaches. Different AI-based techniques employed to resolve interconnected problems related to heating, ventilation and air conditioning (HVAC) systems and enhance building performances were reviewed, including energy forecasting and management, indoor air quality and occupancy comfort/satisfaction prediction, occupancy detection and recognition, and fault detection and diagnosis. The present study explored existing AI-based techniques focusing on the framework, methodology, and performance. The literature highlighted that selecting the most suitable machine learning and deep learning model for solving a problem could be challenging. The recent explosive growth experienced by the research area has led to hundreds of machine learning algorithms being applied to building performance-related studies. The literature showed that existing research studies considered a wide range of scope/scales (from an HVAC component to urban areas) and time scales (minute to year). This makes it difficult to find an optimal algorithm for a specific task or case. The studies also employed a wide range of evaluation metrics, adding to the challenge. Further developments and more specific guidelines are required for the built environment field to encourage best practices in evaluating

* Corresponding author.
   E-mail addresses: paige.tien@nottingham.ac.uk, paige.tien@gmail.com (P.W. Tien).

and selecting models. The literature also showed that while machine and deep learning had been successfully applied in building energy efficiency research, most of the studies are still at the experimental or testing stage, and there are limited studies which implemented machine and deep learning strategies in actual buildings and conducted the post-occupancy evaluation.

Nomenclature and abbreviations

AI	Artificial intelligence
ACEEE	American Council for an Energy-Efficiency Economy
ANN	Artificial neural network
AWS	Amazon Web Service
BD	Big data
BES	Building energy simulation
BEMS	Building energy management systems
BICT	Building-integrated control testbed
BIM	Building information modelling
BN	Building network
BMS	Building management systems
BPS	Building performance simulation
CART	Classification and regression tree
CIBSE	Chartered Institution of Building Services Engineers
CL	Cooling load
CNN	Convolutional neural network
CO_2	Carbon dioxide
CV	Coefficient of variance
DCC	Demand-driven cooling control
DM	Diebold-Mariano
DNN	Deep neural network
DL	Deep learning
EFB	Energy-efficient buildings
ELM	Extreme learning machine
EUG	End-user group
GA	Genetic algorithm
GPR	Gaussian process regression
GPS	Global positioning system
GPU	Graphical processing unit
GM	Gaussian Mixture
HL	Heating load
HME	Hierarchical mixture of experts
HMM	Hidden Markov model
HVAC	Heating, ventilation, and air conditioning
IAQ	Indoor air quality
IR	Infrared
IoT	Internet of Things
KNN	k-nearest neighbour
Labs	Laboratory
LR	Linear regression
LSTM	Long short-term memory
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MARE	Mean absolute relative error
MCC	Matthews correlation coefficient
ML	Machine learning
MLP	Multilayer perceptron
MOGA	Multi-Objective Generic Algorithm
MRE	Mean relative error
MRT	Mean radiant temperature
MSE	Mean squared error
MSPE	Mean square percentage error
NLP	Natural language processing
NMBE	Normalized mean bias error
NN	Neural network
OSELM	Online sequential extreme learning machine
PIR	Passive infrared
PMV	Predicted mean vote
PSO	Particle swarm optimisation
Rad	Radiation
RBF	Radial basis function
RF	Random forest
RH	Relative humidity
RICS	Royal Institution of Charted Surveyors
RMSE	Root mean squared error
RMSPE	Root mean squared percentage error
RNN	Recurrent neural network
SD	Standard deviation
SNA	Social network analysis
RT	Random tree
STD	Standard deviation
R^2	R squared
SVM	Support vector machines
SVR	Support vector regression
TS	Thermal sensation
UK	United Kingdom
3D	Three-dimensional

1. Introduction

According to the latest projections, the global climate is predicted to continue to change, and the frequency of extreme weather and climate events is expected to increase. This will significantly impact the built environment sector, and new building designs should be able to cope with climate change effects and meet future energy needs [1]. For instance, heating, ventilation and air-conditioning (HVAC) systems are responsible for up to 40% of the energy consumed by buildings in the commercial sector [2]. Developing new technologies and solutions that can minimise its consumption can significantly reduce emissions from the built environment sector. However, thermal comfort and indoor air quality are also important factors that must be considered in the design of buildings and HVAC and should not be neglected when seeking

solutions for reducing the building energy demand [3].

An example of this solution is the integration of building energy management systems (BEMS) to automatically control building operations, including HVAC [4], lighting and equipment [5–7]. According to the report [8], energy savings of 18% for offices and 14% for retail stores can be achieved by employing smart technologies and analytics. BEMS ensures building services, systems, and equipment operate optimally by reducing energy consumption, along with the reduction in operational costs and emissions while providing a better-quality environment for occupants. BEMS are more automated and limits the need for manual procedures in monitoring and controlling HVAC systems. They have a significant role in improving the efficiency of buildings and will be a crucial strategy for developing truly intelligent buildings.

Due to computer and software technology advancements, different sectors are being taken over by automated machines and software, and

the field of artificial intelligence (AI) is becoming more important [9]. This has led to smart solutions for buildings that optimise energy performances and reduce resource waste [10,11] without compromising comfort, health or security [12]. The increasing adoption of the Internet of Things (IoT) and AI technologies for building monitoring and controls will drive the smart building market's growth. According to the Google Trends results [13], since 2017, the popularity of both machine learning and deep learning has overtaken IoTs.

More and more academic researchers and building professionals are developing and utilising AI-based solutions for the design and construction [14], operation and maintenance [14,15] of the built environment. An example is integrating AI algorithms and sensors into the indoor environment to optimise the process in real-time, such as monitoring and controlling the indoor climate. These systems can automatically analyse the data and provide future predictions of the building's behaviour and facilitate and assist the decision-making [16, 17]. However, the study by the McKinsey Global Institute [18] presented a statistical comparison of AI's use within various sectors. It is acknowledged that AI within the building and construction sector has been slow to employ AI and digital tools in comparison to other sectors. Accordingly, there is a need to study AI techniques for enhancing building energy efficiency and solving building-related problems, identifying the reasons for its slow adoption and potential solutions. This stresses the urgency for an in-depth review and exploration of the
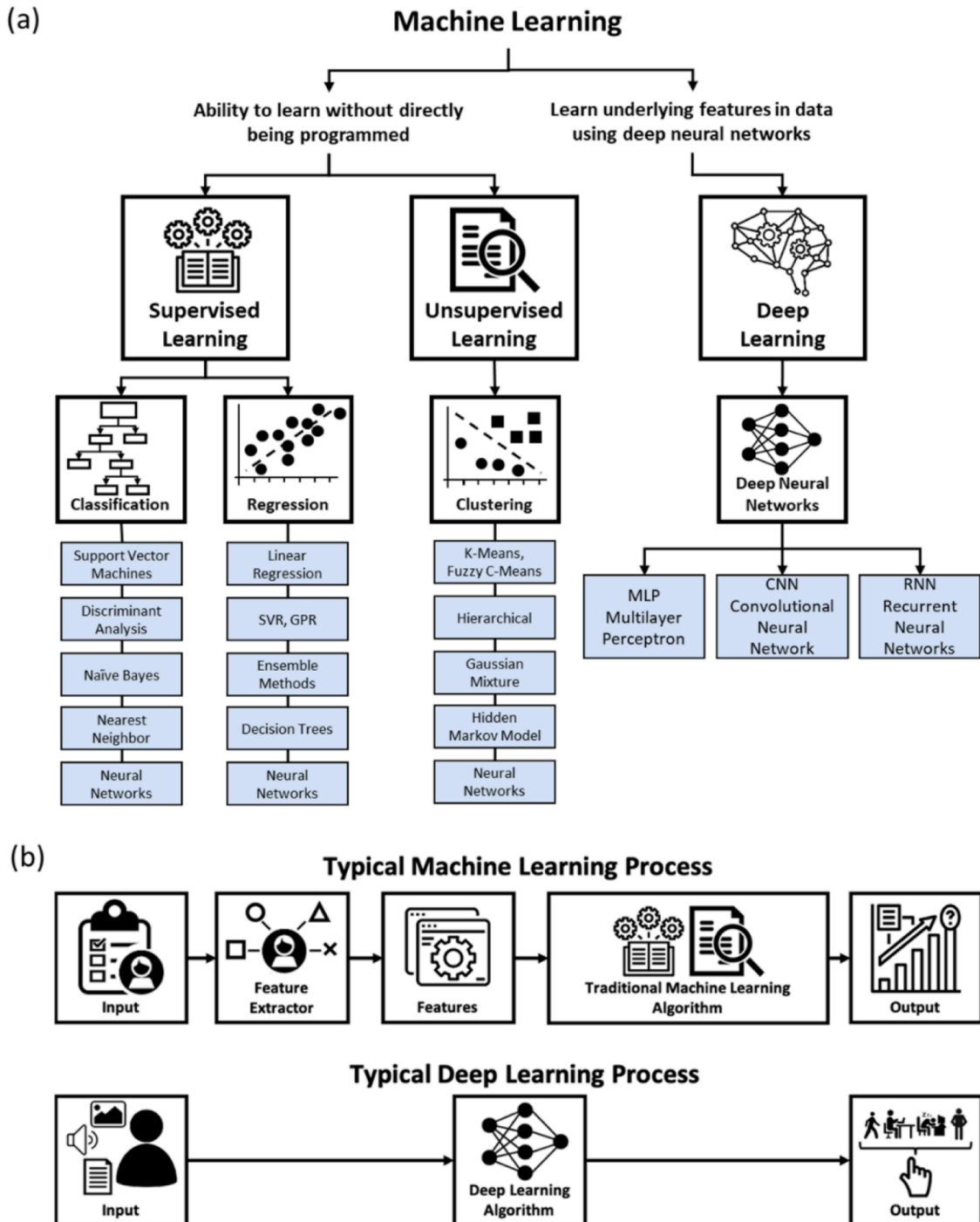

Fig. 1. (a) AI-based machine learning and deep learning techniques reviewed in this study (b) typical machine vs deep learning.

current use and how it can inform the development of AI solutions for future buildings.

Fig. 1a summarises the most common AI-based machine and deep learning techniques currently used within the built environment sector, particularly energy efficiency-related applications, which are reviewed in this study. In machine learning, data presented in numerical, categorical, time series and text are used as input [19] with the selection of an algorithm as a computational method to "learn" information directly from data. Deep learning interprets data features and their relationships using neural networks to form a unique model based on a wider range of data, including images, videos, and sounds. To a greater extent, deep learning provides higher accuracy than other methods as the feature extraction process is performed automatically from raw data. However, deep learning would require more data points to improve its accuracy. Several studies have suggested that deep learning surpassed machine learning and other learning algorithms in various applications [20]. It should be acknowledged that there are more AI techniques, and the present work will mainly focus on these.

Fig. 2 presents a summary of the applications of the machine and deep learning-based methods in the design of energy-efficient, comfortable, and healthy buildings, evaluated in this paper. The review focuses on supervised and unsupervised machine learning, and deep learning techniques that were applied to enhance building and HVAC system energy-efficiency, and improve indoor environment quality. The review of different machine and deep learning techniques will help identify the specific techniques that are more suited for each area. This enabled the formation of the connections shown in Fig. 2, detailed in Sections 3. and 4.

Supervised machine learning designed for classification and regression problems consists of algorithms that are trained using datasets that are fully labelled, i.e., features' data, providing an answer key that can be used to assess its accuracy. While in unsupervised machine learning, the algorithm attempts to make sense of the unlabeled data by extracting patterns and features on its own without clear instructions on what to do with them. This is useful when fully labelled datasets are not available and, in some cases, when the desired outcome or answer is not known. Another subset of machine learning that uses a multi-layered structure of algorithms to create an artificial neural network (ANN) is deep learning (DL). Deep learning offers several advantages over traditional machine learning methods and, in some cases, outperforms them. Deep learning networks do not require human intervention and can learn from their own mistakes (Fig. 1b). However, it can be costly in terms of computational power and time. Deep learning is usually applied to problems that require complex and unstructured data such as images,
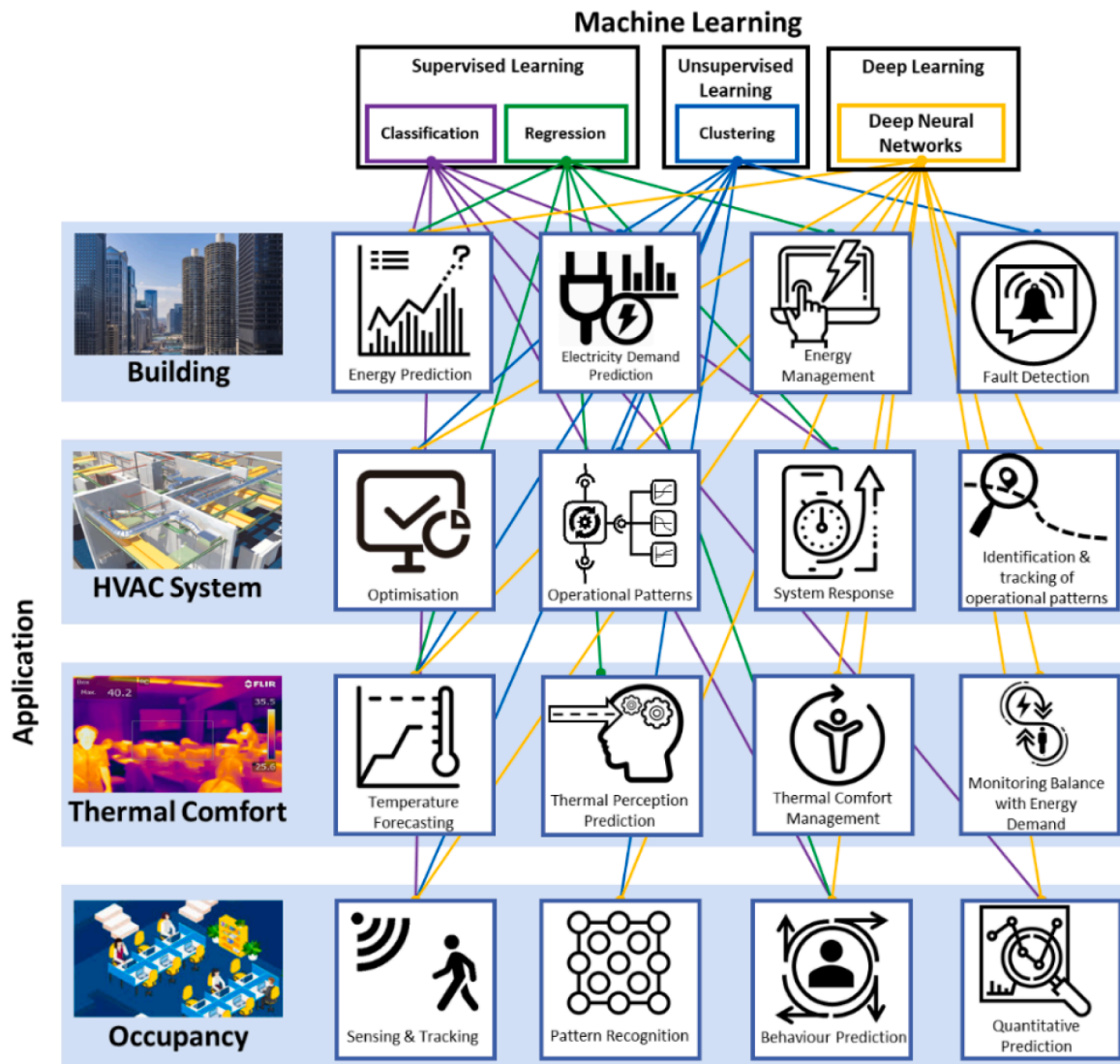

Fig. 2. Summary of the AI-based methods employed in the design of energy-efficient, comfortable, and healthy buildings, reviewed in this paper. Connections made correspond to the studies evaluated in Sections 3. and 4.

videos, and sound to perform tasks.

An example is detecting occupants in indoor spaces and using the information to control the operation of HVAC. In recent years, there have been significant developments in deep learning due to the increased available computing power and graphical processing unit (GPU) computing. Compared to supervised and unsupervised machine learning, there is limited research on deep learning techniques for building and energy-related applications; however, deep learning has recently gained more popularity [13]. This stresses the need to review its development and applications in the built environment.

1.1. Literature Gap and Aims and Objectives

Numerous review papers evaluated various aspects and applications of machine learning in the built environment. Published review papers specifically focus on evaluating an individual area within the wide range of built environment problems solved using AI approaches. This includes building energy consumption forecasting [21-23,30-34], integration with building energy management systems (BEMS) [23–25], building design optimisation [26] and occupancy detection [26–29]. Bordeau et al. [21] reviewed data-driven and machine learning techniques for modelling and forecasting buildings' energy consumption. The study focused on the characteristics of input data and methods for pre-processing. It concluded that a standardised protocol that can solve various problems is still lacking.

Similarly, Rätz et al. [22] reviewed machine learning algorithms for modelling building energy systems but focused on the framework and optimisation methods to develop an automated concept or toolbox. Amasyali and El-Gohary [23] explored the type and size of data used and the features that were selected for training. Other reviews, such as [30–33], focused on different AI-based algorithms for predicting energy consumption.

The studies [23–25] reviewed AI applications based on their design and integration with BEMS. The review covered different AI frameworks and workflows used for HVAC design and optimisation process and control. While the review of Machairas et al. [26] covered machine learning algorithms coupled with building simulation programs, which focused on optimisation methods for building design. Studies [27–29] reviewed various occupancy sensing and detection techniques based on integrating sensors and machine learning algorithms. The reviews focused on evaluating the 'best' method for occupancy sensing and detection. However, these did not fully consider how the occupancy information can be used to influence the energy demand of the buildings or the indoor environment.

The primary aim of this paper is to provide a critical summary of the existing literature on the machine and deep learning methods for the built environment over the past decade, with special reference to holistic approaches. The present study will explore existing AI-based techniques focusing on the framework, methodology, and performance, including the data acquired, model formation process, accuracy, and speed. This paper will review different AI-based techniques employed to resolve interconnected problems related to HVAC systems and enhance building performances, including energy forecasting and management, indoor air quality and occupancy comfort/satisfaction prediction, occupancy detection and recognition, and fault detection and diagnosis.

An extensive literature search was performed to identify publications on existing studies on the application of machine and deep learning methods for the built environment. Peer-reviewed journals, conference papers, technical reports and books from the last decade (with some exceptions made) were searched using the Scopus and ScienceDirect search engines. The search was carried out using keywords such as "artificial intelligence in buildings", "machine learning in built environment", "deep learning in built environment". We selected the articles based on the publication title and the abstract. Following a data collection process of identification, screening, eligibility analysis and inclusion (PRISMA method), we selected and reviewed 171 articles

(from an initial list of 362 articles).

2. Applications of AI within the Built Environment

AI is being adopted widely in various areas to perform tasks more efficiently while reducing the need for human effort. With the ever-increasing computational power and data availability in today's digital society, significant progress has been made in the field of AI in recent years [35]. Within the construction sector, building information modelling (BIM) is becoming the norm for developing new buildings and facilities. As an enabler of innovation and digitalisation in the sector, BIM provides a foundation for a digital world in which AI can help optimise design, construction and operation/facility management [36]. For example, with the help of AI, BIM can utilise a large amount of data from previous construction projects and automatically suggest solutions to optimise the design.

AI is driving the development of smart buildings, making them self-learning and adaptive rather than just automated. Smart buildings utilise advanced technologies to automatically control building operations, including HVAC systems, lighting, and security [37,38]. Fig. 3 shows the evolution of buildings from conventional to intelligent, along with the integrated systems and techniques such as AI and machine learning (ML) which equip buildings with an ability to learn and adapt [39,40]. Much research has been dedicated to using AI technologies in smart buildings, focusing on improving energy efficiency, thermal comfort, health, and productivity in the built environment. This section explores existing AI-based techniques which aim to achieve energy-efficient, comfortable, and healthy buildings.

2.1. Building Energy Demand Forecasting

Building energy demand forecasting is vital for optimising building energy performance. It assists energy planning, management, and conservation to inform strategies for reducing energy consumption and CO_2 emissions [23]. Energy forecasting is also used to evaluate building design alternatives and operational strategies to improve demand and supply management [42].

To perform predictions of future building energy usage using AI-based methods, existing historical data must be collected. Currently, data are collected via energy meters and sensors. Ahmad et al. and Avancini et al. [43,44] highlighted the technological advancements in building energy metering and environmental monitoring. Chammas et al. and Terroso-Saenz et al. [45,46] presented the application of wireless networks, sensors and IoT-based techniques to enable energy monitoring solutions that are low cost, highly accurate and easy to deploy. However, IoT devices can generate vast amounts of data; hence, integration with AI can help deal with such huge volumes [47]. Din et al. [48] identified that machine learning techniques are expected to pave the way for IoT networks, generating sophisticated visions and ideas for IoT systems. Wang and Srinivasan [32] highlighted that AI-based approaches had recently gained popularity due to the ease of use and adaptability to obtain optimal solutions rapidly while requiring less detailed physical parameters and information about the building.

Several works highlighted the importance of different external and internal parameters on prediction performance. Zhao and Liu [49] developed a machine learning-based building energy load forecasting solution with the proposed model achieving a high accuracy prediction of energy loads with a MARE of 2.60% for cooling and 3.99% for heating 1-h (one hour) ahead [50]. The study highlighted the importance of sufficient training of the model and selecting the types of input data to achieve such accuracy. The weather forecast precision affected the proposed model. When an MAE between the actual temperature and the forecasted temperature was 1°C, the MARE of the 24-h ahead loads raises to 2.01% for heating. Hence, dynamic load forecasting for different time horizons from 1 to 24-h ahead could be advantageous to HVAC control system optimisation.
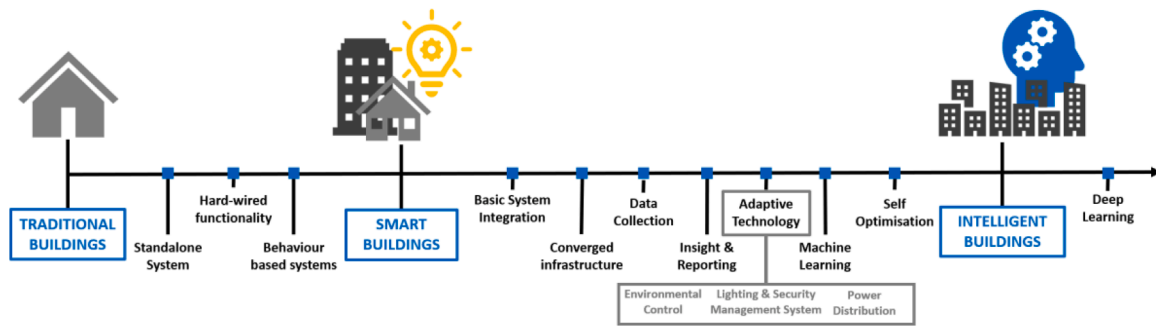
Fig. 3. Development and evolution of smart and intelligent buildings. Fig. adapted from [41].

While Kwok and Lee [51] highlighted the importance of occupancy in predicting building cooling load using an ANN model, addressing the issues and limitations identified in previous studies. Their results showed that the use of building occupancy data could significantly increase the accuracy of the cooling energy predictions compared to using fixed schedules or historical data to represent the occupancy in the building. While the study by Ding et al. [52] highlighted the importance of the interior variables when predicting the building heating load using

Table 1
Summary of AI-based techniques for energy management and predictions in the built environment.

Ref.	Building Type	Energy/System	Prediction Rate	Evaluation Metrics	Key finding/summary	Performance/Accuracy
[42] Fan et al. 2019	Research Academic	Cooling	Predictions for 24-h ahead	MAE, CV-RMSE	Recurrent models achieved the most accurate predictions without increasing computational load.	RMSE was 30.8% for cooling load prediction
[55] Pham et al. 2020	Office	HVAC	Predictions for 1 step, 12 steps and 24 steps ahead	MAE, RMSE, MAPE	Random Forest (RF) model was superior to the M5P (M5 Model Tree) and Random Tree (RT) models.	1-step-ahead prediction of RF was 49.21% better in MAE and 46.93% in MAPE than RT. 12 and 24-steps-ahead predictions of RF were 49.95% better in MAE and 29.29% in MAPE than M5P.
[45] Chammas et al. 2019	Office	Building energy Lighting	Not specified	R^2, RMSE, MAE, MAPE	Multilayer Perceptron (MLP) performed better against four classification models	64% R^2, RMSE 59.84%, MAE 27.28%, MAPE 27.09%.
[49] Zhao and Liu. 2018	Office	Heating Cooling	Predictions for 1-h, 2-h, 2-h, 24-h ahead	MARE, MAE	Dynamic load forecasting for different time horizons to optimise controls of HVAC systems.	1-h ahead: Cooling MARE 2.60%, Heating 3.99%, 24-h ahead MARE increases by 2.01%.
[51] Kwok and Lee. 2011	Office	Cooling	24h	RMSPE	Occupancy significantly affects cooling load predictions. Understanding occupancy behaviour improves the predictive accuracy of the models.	RMSPE 40.376 – 52.047 not employing occupancy factors, RMSPE 14.836 -30.090 not employing occupancy factors
[52] Ding et al. 2011	Office	Heating	Real-time prediction 24-h	MRE R^2	The prediction of the heating load is influenced by exterior variables and interior variables, including occupancy level and lighting/equipment use.	Prediction of heat load with only exterior variable: 84% R^2 Prediction of heat load with exterior and interior variable: 94% R^2
[53] Singaravel at al. 2018	Research Academic	HeatingCooling	Monthly prediction	R^2	Cooling energy prediction using the AI method was significantly faster than BES while providing accurate results.	Building energy simulation required 1145s to simulate while only 0.9 s for the AI model.
[54] Kumar et al. 2018	-	Heating Cooling	Real-time prediction	MAE	The extreme learning machine (ELM) models learned better and outperformed other popular machine learning approaches.	MAE (kW) 0.0348 and 0.0389 for both cooling and heating, along with the lowest heating load prediction time of 0.06s.
[56] Xu et al. 2019	Multi-building	Building Energy Use	Monthly, Annual	MAPE, RMSE	SNA-ANN model predicted the multi-building energy use with satisfactory accuracy providing an empirical approach to urban building energy use prediction.	MAPE 10.72% RMSE 14.52%, Accuracy of 90.28% for the predicted energy use for all building groups.
[57] Chou and Bui 2014	Residential	Heating Cooling	Not specified	MAPE, RMSE	The ensemble approach and support vector regression (SVR) were the best models for predicting heating and cooling load.	Ensemble approach: MAPE below 4%, 39.0%-65.9% lower RMSE compared to previous works.
[58] Tsanas and Xifara. 2012	Residential	Heating Cooling	Not specified	MAE, MSE, MRE	The heating load can be more accurately estimated than the cooling load.	Heat load estimation with 0.5 points deviation from ground truth, cooling load with 1.5 points deviation from the ground truth.
[59] Zhou and Zheng 2020	Residential, High-rise building	HVAC	Real-time prediction	NMBE, RMSE	Energy peak power reduction of up to 21.9% was achieved using ML for building demand prediction, integrated with a hybrid controller.	NMBE <10% CV-RMSE<30%.

AI models such as ANN and support vector machine (SVM). The interior variables include the indoor environment parameters, occupant level, artificial lighting and equipment operation. Their results showed that only considering exterior variables can ensure a high prediction accuracy ($R^2 = 84\%$) while considering both indoor and outdoor improved the accuracy further ($R^2 = 94\%$).

While building engineers and architects commonly use building energy simulation (BES) to predict the energy consumption of buildings, several factors/issues can lead to low-energy design solutions or performance energy gaps. This includes skills and knowledge of the modeller, use of simplification methods, assumptions, and the tools' quality. Hence, more researchers are attempting to address this using data-driven AI and machine learning approaches, which do not require detailed information about the building. Singaravel et al. [53] compared AI methods with BES in terms of accuracy and speed in predicting the building energy demand. Based on the results of 201 cases, the AI model predicted cooling energy with similar accuracy as BES, while it was slightly less accurate in terms of heating energy prediction. However, the AI model significantly reduced simulation time as compared to BES, a reduction of simulation time from 1145 seconds to 0.9 seconds. Finally, they also showed that the deep learning models performed slightly better than simple ANN models. The high-speed prediction compared to BES means more evaluation of design options and optimisation can be carried out or allow real-time predictions.

Kumar et al. [54] employed ML methods to improve the real-time heating and cooling load predictions' accuracy and efficiency. They used extreme learning machine (ELM) for applications when complete data is available and online sequential ELM (OSELM) methods for applications when data comes in real-time. In addition, they highlighted the importance of using significant building design and structural attributes in predicting heating and cooling loads, such as relative compactness, glazing and roof area and orientation. Their results showed that the models learned better and outperformed other popular machine learning approaches. The proposed model took less than 0.5 s time to predict.

Table 1 presents a summary of the previous works reviewed in this section. This explores the different AI-based techniques used for building energy forecasting and the different building types, energy systems, prediction interval and evaluation metrics used in previous work. The evaluated studies suggest that many methods are evaluated or tested in office and academic buildings. It can also be seen that many works used different types of evaluation metrics to assess and compare the performance of the models. The studies have shown the advantage of AI methods for predicting energy loads compared to conventional BES models. It requires fewer details and information about the building, which reduces the time of developing the model and, at the same time, AI-based models are significantly faster. However, it is important to note that the AI-based model's accuracy and reliability rely on the input data, and users must select a suitable learning algorithm for their prediction model. Due to the reliance on historical building data, AI-based models application in the design stage are limited. Furthermore, one cannot extrapolate the prediction results once changes are made to the design and operation of a building.

Furthermore, as pointed out by [51,52], indoor parameters such as occupancy level and behaviour can significantly impact the building energy use and the prediction results. Studies such as [60] suggest integrating occupancy behaviour pattern recognition with the energy load forecasting model to enhance prediction performance. The following section explores occupancy behaviour within buildings and different methods to collect occupancy information, including AI-based prediction and detection strategies.

2.2. Occupancy Behaviour & Detection Strategies

The amount of energy consumed by buildings is influenced by various factors, from thermo-physical properties of the building elements to the location, occupancy behaviour and the HVAC systems [61]. Although outdoor environmental conditions significantly impact building energy consumption, the variations in occupancy rate and their behaviour are equally important. The number of occupants, their activity level and how they use the equipment can impact the internal heat gains, indoor environment and energy demand. Occupants also interact with the building and make personal adjustments such as the thermostat or opening the windows. In practice, conventional HVAC is typically controlled using "static" or "fixed" operation schedules, leading to unnecessary energy usage, such as when spaces are left unoccupied [62]. Similarly, traditional building energy models use "static" and deterministic occupancy inputs, leading to prediction errors [63]. This can be seen in the examples in Fig. 4, which compares the occupancy heat gains profile for an office building and an assumed "static" heat gains profile.

Clearly, this can result in uncertainties in the building energy prediction, difficulty in sizing and controlling HVAC systems [64], and not meeting the desired indoor conditions and comfort requirements [65]. Hence, occupancy behaviour and its impact on the energy performance of buildings have gained significant interest within the scientific community [66]. This led to the development of advanced occupancy detection techniques and occupancy simulators [67]. The occupancy data can help determine the effects of occupant presence and their activities within buildings, which can be used to optimise HVAC and lighting controls [68]. Conventional occupancy detection methods such as motion sensors can estimate the number of people within the desired space. While recently, more advanced methods such as WiFi-enabled IoT devices are used to identify occupants' activities [69,70] automatically. This is made feasible by the wide availability of WiFi infrastructure and the occupant's mobile WiFi-connected devices [71].

The activity recognition solution proposed by Zou et al. [69] called the 'Deep Hare' was integrated with a deep learning technique to enhance occupancy activity recognition. The approach can distinguish between the different activities performed over time with an accuracy of up to 97.6%. Wang et al. [72] proposed a WiFi probe-based occupancy detection method, which uses a Markov-based feedback recurrent neural network algorithm. The study showed that it could predict occupancy with accuracies between 80.9%-93.9%. In a recent study, Wang et al. [73] employed the WiFi probe-based occupancy detection method and showed that it could save up to 26.4% of energy demand, based on experiment and simulation results.

Other methods used more conventional sensors such as RFID and environmental sensors. Carreira et al. [74] used radio frequency identification (RFID) to estimate the occupancy number in a room. Like the previous works, machine learning was incorporated and automatically enabled HVAC management to reduce energy demand while maintaining comfort levels. Jiang et al. [75] estimated the number of indoor occupants in real-time based on the carbon dioxide (CO2) levels and an extreme machine learning model. The results showed that the proposed method could accurately estimate the occupant number up to 94% based on field tests.

Some researchers employed cameras integrated with AI for sensing occupancy. Zou et al. [76] used existing surveillance video data and a deep learning approach to measure occupancy for building energy conservation. The experimental results showed an accuracy of up to 95.3% achieved using the approach with low computational requirements. While Diraco et al. [77] used 3D depth sensors to count and localise occupants in buildings while assuring the occupant's privacy as the depth information. Table 2 summarises the different occupancy detection techniques developed and used in current research, mainly for building applications. The benefits varied between each type of sensor depending on the desired applications.

The newer techniques such as Wi-fi, wireless sensors and cameras are increasingly being employed in research studies for occupancy studies and, at the same time, integrated with AI techniques. The camera is one of the most popular sensing techniques for indoor environments and human recognition. Similar limitation problems arise from using a

Fig. 4. (a). Example of an actual occupancy heat gains profile for a typical office day compared to (b). A typical static occupancy heat gains profile [62].

camera for detection; however, significant effort has been carried out in recent research to enhance the abilities to use the camera through AI adaptation [99].

The utilisation of camera-based techniques for occupancy detection has been increasing recently due to the advancement of deep learning-based techniques [100], such as convolutional neural network (CNN). Deep learning interprets data features and relationships solely using neural networks to form a unique model designed for the desired application, ultimately providing greater flexibility, performance, and accuracy.

The proposed framework process that Ijjina and Chalavadi [99] used for human action recognition emphasises motion in different temporal regions to achieve better discrimination among actions. It suggests that a video is used as input data into a CNN model, extracting its features. Within the model, a classifier is trained to recognise the human actions to predict activity. The strategy developed by Castro et al. [101] to predict occupants' daily activities using egocentric images is similar to [99]; corresponding stages incorporated within the framework were used to develop a workflow to enable activity prediction. Several stages of training were performed to enable in-depth feature extraction refinement before producing the final output classifications. Overall, applying the CNN model alone gave an accuracy of 78.56%, and maximum accuracy of 83.07% could be reached when an ensemble approach was applied. Fig. 5 provides an example of a workflow process for the development of an AI-based technique for occupancy detection in an indoor environment.

Fig. 6 and Video 1 present an example application of an AI vision-based camera detection within an office environment. The cameras employed an AI-based model, which was trained following the procedure given in Fig. 5, enabling the camera to provide equipment detection by identifying the presence of PC monitors that were turned ON. The camera can also provide the application of occupancy activity detection. This enabled the identification of times when occupants are performing the activities of sitting, standing, and walking and when there is no occupancy.

Video 1 Example occupancy and equipment detection within an office environment.

Based on the review of literature, different sensors and detection solutions have different merits and limitations. Through evaluation, camera detection with AI techniques seems to be a promising approach for indoor occupancy detection. As identified, existing AI-based occupancy detection methods mostly utilise the camera for detection and recognition purposes. Most studies have not attempted to integrate the vision-based occupancy detection approach with the HVAC control systems. Furthermore, the impact of the application of such approaches on energy demand and thermal comfort has not been well studied. The development of AI camera-based techniques for occupancy detection is further discussed in Section 4.

2.3. Thermal Comfort and Air Quality

People spend most of their time indoors, and hence comfortable and

Table 2
Overview of various types of sensors used for obtaining occupancy information.

Sensor Types	Specific Sensor Type	Features	Advantages	Limitations	Ref.
Environmental	Temperature Humidity CO_2	Detect change in an environment due to the presence and activities of occupants	Low cost, Commercially available, Non-intrusive	Requires regular calibration to ensure stability Can be affected by varying airlfow rates, Time delay	[78] [79]
Dual-Technology	PIR & ultrasonic	Combined PIR and ultrasonic technology minimise false-alarm	Self-adaptive, Eliminates falsealarm	Reduced detection probabilities, More costly	[80]
Passive Infrared (PIR)		Use IR to detect a difference in heat emitted by moving people and background heat	Low cost, Commercially available, Non-intrusive, Easy detection	Limited to movements, Requires direct line of sight	[81]
Ultrasonic		Observe frequency changes caused by occupants	Detects minor motion, Does not require an unobstructed line of slight	High levels of vibration or airflow complicates their application	[82]
Electromagnetic (EM)e	Microwav	>Emits and receives microwave signals to detect motion	Suitable for various environments, including high heat, Wide detection range	More prone to false alarm, High operation cost, Continuous power draw	[83]
Acoustic/Sound	Audio	Detect changes in the acoustic/sound wave, associated with certain occupancy activity	Low cost, Commercially available, Less intrusive	May require other sensors, Not suitable for some environments such as labs and libraries	[84, 85]
Radio Frequency Identification		Uses radio waves to transmit information from tag to reader.	Automatic , Real-time response	Require users to carry a card/tag	[74, 86]

Table 2 (*continued*)
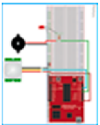
Sensor Types	Specific Sensor Type	Features	Advantages	Limitations	Ref.
Door	Beam counter	IR based electronic device to measure footfall trends	Non-intrusive, Low-cost	Accuracy tends to decrease when several people are detected	[87]
Pressure	Pressure pads	Monitors occupancy, Automatically alarm when detected unexpected activity	Sensitivity, Simple and robust	Can only be applied to specific locations	[88]
	Air pressure change	Detect change in air pressure within the environment due to the behaviour of occupants	Non-intrusive, Senses movement between spaces	Relationship to occupancy can be indirect	[89]
Smart Meters		The energy consumption patterns from smart meters can be used to infer occupancy	Does not require other dedicated sensors	Privacy concerns, It can give false results when appliances or systems are in use when occupants are away	[90]
Wireless	Bluetooth	Collects and monitors bluetooth signals (related to occupancy) in real-time, Transmit data to the sensor or IoT based infrastructure	Low-power, Can utilise exisitng infrastructure, Support intelligent system	It may requires integration with other technologies, such as PIR, Require each occupant to carry a bluetooth-enabled device	[91]
	Door sensor	Wireless IoT based sensor which detects when a door or window is opened	Non-intrusive, Sensitive and accurate, Low-power	It may requires other sensors for efffective occupancy detection	[92]
	Wi-Fi/Smart device tracking	Use of Wi-Fi signal or GPS for occupancy tracking	Real-time detection of occupancy, Can utilise exisitng infrastructure,	Privacy concerns, Requires a device to be carried by the occupant	[88]
	Wireless sensor network (WSN)	Monitor changes in the environment using sensor nodes	Easy to use, Cross-layer design, Flexibiliy, Low-power	Must be connected through a specific infrastructure or a central device, Security concerns	[93]

Table 2 (*continued*)

Sensor Types	Specific Sensor Type	Features	Advantages	Limitations	Ref.
	Sensor belief network	AI-based network of sensors for determining occupancy information	High accuracy, Missing data entries can be handled successfully	Complex model or algorithm required for processing data	[94]
Camera	Thermal camera	Measure thermal energy for occupancy detection	Less privacy issues compared to standard cameras, Independent of light,	High costs,Resolution , Require line of sight	[96, 97]
	AI-based cameras	Use of AI algorithm integrated with camera devices for detection	Doesn't create disruption High detection level and accuracy, Real-time reponse	Require line of sight, Privacy concerns, Computationally challenging	[95, 98]

healthy spaces must be provided to occupants. Thermal comfort can be defined as a condition of mind that expresses satisfaction with the thermal environment (BS EN ISO 7730). Thermal comfort is traditionally evaluated using the predicted mean vote (PMV) method, which considers environmental and personal factors. It is vital in the design of buildings and HVAC systems to strike a balance between providing adequate thermal comfort and reducing the energy consumed [3]. Like in the previous sections, which showed the emerging developments and adoption of AI methods in energy forecasting and occupancy prediction, recent research has focused on AI methods for predicting and enhancing thermal comfort in buildings.

To address the limitations of the PMV method for thermal comfort assessment in buildings with natural ventilation, Chai et al. [102] employed machine learning algorithms to predict the occupant's thermal comfort and sensation in a naturally ventilated building. The ML algorithm used a combination of indoor and outdoor environmental parameters and personal factors as input. The study highlighted the quick ability of ML to analyse the input and output parameters relationships. They concluded that the ML method performed better than conventional and established models such as PMV. Similarly, Hu et al. [103] used ML techniques to develop a learning-based approach for thermal comfort evaluation. The results showed that all the ML methods achieved better performance than PMV. Specifically, the proposed method outperformed the PMV by up to 17.8%. While Chaudhuri et al. [104] also employed several classification algorithms for developing a thermal comfort prediction model and showed that the ML method outperformed traditional and modified PMV models, achieving prediction accuracy of up to 81.2%.

ML methods can be integrated with control systems to adjust indoor thermal conditions according to the occupant's thermal preference or comfort requirements while enhancing energy efficiency. Peng et al. [105] used ML to develop a framework consisting of multi-learning processes with specified rules for a demand-driven control strategy, which can automatically adapt to occupancy behaviour. The control technique uses the learned occupancy information to operate the cooling system by adjusting the setpoints in real-time. An energy saving of up to 52% was achieved by the proposed control as compared to a conventional method. Yang et al. [106] proposed an optimisation method, which uses model predictive control (MPC) integrated with an ML technique to maintain thermal comfort while consuming the least amount of energy. A reduction of up to 58.5% in cooling energy was achieved in an office compared to conventional controls.

Some of the works combined AI-based thermal comfort prediction and management methods. Such studies use thermal comfort prediction as feedback for HVAC control. Lu et al. [107] used a combination of a thermal comfort prediction model based on machine learning algorithms and a reinforcement learning-based temperature set-point control system to develop a data-driven comfort-based controller for HVAC. They concluded that the ML thermal comfort model outperformed that of PMV. While some studies also looked at optimising other parameters such as indoor air quality. Vallardes et al. [108] developed an HVAC controller based on deep reinforcement learning to reduce energy consumption while maintaining good thermal comfort and air quality in a university building. The results showed that the PMV was maintained within the range of −0.1 to +0.07 while having a 10% lower CO_2 level and reducing the energy by up 4-5% compared to a conventional controller.

The study of Gao et al. [109] also employed deep reinforcement learning to optimise HVAC energy demand and thermal comfort of occupants. The deep neural network method was used to predict the thermal comfort, and the results are then used as input for the controller, which adopts a deep reinforcement learning method. The results showed that the proposed method achieved higher thermal comfort prediction accuracy as compared to other methods such as linear regression and SVM. The study showed the impact on the cooling load of adjusting the thermal comfort threshold and weighting of energy cost, which can be
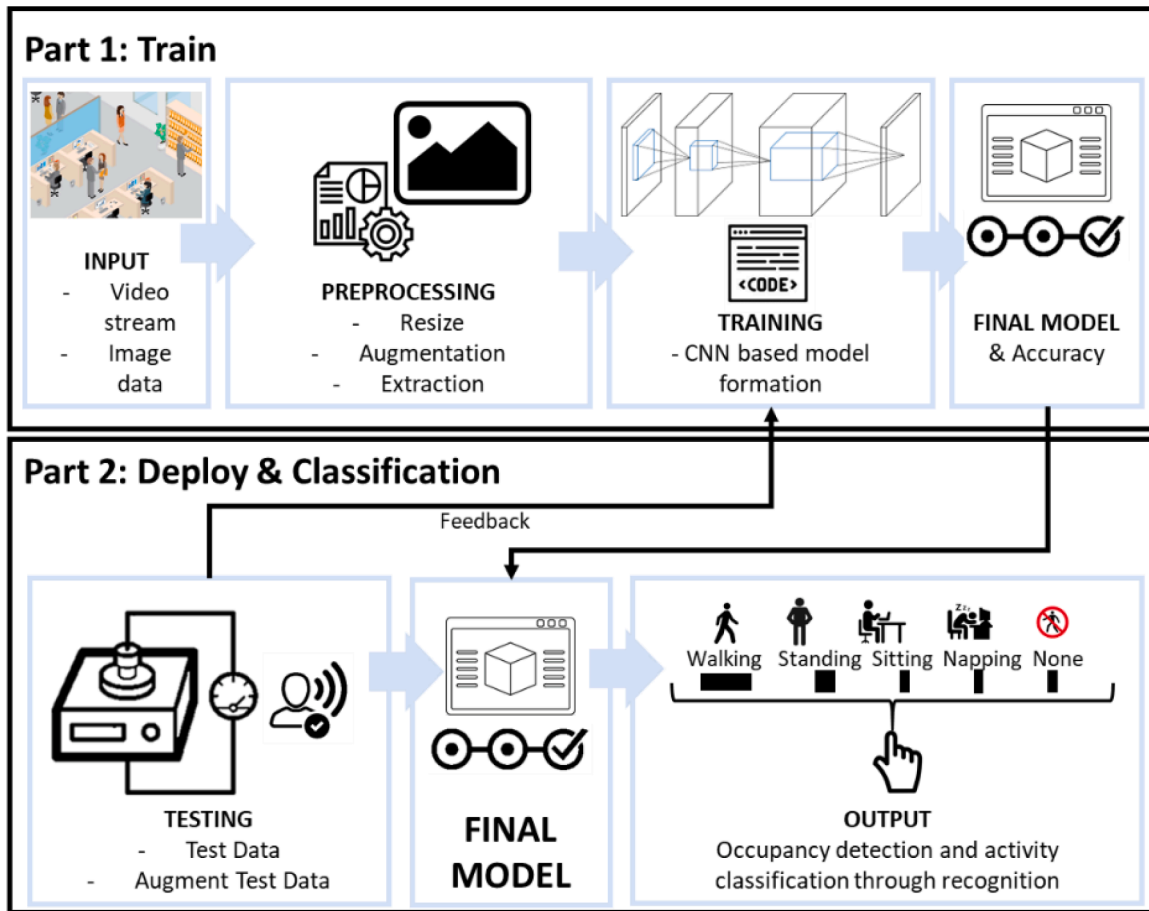
Fig. 5. Example workflow process of the development of AI-based technique used for occupancy activity detection within an indoor environment.
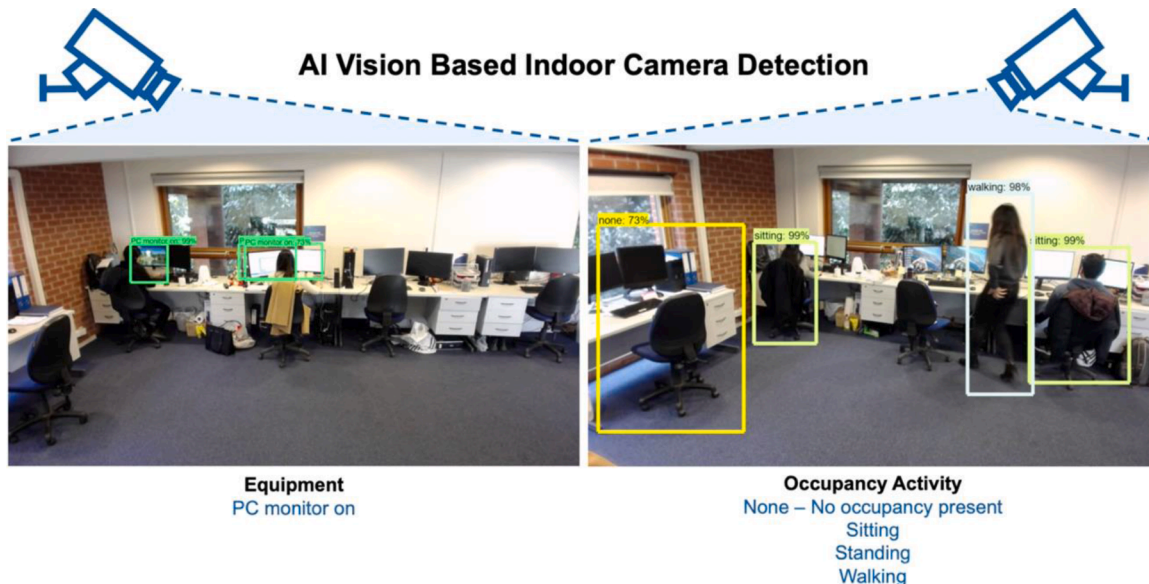
Fig. 6. An example application of an AI vision-based approach for occupancy and equipment detection within an office environment. Refer to **Video 1** to view an example of the detection and recognition.

set depending on the priority.

The applications of ML and DL in thermal comfort studies have been growing, such as in thermal comfort prediction and management [110]. Studies have shown that ML outperformed conventional and modified PMV models. However, studies have also shown the importance of the input parameters and the data size. ML prediction's higher accuracy and speed make it suitable for integration with demand-driven or occupancy-responsive HVAC controls, providing real-time feedback. Several ML and DL methods were used to develop control strategies to ensure a trade-off between energy efficiency and thermal comfort. Based

on the reviewed studies [102–110], a flow chart which summarises the procedure for developing thermal comfort prediction and management models is shown in Fig. 7. Table 3 summarises the AI strategies developed for thermal comfort management and predictions related to the above review.

In addition to ensuring good thermal comfort in buildings, the provision of good indoor air quality (IAQ) is equally important. A good IAQ is essential to ensure the health and well-being of occupants [111]. Like comfort-based systems, various building ventilation systems and control strategies aim to optimise IAQ with the aid of AI techniques such as predicting concentrations of pollutants and managing the indoor environment. Cho and Moon [112] developed an ANN model to predict indoor pollutant concentrations such as carbon dioxide (CO_2), PM_{10} and $PM_{2.5}$. They developed a prediction model sufficiently accurate for integrating control systems in school buildings. The results showed that the model achieved high accuracy with RMSE of 0.8816 for CO_2, 0.4645 for PM_{10}, 0.6646 for $PM_{2.5}$. The study only used simulation results, and field experiments are required to test the approach further. Similarly, Kim et al. [113] predicted the indoor CO_2 concentration for the demand-drive and proactive control of ventilation systems. The study employed machine learning models, including ridge regression, decision tree, random forest, and multilayer perceptron. The study found that the random forest model was the most accurate, and the decision tree was almost as accurate but less computational resource intensive. Hence it is more suitable for lightweight applications.

Several works, such as Vallardes et al. [108] and Yu et al. [114] employed AI algorithms to optimise the operation of HVAC in terms of comfort, air quality and energy. In the study [114], a control algorithm based on deep reinforcement learning was employed to balance the IAQ, thermal comfort and energy demand of air conditioning and exhaust fan

systems. The results showed that the proposed approach achieved an energy saving of up to 43% while reducing the CO_2 level by 24%, as compared to an air conditioning system with a fixed temperature schedule. A demand-controlled ventilation system can benefit from using the two AI-based methods here; accurate pollutant prediction and control optimisation models. This would benefit buildings with irregular occupancy by utilising the forecasted pollutant concentration to control the ventilation to minimise or prevent the rapid increase of CO_2 levels and operating ventilation systems at max capacity.

While the studies covered here are mainly for mechanical systems, AI methods can also be applied in naturally ventilated spaces. For example, camera-based AI techniques can be used to detect occupancy information such as presence, location, activities, and interaction with natural ventilation strategies. For example, Fig. 8 and Video 2 shows the vision-based detector introduced earlier, which can also be used to detect window status in a building. An alert system can be developed to inform/suggest occupants open or close windows depending on the detected (occupancy) and predicted (carbon dioxide level) information.

Video 2 Example of window detection within an indoor environment.

Although not as developed as AI methods for thermal comfort optimisation, the applications of ML and DL in IAQ studies have been recently growing, such as for IAQ prediction and management. This is probably driven by the COVID-19 pandemic and increased awareness of IAQ. Similar to the reviewed thermal comfort prediction and management models, the flowchart in Fig. 7 provides a typical procedure for the development of AI-based IAQ solutions.
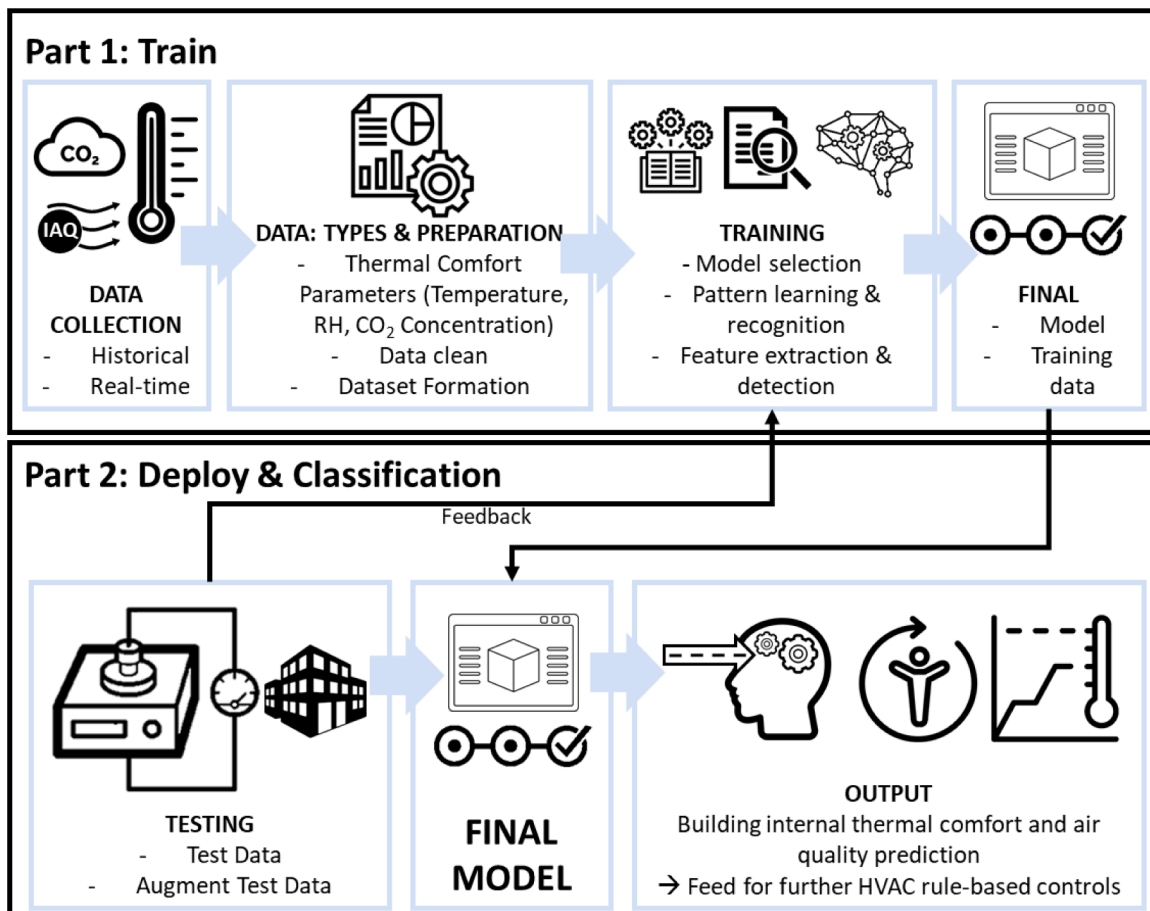

Fig. 7. Example of a framework strategy for building indoor thermal comfort and air quality prediction.

Table 3
Summary of the studies reviewed that used AI for thermal comfort and IAQ prediction and management.

Ref.	Application	Building Type	Technique	System Type	Evaluation Metrics Used	Key Findings	Model Performance /Accuracy
[102] Chai et al. 2020	Thermal comfort prediction	Residential	ML	Natural ventilation	R^2	ML method performed better than conventional PMV.	R^2 0.4872
[103] Hu et al. 2018	Thermal comfort prediction and management	Office	ML	Airconditioning	Prediction accuracy %	ML method outperformed the PMV model	The black-box neural network outperformed PMV by up to 17.8%
[104] Chaudhuri et al. 2017	Thermal comfort prediction	Various	ML	Cooling/Free-running	Prediction accuracy %	ML method outperformed traditionally and modified PMV models	ML achieved a prediction accuracy of up to 81.2%.
[105] Peng et al. 2018	Thermal comfort prediction and management	Office	ML	Cooling	Not specified	Energy-saving of up to 7–52%.	-
[106] Yang et al. 2020	Cooling and thermal comfort management	Office	ML	Cooling and Ventilation	R^2	Energy-saving of up to 58.6% while maintaining thermal comfort level.	Thermal comfort model R^2 0.998
[107] Lu et al. 2017	Thermal comfort prediction and management	Office	ML	Airconditioning	Recall	The recall of the model with the tree algorithms has increased to 6.3%.	ML methods outperformed PMV model up to 6.3%
[108] Vallardes et al.	Thermal comfort management	University	DL	Airconditioning	Not specified	PMV values between −0.1 to +0.07, 4–5% better power usage and 10% lower CO_2	-
[109] Gao et al. 2019	Energy and thermal comfort management	Laboratory building	DL	Cooling	MSE	Deep neural network method outperformed other AI methods.	MSE 1.16
[112] Cho and Moon et al. 2022	IAQ prediction	School building	DL	Electic Heat Pump with Ventilator	RMSE, R^2	Prediction model sufficiently accurate for integration with control system in school buildings.	RMSE 0.8816 for CO_2, 0.4645 for PM_{10}, 0.6646 for $PM_{2.5}$, R^2 1, 0.9991, 0.9979
[113] Kim et al. 2021	IAQ prediction	Not specified	ML	Ventilation	MAE	Decision Tree was found to be almost as accurate as the computationally heavier Random Forest model.	MAE 1.8-14.8 for all the evaluated models
[114] Yu et al. 2021	IAQ prediction and management	University	DL	Airconditioning and exhaust fan	Not specified	Energy saving of up to 43% while reducing the CO2 level by 24%	-

Fig. 8. An example application of an AI vision-basedapproach for window detection within an indoor environment. Refer to **Video 2** to view an example of the detection and recognition.

3. Applications of Machine Learning in the Built Environment

Machine learning is a subset of artificial intelligence. It performs tasks using computer systems that can automatically learn from previous data and improve from experience without following specific instructions [115]. It uses algorithms and statistical models to complete tasks such as modelling, prediction and control [116]. Machine learning has already received much attention in the past decade, but it is expected to continue driving the next big wave of innovations, services and products in many sectors [18]. As established in Section 2, machine learning is one of the most common AI techniques adopted to help solve HVAC and building-related problems. Machine learning can answer the demand of the built environment sector for accurate and quick prediction models, necessary for optimising the design and operation of buildings and energy systems which can lower costs and carbon emissions. At the same time, providing an optimum balance between energy demand, comfort and air quality. This section aims to review the studies that employed machine learning techniques for the built environment.

3.1. Types and Workflow of Machine Learning Techniques

Machine learning consists of three main types of learning; supervised, unsupervised and reinforcement. ML can determine non-linear relationships, such as the relation between the cooling load and related factors such as outdoor temperature and occupancy activity, through mapping functions from a dataset. In supervised learning, a pattern is learned from a labelled dataset (input and output data), and the correct output is expected to be predicted when a different input is entered into the model based on this pattern. Supervised learning algorithms deal with two types of problems: classification and regression problems. Classification algorithms predict a discrete or distinct value, such as when the output is a category, while regression algorithms are used to determine continuous values or quantities. In energy demand forecasting, regression models can be used to understand the factors that drive energy consumption, such as building shape, material and orientation [116]. Fig. 9 presents an example workflow diagram for training and deploying supervised machine learning models, generated based on the reviewed works in Section 2, incorporating all the main steps, including data preparation, model selection and development, and the trained model's application.

On the other hand, unsupervised learning is used to identify patterns in unlabelled datasets and predict the output. Unsupervised learning algorithms are typically utilised for tasks including clustering, association, and dimensionality reduction. Clustering, which can find a structure in a collection of unlabeled data, is the most common algorithm for unsupervised learning techniques [116]. It is the most common unsupervised learning method applied to categorising building performance data [117]. Supervised and unsupervised learning models are different in how they are trained and in terms of the conditions of the required training data. In some cases, such as [118], a strategy which uses a combination of both learning techniques, also called semi-supervised learning, was used to characterise the energy consumption in smart buildings. This type of algorithm learns from datasets simultaneously containing labelled and unlabeled data. For example, a large quantity of unlabeled data is used to enhance the prediction for the labelled data. It can also work independently, performing classification or clustering tasks separately.

Reinforcement learning [119] algorithms learn to react to an environment independently. It has an agent that learns how to map situations to actions, aiming to maximise a numerical reward (gained from a correct output) signal by trial and error. This way, the algorithm improves over time. Over the years, reinforcement learning methods have been applied to the building sector, particularly in HVAC control systems. It has been applied to find the optimal strategies to help decrease building energy demand. For example, a reinforcement learning-based HVAC control will continuously adapt to the controlled indoor environment using real-time data. This presents advantages compared to conventional approaches such as rule-based and model-predictive controls.

Different ML algorithms are suitable for different types of datasets or problems in the field of the built environment and building
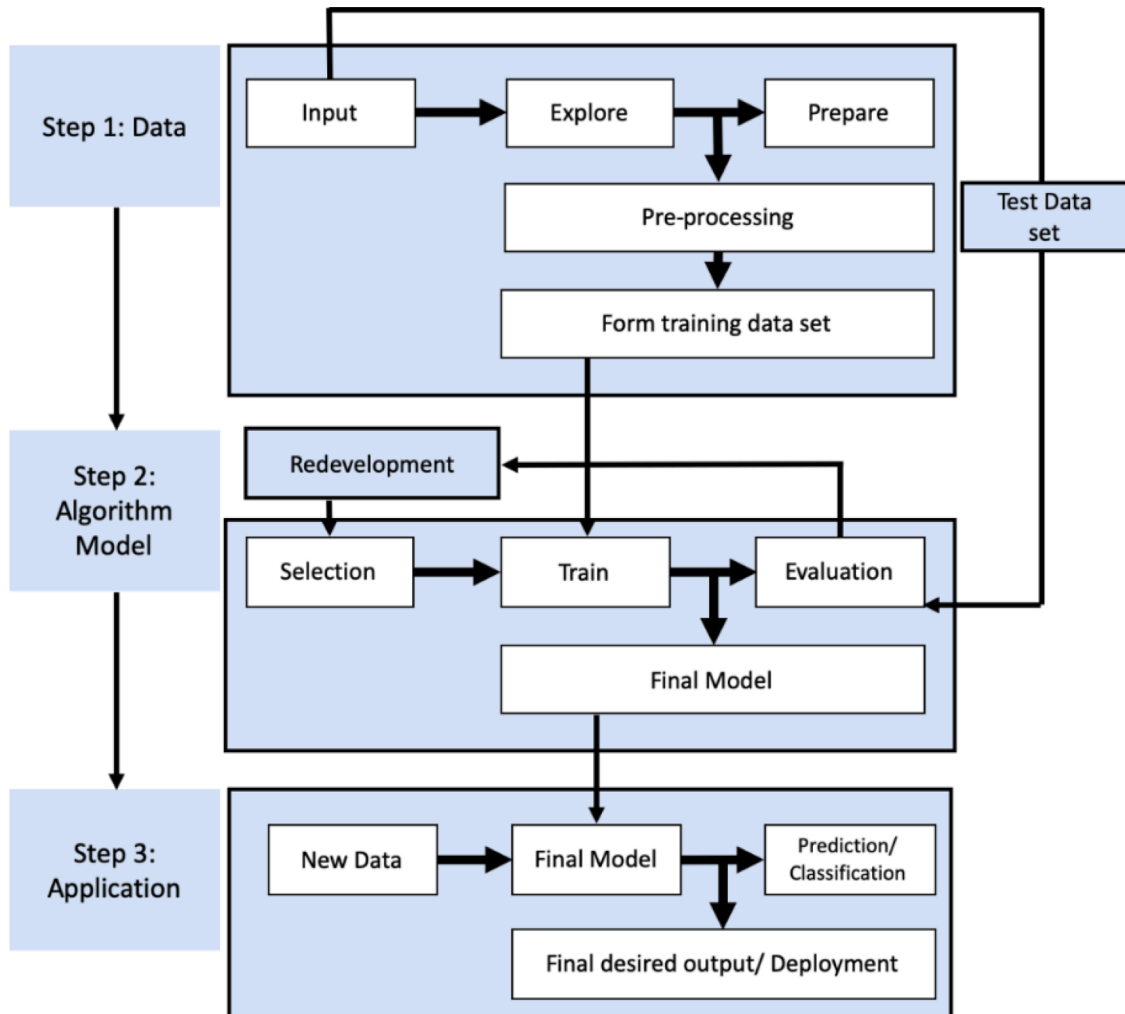

Fig. 9. Example machine learning framework.

performance. Also, different models have different impacts on various problems [120]. Hence it is important for modellers to analyse the available data and application when choosing an ML model and to determine the data pattern that should be learned. In addition, it is important to assess the necessity of implementing ML to solve a specific problem and whether it will be better and more practical than conventional and simpler approaches, considering the modelling effort. The steps required to build, train and deploy an ML model will vary depending on these factors. In general, a typical workflow in built environment studies that use ML models consists of 3 phases; generating data, training the model and deploying the model [121,122].

The data generation involves acquiring input data, which are parameters that impact or correlate with the output data. Methods such as stepwise regression and statistical analysis can be employed to help select useful variables for the prediction. For building performance-related models, inputs could include the climate, building form, occupancy and material properties, while the outputs are parameters which represent building performance. Depending on the prediction time scale required, the sampling period can be from minutes to years. Then the collected data is pre-processed into a format that is suitable for the training process. This involves techniques which improve the quality of the input data. In the training process, appropriate prediction targets and predicting parameters are selected for the model. The input variables, size of the training data, and performance indicators should be considered when selecting the parameters. Once trained, the model is tested to evaluate its prediction performance and determine its suitability for deployment.

As detailed in Tables 1 and 3, various performance indicators can be used for evaluation. This will be further discussed in Section 3.3. The following section will further explore the types of machine learning approaches and models applied to buildings to meet one or more of the objectives detailed in Section 2. Although many studies also applied reinforcement learning, the present study will focus mainly on supervised and unsupervised learning.

3.2. Applications of Machine Learning Techniques in Enhancing Building Performance

As detailed in the previous sub-section, different machine learning models are suitable for several applications related to optimising building performance. The review of the literature showed that it could

be challenging to select the most suitable ML model for solving a problem [123]. The explosive growth experienced by the ML research area in the last 10 years has led to hundreds of ML algorithms being applied to building performance-related studies. This makes it difficult to find an optimal ML algorithm for a specific task or case. Hence, several researchers in the ML field have developed guidelines to encourage best practices in the evaluation and selection of ML models [124,125]. However, further developments and more specific guidelines are required for the built environment field [126].

Many studies have employed different ML algorithms, with many achieving significantly better performance than conventional methods, such as in building energy forecasting [53] and thermal comfort prediction [107]. Supervised learning models are useful for applications such as building energy demand and IAQ forecasting [116]. While unsupervised learning models are helpful for applications such as load profiling, detection and diagnostics of problems occurring in buildings and occupancy detection [117]. In some cases, a combination of both learning techniques, also called semi-supervised learning, is employed [118] to take advantage of the benefits of both methods. It is useful when learning from large datasets containing labelled and unlabeled samples, performing classification and clustering tasks simultaneously. Although advantageous, it would increase the difficulty of the learning process.

Tables 4 and 5 highlight some of the unsupervised and supervised machine learning-based algorithms applied to enhance building performance. The literature review shows that current uses of unsupervised ML in enhancing building performance mainly include, HVAC system management, fault detection and diagnosis, and occupancy detection. The most popular among the domain are clustering algorithms, notably the K-Means algorithm. Clustering methods can provide information on the underlying data structures that is initially hard to detect [126]. Several works have used unsupervised clustering data methods to enhance the performance of their models. In the study [127], unsupervised learning clustering algorithms are integrated with supervised learning algorithms to enhance the prediction of indoor temperature.

The growing interest in the internet of things (IoT)-enabled buildings and a large volume of data generated by IoT sensors presents an opportunity to use unsupervised ML models. For example, training ML models on very large datasets is time and computational resource intensive. One method to overcome this is to reduce the number of training samples such as by employing unsupervised learning algorithms

Table 4
Examples of unsupervised learning models for enhancing building performance.

Ref.	Model Type	Evaluation	Eval. Metric	Application	Key findings
[74] Carreira et al. 2018	K-Means algorithm	Simulation	CV, STDV	Enables automatic configuration of an HVAC system.	The k-means algorithm helped optimise the HVAC system to minimise energy consumption while maintaining user comfort.
[127] Mateo et al. 2013	K-means, Fuzzy c-means, Cumulative Hierarchical Tree, DBSCAN, K-medoids	Simulation	MAE	Improving building indoor temperature prediction of linear and non-linear methods.	Clustering techniques did not show significant improvement for linear or non-linear methods.
[128] Li et al. 2010	Fuzzy C-mean clustering algorithm	Historical data	MAPE, RMSE	Short-term cooling load forecasting	The clustering technique helped reduce the number of training samples and avoid noise samples.
[129] Liu et al. 2021	Two single-step clustering based on K-Means and DBSCAN	Electricity consumption dataset	Visual comparsion, Dunn index	Anomaly detection in building electricity consumption data	The proposed two single-step clustering method outperformed k-means and gaussian mixture model in terms of detecting outliers and discovering typical electricity usage characteristics
[130] Habib et al. 2016	K-Means, Bag of words representation with hierarchical clustering	Building operational data	Cluster evaluation	Detects various operational patterns in a building energy system	The method can automatically find various patterns using as little configuration or field knowledge as possible.
[131] Hong et al. 2015	Gaussian mixture model with partitional clustering	Building sensor information	Classification Accuracy %	Differentiate sensors in buildings by type	The approach is able to achieve more than 92% accuracy for type classification.
[132] Guo et al. 2013	Hidden Markov Model and a clustering algorithm	Experiment	Not specified	HVAC system fault detection and diagnosis	The method not only identified system faults that were modelled within the training process but also can be applied for diagnosis.
[133] Trabelsi et al. 2013	Hidden Markov Model, Expectation - Maximization (EM) algorithm	Experiment	Classification, Precision, Recall	Recognise human activities from wearables	The proposed method achieved a high classification rate of 91.4% and competitive with a well-known supervised approach.

Table 5

Examples of supervised learning models for enhancing building performance.

Ref.	Model Type	Evaluation	Eval. Metric	Application	Key findings
[134] Ciulla and D'Amico 2019	Multiple linear regression	Simulation	R^2, MSE, RMSE, MAE, MAPE	Prediction of annual heating and cooling energy demand	Predicted heating, cooling and comprehensive energy requirements of a building with a high degree of reliability. R^2 of 0.9 and the Mean Absolute Error and Root Mean Square error are lower than 10 kWh/m^2 year.
[135] Bilous et al. 2018	Multivariate regression model	Simulation	R^2, Fisher's criterion	Prediction of hourly internal air temperature	The model provided high accuracy predictions with R^2 of 0.981.
[57] Chou and Bui. 2014	Support vector regression (SVR), Ensemble, General linear regression, Classification and regression tree, ANN	Simulation	R^2, MSE, RMSE, MAE, MAPE	Heating and cooling load predictions	The ensemble approach (SVR +ANN) and SVR were the best models for predicting heating and cooling load.
[136] Zhang et al. 2018	Nonlinear ML algorithms, SVR with nonlinear radial basis function (RBF) kernel and neural networks	Experiment	MAE, R^2, Time	Different ML techniques for modelling thermal comfort levels	The nonlinear models performed significantly better than the linear models. The neural network had the best performance.
[137] Wang et al. 2018	Combined ANN with an ensemble approach	Simulation	R^2	A novel dynamic forecasting model for building cooling loads that combine ANN with an ensemble approach	The proposed ensemble model greatly improved the forecasting accuracy.
[138] Wu et al. 2018	Ensemble method	Experiment	R^2, RMSE, MAE, r	An intelligent ensemble ELM method was developed for thermal perception prediction	The proposed ensemble model performed better than ANN and SVM.
[139] Ryu et al. 2016	Decision Tree	Experiment	RMSE	Indoor environmental data-driven model for occupancy prediction	The decision tree model was capable of estimating the occupancy state. Depending on the number of predictors used, the RMSE ranged between 0.3673 and 0.2202.
[140] Wang et al. 2018	K-nearest neighbour (KNN), Support vector machine (SVM), Artificial neural network (ANN)	Experiment	RMSE, MAE, MAPE	Occupancy prediction through ML, environmental and Wi-Fi sensors	The use of multiple sensing data sources significantly improved the reliability and accuracy of ANN-based models while it enhanced the reliability of SVM and KNN.
[141] Johannesen et al. 2019	Random Forest Regressor, k-Nearest Neighbour Regressor and Linear Regressor	Energy load dataset	MAPE	Urban area electrical energy demand forecasting	Random Forest Regressor provides better short-term load prediction, and kNN offers relatively better long-term load prediction.
[142] Song et al. 2017	K-means for building energy prediction. ANN for end-user group prediction.	Energy and occupancy dataset	CV-RMSE	Hourly energy prediction considering occupancy characteristics	The prediction accuracy is improved when considering diverse occupancy and its correlation with energy use.
[143] Peng et al. 2017	K-nearest neighbour	Experiment, Application	Not specified	Occupancy information prediction employed for the control of the cooling system	Average control accuracy of 88.1%. Energy saving of up to 20.3% is achieved with the use of demand-driven control.
[144] Xiong and Yao, 2021	K-nearest neighbour	Experiment	Accuracy %	Thermal comfort model to establish a personalised adaptive thermal comfort environment.	The KNN-based thermal comfort model with 1000 sets of training data can have an accuracy of 88.31%.
[145] Qiong et al. 2017	Support vector machine	Energy load dataset	RMSE, MRE	Prediction model of annual energy consumption of residential buildings	The support vector machine achieved better accuracy and generalisation than evaluated neural network techniques.
[146] Liu et al. 2021	Random Forest	Simulation	R^2, RMSE	Prediction model of energy consumption of University buildings	The Random Forest model exhibits notable advantages in building energy consumption prediction compared to SVM

to pretreat the training sample sets. This could reduce the number of training samples and also avoid noise samples [128]. Another application of ML is AI-based anomaly detection, which can help in making better decisions to reduce energy use wastage and promote energy-efficient behaviour in buildings. The study [129] proposed a two-step clustering method composed of DBSCAN and k-means algorithm for a framework which identifies daily electricity usage patterns and detects anomalies in building electricity consumption data. For analysing actual building operational data, the use of unsupervised learning is more practical since anomaly labels are typically not available.

Some studies employed unsupervised learning methods to analyse and optimise the operation of building technologies and systems. The work [130] used a combination of K-means and hierarchical clustering to detect various operational patterns in a solar adsorption chiller. The proposed method was capable of automatically finding various patterns using as little configuration or field knowledge as possible. Several studies [132] also applied unsupervised ML in the development of the fault detection and diagnosis (FDD) approach for HVAC systems in

buildings. While many studies have highlighted the good performance of supervised learning techniques-based FDD methods for HVAC, they rely on a balanced dataset containing normal and fault data. In practice, obtaining a sufficient amount of fault dataset could be challenging; hence, many researchers are exploring unsupervised learning techniques.

Similarly, for the detection and recognition of occupants in a building, obtaining sufficient actual (ground truth) information can sometimes be challenging, and hence several researchers have explored unsupervised approaches. For example, the work [133] used an unsupervised learning approach for occupancy activity recognition from wearable sensor data. The results showed that the proposed method achieved a higher classification rate of 91.4% as compared to well-known unsupervised classification approaches and was competitive with well-known supervised approaches. In Section 4, we will dive deeper into other approaches, such as deep learning and computer vision methods for occupancy detection.

In the last decade, many works have used supervised learning techniques to conduct various types of building energy use forecasting,

including heating and cooling load, thermal comfort prediction and occupancy prediction. Table 5 lists typical supervised learning studies in the literature. Most of the studies focus on heating and cooling energy demand, which accounts for a significant portion of the total energy use in buildings and, at the same time, impacts the IEQ. Focusing solely on energy prediction studies, it was observed that existing studies have a wide range of prediction time scales from minute to year basis. The study [134] employed a multiple linear regression model to conduct an immediate assessment of annual (long-term) heating and cooling building energy requirements, which can be used as a decision support tool for the preliminary evaluation of a building. While it predicted the energy requirements of a building with a high degree of reliability, however, they did point out that it is not intended to replace a dynamic simulation model. While some of the studies showed the capabilities of supervised learning techniques in short-term predictions. The study [135] used a multivariate regression model to predict the hourly indoor air temperature. They used a multivariate regression model which considered various external environmental factors affecting the thermal performance of the building. The results indicated that the model provided high accuracy predictions with $R^2 = 0.981$.

Another important application of supervised ML is thermal comfort prediction. The study used supervised learning methods including support vector regression (SVM), ensemble, general linear regression, classification and regression tree to predict indoor predictive mean vote (PMV) factors. Unlike previous works, the study investigated the relationship between controllable HVAC operation parameters and indoor PMV factors. The study showed that non-linear models such as support vector regression with non-linear RBF kernels and neural networks performed better than linear methods.

Several works used the ensemble learning approach to improve the performance of prediction models. It combines the predictions from multiple algorithms to achieve better prediction accuracy. The study [137] proposed a one-step-ahead forecasting model based on an ensemble technique for cooling loads, which can help tackle the time-lag issues of HVAC control. The results showed that the proposed ensemble model greatly improved the forecasting accuracy. While the work [138] also used an ensemble method for thermal perception prediction. The study concluded that it was more accurate than ANN and SVM in the prediction of thermal perception and outperformed the traditional PMV in terms of thermal sensation estimation.

Occupancy detection and prediction is a popular application of supervised ML. Such methods estimate occupancy using data from building sensors, energy meters, Bluetooth and WiFi signal. The study [139] used a decision tree model to detect the occupancy at the current state based on energy consumption and environmental data. Based on the result, the decision tree model was capable of estimating the occupancy at the current state. Depending on the number of predictors used, the RMSE ranged between 0.3673 and 0.2202. While Wang et al. [140] used environmental data, WiFi and fused data combined with ML to develop an occupancy prediction. Examined with an on-site experiment, the results suggest that the ANN-based model with fused data has the best performance, while the SVM model is more suitable with Wi-Fi data.

The literature showed that existing studies have a wide range of research object scope or scale, from energy prediction for a cooling system within an individual building to energy prediction for an urban area. For example, the study [141] evaluated supervised learning regression tools for urban area electrical energy demand forecasting. The proposed method used less data as compared to continuous time-series and neural network techniques. The impact of meteorological parameters was considered. The results showed that the Random Forest Regressor provided better short-term load prediction while the K-nearest neighbour offered relatively better long-term load prediction.

As observed, each ML algorithm has its advantages and limitations, and it is imperative that several factors such as the task, data availability, practicability and computational cost are considered when selecting a suitable model for a project [147,148]. With regards to building performance-related projects, although knowledge of artificial intelligence/machine learning is key, it is also paramount for model developers to have knowledge in building engineering/energy systems [126]. This is important when selecting an appropriate prediction target and predicting parameters for the ML model.

This section highlighted examples of unsupervised and supervised machine learning-based algorithms applied in building and energy efficiency research. Most of the ML methods have been applied to forecast energy consumption, thermal comfort and occupancy information, and fault/anomaly detection. It can be observed that the majority of the reviewed studies were focused on enhancing energy efficiency and thermal comfort, while there are limited studies on air quality. Furthermore, there are limited works which aim to optimise all 3 variables [108,114]. The literature analysis showed that existing research studies considered a wide range of scope/scales (from an HVAC component to urban areas) and time scale (minute to year).

It can be observed that many of the studies used several unsupervised and supervised ML techniques comparing the capabilities of each algorithm for different applications. Many studies combined several ML methods, with some using a combination of supervised and unsupervised algorithms (also called semi-supervised) to take advantage of the benefits of both approaches. This can be seen in the reviewed works which developed ML methods with multiple stages. For example, the first stage uses an ML method to detect current-state occupancy, then the second stage feeds this information to another ML method to carry out another task, such as controlling the HVAC or forecasting occupancy in a future state.

Some of the works used multiple algorithms to achieve better performance, such as the ensemble learning approach. The literature also showed examples of unsupervised and supervised ML performing well in carrying out a similar task. For example, occupancy detection and prediction is a popular application of unsupervised and supervised ML. While studies have shown supervised learning ML performing well in occupancy detection and prediction tasks, obtaining sufficient actual information can sometimes be challenging, and hence several researchers have explored unsupervised approaches.

While the literature showed that ML had been successfully applied in building energy efficiency research, most of the studies are still at the experimental or testing stage, and there are limited studies which implemented ML strategies in actual buildings and conducted the post-occupancy evaluation. Finally, most of the studies are focused on individual buildings or few building spaces, while there are limited large-scale applications and implementation in other types of buildings such as industrial and retail buildings. This is probably due to the constraints and challenges of data acquisition.

3.3. Evaluation of Machine Learning Algorithms for Building Performance-Related Applications

As shown in Fig. 9, the next stage following the training of the ML model is the evaluation or experimental stage. At the experimental stage, the performance of the selected ML method is evaluated based on how well it performs on new or unseen data. Typically the development of ML methods involves several experimentations, including testing of several types of algorithms and optimisation or tuning of the hyperparameters [124,125]. As detailed in Tables 1, 3, 4 and 5, various performance metrics or indicators related to model accuracy, sensitivity and robustness are used for model evaluation. These metrics are used to evaluate a model's performance and provide feedback to allow improvements until a desirable performance is achieved [149]. Based on the reviewed literature, the typical evaluation metrics employed [149] include accuracy, mean squared error (MSE), root mean square error (RMSE), mean absolute error (MAE), R squared (R^2), and mean absolute percentage error (MAPE) [57]. The evaluations are carried out using data from simulations, historical data and experiments.

Other performance indicators such as computational efficiency, data

requirements and complexity are also equally important but have not been assessed or discussed in many of the reviewed studies. This information would be useful for researchers and model developers that have certain requirements for their projects, such as high-speed prediction, low computational cost and complex datasets. For example, ML methods such as the ensemble approach could achieve higher prediction performance, but their complexity and higher computational resource requirements may hinder their applications in practice. In addition, the challenges faced in practice related to the pre-processing and uncertainty of data should be discussed in future works. More focus should be given to these performance indicators and challenges in future studies to ease the selection of ML models for different applications.

It should be noted that the information presented about the evaluation metrics in Tables 1, 3, 4 and 5 are not meant to assess which is the best evaluation metric, but rather to show that existing studies are using different evaluation methods. This inconsistency makes it challenging to assess or compare the different ML algorithms and select an optimal method. While most of the studies covered in the review showed that the proposed ML methods are suitable and achieved good performance for various building-related applications, it also adds to the challenge of finding an optimal ML algorithm. A common approach practised in the literature is the testing of different models on the same data to evaluate the best performing method. Hence, guidelines on the standardisation of the evaluation methods are required to promote the optimal selection of ML algorithms.

At the application stage, once the model is trained and validated, the ML method is then applied to an actual building, for example, its energy management or HVAC control system [143], or used to support the decision-making during the building design process. While this stage is important for further validating the performance of the ML method in real implementations, previous works have not paid much attention to this. The review of [126] highlighted that many of the building performance-related ML studies are at the experimental stage. This can be observed in Tables 4 and 5, with most of the studies are either evaluated using simulation results, previously collected datasets, and experiments. Many thermal comfort and occupancy detection studies used experiments to validate the model, while energy prediction studies used historical datasets and simulation results. The use of commercial BES tools such as EnergyPlus is a common compromise to generate training data and prove the theoretical feasibility of the ML model, which may not be sufficient to put the ML model into practice.

Furthermore, they [126] also highlighted that most studies did not validate their ML models using real post-building-completion energy or post-occupancy data. Among the reviewed studies, the work of [143] implemented the proposed demand-driven strategy for the controls of a cooling system of several offices in a commercial building in Singapore. The two months field experiment showed that the proposed strategy could result in energy savings of up to 20.3%. While this study showed the capabilities of the K-nearest neighbour model in occupancy prediction strategy in real offices, the authors did acknowledge the limitations of the experiment. The limited experimental time and a small number of case study offices may restrain the universality of their results.

While modellers aim to develop ML models with good performance, it does not necessarily mean that it will lead to successful real-world implementations. Future works should consider implementing their proposed ML-based strategies in real buildings and systems. More studies should be carried out on the integration of existing ML strategies with building management systems and control strategies.

4. Applications of Deep Learning in the Built Environment

While deep learning (DL) is a subset of machine learning, we have dedicated a section covering some of the latest advancements and applications of deep learning techniques in the built environment, in particular, energy prediction and occupancy detection. As defined by [150], DL allows computational models comprised of multiple processing layers to learn data representations with multiple levels of abstraction, imitating the human brain. While it can be considered an ML concept based on artificial neural networks (ANN), DL typically contain advanced neurons, which allows deep neural networks to be fed with data in raw form and automatically discover a representation required for the corresponding learning tasks [115]. Such functionalities are not provided by simple ANNs and ML algorithms.

Hence, DL can address some of the limitations of conventional ML methods, including ML's inability to process natural data in their raw form, as discussed in the earlier sections. It eliminates some of the data pre-processing tasks typically required in developing ML methods, reducing the reliance on expert knowledge (see Fig. 1b). Besides numerical and text data forms, DL can also process other forms of data such as images, videos, and sounds. Like ML, DL can be supervised, unsupervised or reinforcement, depending on how the neural network is utilised. For instance, when both input and output are known, such as when performing classification or object detection tasks, supervised learning is used to carry out the prediction task. While unsupervised learning can be used to cluster images based on their similarities.

While ML has shown good performance in various applications such as energy forecasting, thermal comfort prediction and occupancy detection. The literature has shown that these applications are increasingly making use of deep learning techniques [151,152]. According to [115], DL outperforms conventional shallow ML algorithms and traditional data analysis methods in numerous applications that require text, images, audio, and video to be processed, such as in natural language processing, image classification, speech recognition and computer vision [150]. It is particularly useful in applications which require handling large and high-dimensional data [153]. The increasing amount of data generated from many sources, the advances in computational power, and the developments of DL algorithms have led to the increasing popularity of DL. However, it also means that DL would require high-end machines and could take a long time to train a model. While in applications which deal with low dimensional and small datasets, conventional ML is still preferable as it could produce superior and more interpretable results [154].

From self-driving cars to medical imaging solutions to Google Translate, DL applications have revolutionised many sectors and will continue to do so. While this disruptive technology is becoming more common across a wide range of industries, researchers and practitioners in the built environment sector are striving to keep up with the pace of applying DL. Table 6 highlights some deep learning-based algorithms applied to solve building performance-related problems. The literature review shows that the current uses of DL are in enhancing building performance, mainly including energy demand and thermal comfort prediction, and occupancy detection.

While DL algorithms can be applied to a range of prediction or classification tasks, like ML, modelers should consider several factors when selecting a DL model for a project, such as the type of tasks, learning method, data availability, practicability and computational cost. The literature showed that DL techniques had been successfully applied in various types of building energy use forecasting. For example, recurrent neural network (RNN) models have shown favourable performance in electric load forecasting [151]. RNN is one of the neural network architectures best suited to tackle datasets with sequential correlations and has demonstrated good performance in processing time-series data such as text and speech. The time series nature of operational data of buildings makes RNN a suitable technique for energy consumption prediction tasks. However, traditional RNN models are not effective in capturing long-term temporal dependencies. Models such as long short-term memory (LSTM) and gated recurrent units (GRU) are used to overcome some of the shortcomings of RNN models. Hence the work [42] proposed several techniques to enhance the performance of recurrent models in building cooling energy predictions.

The study [152] compared the performance of RNN and convolution neural network (CNN)-based algorithms for energy forecasting in

Table 6
Examples of deep learning-based methods for enhancing building performance.

Ref.	Model Type	Evaluation	Eval. Metric	Application	Key findings
[151] Kim et al. 2019	Recurrent neural network (RNN), recurrent inception convolution neural network (RICNN)	Smart meter dataset	RMSE, MAPE	Multi short-term load forecasting model	RICNN model outperforms the benchmarked multi-layer perception, RNN, and 1-D CNN
[42] Fan et al. 2019	Recurrent neural network (RNN)	Building operational data	RMSE, MAE, CV-RMSE	Short-term building energy predictions	Recurrent models achieved the most accurate predictions without increasing computational load.RMSE was 30.8% for cooling load prediction
[152] Cai et al. 2019	Gated RNN, CNN	Electricity dataset	CV, MAPE, Computational efficiency	Day-ahead building-level load forecasts	24-h gated CNN model performed better than conventional deep learning models.
[153] Kim and Cho et al. 2019	CNN-Long short-term memory (LSTM)	Electricity dataset	RMSE, MAPE, MAE, MSE	Residential energy consumption prediction	CNN-LSTM network estimated the electric energy consumption in real residential houses with a stable performance of 0.37 MSE
[155] Somu et al. 2019	CNN-LSTM	Energy consumption data	RMSE, MAPE, MAE, MSE	Building energy consumption forecast	CNN-LSTM captured the spatio-temporal characteristics in building energy consumption data
[156] Suryanarayana et al. 2018	Deep neural networks	Heat load dataset	MAPE, F_1 score, F_2	Thermal load forecasting in district heating	Deep neural networks performed better than all linear models although more computationally intensive.
[157] Marino et al. 2016	Long short-term memory (LSTM)	Electricity dataset	RMSE	Building energy load forecasting	LSTM performed well for one-hour resolution dataset but not for one-minute resolution data.
[158] Somu et al. 2021	CNN-LSTM	Thermal comfort dataset	Precision, Accuracy, F_1 score, MCC	Thermal comfort modelling with limited dataset	The model proposed provided reasonably accurate predictions and overcome the challenges related to the inadequacy of modeling data.
[159] Gao et al. 2021	Transfer learning based multilayer perceptron (TL-MLP)	Thermal comfort dataset	Accuracy, F_1 score	Thermal comfort prediction in multiple cities	The Tl-MLP model from the same climate zone exceeded the performance of state-of-the-art methods in accuracy and F1-score.
[160] Deng and Chen 2018	Artificial neural network (ANN)	Thermal sensation data, behaviour data	MAE, R^2	Thermal comfort prediction in single-occupant and multi-occupant offices	Comfort zone obtained by the ANN model using thermal sensations was narrower than the comfort zone in ASHRAE Standard 55, while the ANN model using behaviors was wider than the ASHRAE comfort zone.
[161] Wei et al. 20	Convolutional neural network (CNN)	Experiment, Simulation	Accuracy %	Equipment usage detection for estimating sensible heat gains	The CNN-based model detected three typical office appliances – PC, printer and kettle. The initially developed model achieved an accuracy of 89.3% on the test dataset.
[163] Tien et al. 2020	Faster region-based convolutional neural network (RCNN)	Experiment, Video, Simulation	F1 score, Precision, Accuracy, Recall	Occupancy activity detection for estimating sensible heat gains	Using the deep learning detection method, the results showed that the occupancy heat gains could be represented more accurately in comparison to using static office occupancy profiles.
[164] Tien et al. 2021	Faster region-based convolutional neural network (RCNN)	Experiment, Video, Simulation	F1 score, Precision, Accuracy, Recall	Window opening detection for estimating ventilation heat loss	The initial detection results showed that the method was capable of identifying windows that were opened and had an average accuracy of 97.29%.
[165] Wei et al. 2022	Faster region-based convolutional neural network (RCNN)	Experiment, Video, Simulation	IoU, F1 score, Precision, Accuracy, Recall	Occupancy detection for demand-comtrolled ventilation	The proposed approach could provide demand-driven ventilation control data on the dynamic changes of occupancy to improve indoor air quality.
[166] Pincott et al. 2022	Faster region-based convolutional neural network (RCNN)	Video	F1 score, Precision, Accuracy, Recall	Fire and smoke detection in buildings	Results indicated that the fire detection achieved accurate detections while the smoke detection did not perform as well.
[167] Anh et al. 2017	Long short-term memory (LSTM), Gated recurrent units (GRU)	Air quality dataset	Accuracy	Indoor air quality prediction using sensors and DL	GRU model outperformed the LSTM model. The GRU model achieved an accuracy of up to 84.5%
[168] Mutis et al. 2020	Deep neural network	Human action dataset	Accuracy	Indoor air quality control using occupancy information	The model achieved 84% accuracy on human action dataset using a multi-stream fusion network for recognising activities.
[169] Taheri et al. 2022	Deep recurrent neural networks (DRNN)	HVAC dataset	F1 score, Precision, Accuracy, Recall	Fault detection and diagnostic for variable flow refrigerant system	The DRNN model outperformed random forest and gradient boosting regression.
[170] Guo et al. 2018	Deep belief network	HVAC dataset	Correct rate (CR), Hit rate (HR)	Fault detection and diagnostic for variable flow refrigerant system	The fault diagnosis correct rate of the optimized model was 97.7%,
[171] Fan et al. 2018	Autoencoder	Building operational data	Accuracy %	Anomaly detection in building energy data	The proposed autoencoder method successfully identified anomalies in building energy data.

commercial buildings. The study developed the gated RNN and CNN specifically for day-ahead building-level load forecasting. The results showed that the gated CNN model outperformed the gated RNN in terms of accuracy and computational efficiency. While CNN is well known for object/image recognition and classification due to their optimal performance, it has also been successfully used in various applications such as building energy prediction tasks which involve time series data [152].

Although most deep learning architectures are applicable to a range of prediction or classification tasks, some studies have combined deep

learning techniques for better performance. In the study [151], a recurrent neural network (RNN) model was combined with 1-dimensional CNN to enhance the performance of a forecast model for short-term loads. The study showed that the proposed model outperformed other models, including multi-layer perception and RNN. In another study [153], CNN is combined with LSTM to predict residential energy consumption. The CNN layer extracts the features between multivariate variables affecting energy consumption, while the LSTM models the temporal information and maps time series into separable spaces to generate predictions. The proposed hybrid model accurately predicted the electric energy consumption of residential houses and performed better than other DL models such as LSTM and GRU.

In some cases, traditional ML techniques are combined with DL techniques. For example, the study [155] combined CNN and LSTM with $k-$means clustering for building energy consumption forecasts. The unsupervised learning clustering method was used to understand the energy consumption trend before data modelling. The results showed the efficiency of the proposed model over the existing building energy consumption forecast models, such as MLP, CNN and LSTM. Furthermore, it showed the capability of the combined CNN and LSTM in capturing the spatio-temporal characteristics in building energy consumption data.

While many of the works in the literature predicted electricity consumption of buildings, the study [156] extended deep learning methods to the field of heat load forecasting. They compared the performance of the DL architecture, deep neural network with linear models in forecasting thermal loads in district heating networks. A deep neural network is a standard neural network with multiple hidden layers (at least 2 hidden layers) between the input and output layers. It extracts uniquely abstract features to model complex non-linear relationships. The results showed that although more computationally intensive, the deep neural network provided the best accuracy among the tested techniques.

Another important application of DL is thermal comfort prediction [158,159] and management [108,109]. As mentioned earlier, data-driven techniques have shown their advantage over traditional PMV models; however its application can be hindered by unavailability or lack of labelled thermal comfort data from occupants. The work [158] addresses this issue by implementing a transfer learning based CNN-LSTM model for predicting thermal comfort in buildings with limited data. Transfer learning takes relevant parts of a pre-trained model and applies them to a new but similar problem. This allows the training of a model using a smaller dataset while using a large amount of relevant data from a previous task. The hybrid model achieved reasonably accurate predictions and overcame the challenges related to the inadequacy of modelling data. It performed better than other methods such as LSTM, CNN, SVVM, kNN and traditional PMV.

Another study [159] proposed a transfer learning-based multilayer perceptron (MLP) model for thermal comfort predictions for any building in a similar climate with limited labelled data. Transfer learning allowed the transfer of knowledge from a building in a similar thermal environment or climate to another building for thermal comfort prediction. The proposed model exceeded the performance of state-of-the-art methods in terms of accuracy and F1 score. While transfer learning has been successfully utilised in many real-world applications, it can be a promising technique to scale up the implementation of ML and DL models in the built environment sector.

As mentioned previously, CNN is one of the most popular DL techniques and has shown optimal performance when it comes to object and image recognition tasks. Its advancement has benefitted many areas, including computer vision which is typically used in applications such as image detection and recognition, image and video analysis and NLP. While computer vision is not new, it has advanced significantly in the last 10 years since the development of the AlexNet model, which uses CNN. Furthermore, the availability and increase in GPU power have accelerated the development of computer vision solutions. Today, CNN

is used in numerous computer vision applications such as facial recognition, augmented reality, pedestrian detection and autonomous vehicles. The advancement in computer vision has gained the interest of many researchers in several fields, including the built environment. Computer vision-based methods have shown good performance in occupancy detection and recognition tasks in buildings [161,162]. Fig. 10 presents an example workflow of a CNN model applied to an occupancy activity detection problem.

The study of [163] employed a computer vision-based approach and camera for the real-time detection of occupancy and equipment use in buildings. Instead of using typical "fixed" scheduled profiles for HVAC systems, the work used the proposed detection method to collect information on the real-time usage of building spaces and use the information to automatically adjust the operation of HVAC systems (see Fig. 11 and Video 3). Through building energy simulations (BES), the work showed that the proposed approach could help minimise unnecessary building energy demand. The same approach was used to detect and recognise the opening/closing of windows in [164]. The real-time information can be used to alert building operators or users and automatically adjust HVAC to minimise energy wastage and maintain thermal comfort.

In addition, the proposed strategy could be used to enhance indoor air quality and building safety by detecting indoor CO_2 levels [165], and indoor fire and smoke [166]. Using the same occupancy detection method [163], the work [165] used the real-time predicted CO_2 information to assist HVAC operations to provide demand-driven ventilation controls. This allows the HVAC to adjust according to the dynamic changes of occupancy to improve the IAQ and minimise the under- or over-estimation of the ventilation demand [167]. The study [166] applied the method in the detection of indoor fire and smoke, which can potentially provide necessary information for fire services, including identifying the position and size of the fire and how the fire spreads. The authors envisioned that a single vision-based device could provide numerous information about a building space, that can be used to achieve comfortable, healthy and safe indoor environments. While these studies showed the advantages of the vision-based method, they also highlighted some of the issues that should be further addressed, including the security issues, overlapping objects and the impact of parameters on detection performance such as indoor lighting conditions and camera angle.

While the prediction of air quality in oudoor/large scale applications such as urban areas and cities using DL is quite well established, the use of this technology for predicting indoor air quality (IAQ) is only starting to pick up steam. For example, the study [167] employed LSTM and GRU models to handle the time series data of measurements from various sensors and predict indoor air quality. The sensors measured several parameters, including temperature, humidity, light quantity, carbon dioxide, fine dust and volatile organic compounds. The results showed that the GRU technique outperformed (up to 85% accuracy) the other two methods; LSTM and linear regression. Another approach was taken by [165], which employed computer vision to detect real-time occupancy levels, and used the information to estimate the indoor CO_2 level. Similarly, the work [168] employed DL methods to sense occupancy and recognise motion in real-time for indoor CO2 level estimation. The model achieved 84% accuracy on the human action dataset using a multi-stream fusion network for recognising activities. Such techniques will be invaluable for the development of demand-driven HVAC control systems.

Another emerging application of DL is in fault detection and diagnosis of HVAC. While ML has been extensively applied to HVAC fault detection and diagnosis, some techniques barely capture temporal dependencies and dynamic behaviour of the faults [169]. Hence, DL methods such as deep recurrent neural networks are developed to address these shortcomings. The work [169] explored different configurations of deep recurrent neural networks (DRNN) are explored for the fault diagnosis of variable flow refrigerant systems. Through
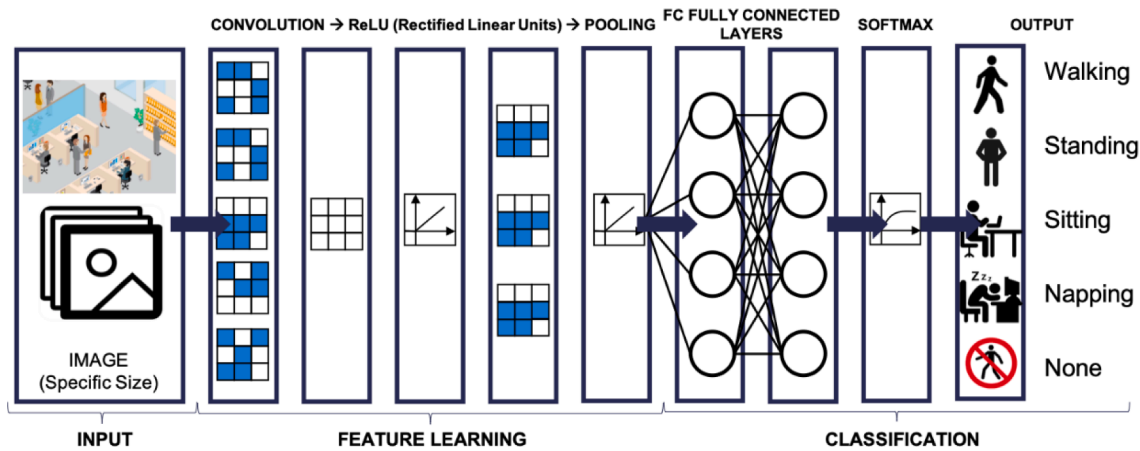
Fig. 10. Example CNN architecture based on deep neural networks [62], applied to an occupancy activity detection problem.
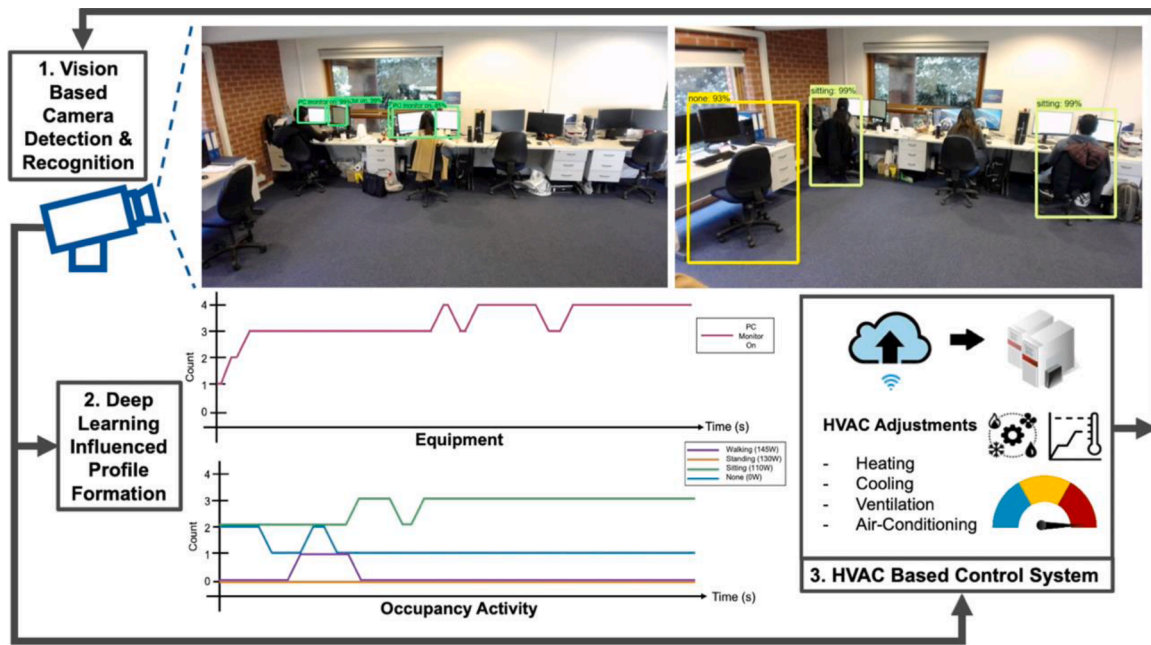

Fig. 11. Example of a computer vision based strategy for the detection of occupancy and equipment use [162,163] for data-driven building energy management systems. Video 3 shows the operation of the proposed approach.

optimisation and evaluation of hyperparameters, the best DRNN model was selected and compared with other established techniques. The results showed that the DRNN model outperformed random forest and gradient boosting regression. Similarly, [170] also used a DL method based on a deep belief network for the fault diagnosis of a variable flow refrigerant system. The study highlighted that the fault diagnosis for variable flow refrigerant systems might not require a very deep model.

This section highlighted examples of deep learning techniques applied in building and energy efficiency research. Most of the DL methods have been applied to short-term/ long-term energy prediction, thermal comfort prediction and occupancy detection. Like in ML, the literature analysis showed that most of the studies were focused on optimising the energy efficiency and thermal comfort in building while less attention was paid to air quality. In addition, although energy use, thermal comfort, and indoor air quality are interrelated, there are limited studies which investigated/considered them all simultaneously.

DL algorithms have recently become the focus of increased attention due to their performance and capabilities. Its popularity has been fuelled by the increasing amount of data generated from many sources, advances in computational power, and the developments of algorithms. DL methods have been promising for the development of building energy prediction models due to their powerful learning and prediction abilities. Many of the reviewed studies showed DL methods outperforming conventional shallow ML algorithms and traditional data analysis methods. Although most deep learning architectures are applicable to a range of prediction or classification tasks, some studies have combined DL techniques or DL with ML techniques for better performance. Such an integrated method typically works in stages, with each stage taking advantage of the abilities of the DL/ML method. For example, one method will carry out clustering tasks, and the other will perform prediction tasks. The literature showed that while ML has been extensively applied to thermal comfort, indoor air quality prediction, and fault detection and diagnosis, the applications of DL in these areas are only starting to pick up steam.

One of the most important developments in the field of DL is the convolutional neural network (CNN), which has shown optimal performance when it comes to object and image recognition tasks. The literature showed that CNN had been successfully used in numerous

computer vision applications in the built environment, such as occupancy counting, occupancy activity detection/recognition, equipment usage detection, window detection and fire/smoke detection. The information generated by such methods can be used to automatically adjust the operation of building systems and manage spaces, which can enhance building performance. The literature also highlighted some of the issues that should be resolved before computer vision-based methods can be widely adopted, such as security/privacy issues and the influence of various indoor parameters on performance.

Like in ML, each DL algorithm has its advantages and limitations, and it is imperative that several factors such as the task, data availability, practicability and computational cost are considered when selecting a suitable model for a project. As detailed in Table 6, various performance metrics or indicators related to model accuracy, sensitivity and robustness are used for model evaluation. However, other performance indicators such as computational efficiency, data requirements and complexity have not been assessed or discussed in many of the reviewed studies. Similar to the review of ML methods, most of the studies covered in the review showed that the proposed DL methods achieved good performance for various building-related applications. This adds to the challenge of finding an optimal DL algorithm. Some studies have carried out a comparative analysis with established ML/DL algorithms to identify the best-performing model.

The literature also showed that most of the studies are at the experimental stage. This can be observed in Table 6, with most of the studies are either evaluated using simulations, previously collected datasets, and experiments. While there is a lack of validation of DL models using real post-building-completion energy or post-occupancy data. Future works should consider implementing their proposed DL-based strategies in real buildings and systems.

5. Summary

The literature review showed that machine learning and deep learning techniques had been successfully applied in various building environment applications, focusing on the improvement of building performance, such as energy-comfort-air quality prediction, fault detection and diagnostics, and occupancy prediction and detection. Table 7 shows a graphical/visual summary of some of the studies covered in the literature review, which categorised them into "what type of building system", "which method/solution", and "who benefits".

While the table does not include all the reviewed studies, it can give the reader an idea about the areas where the existing studies are focused. At the same time, it also highlights that most of the studies were focused on optimising the energy efficiency and thermal comfort in building while less attention was paid to air quality. Although energy use, thermal comfort and indoor air quality (IAQ) are interrelated, there are limited studies which investigated or considered them all simultaneously. While machine learning and deep learning techniques have important roles in aiding the design and operation of buildings to provide energy-efficient, comfortable and healthy indoor environments - it requires a holistic approach. For example, while a certain method could minimise energy use and enhance thermal comfort in a building, it may have an adverse effect on other factors such as the IAQ. Future works should consider energy efficiency, thermal comfort, and IAQ aspects and their interactions when developing ML/DL-based building strategies.

It can be observed that many of the DL and ML methods were employed for energy prediction, with most focusing on cooling and/or heating loads while some considered ventilation loads. This demonstrates a gap in the literature on some topics, such as forecasting of lighting energy, plug loads and dehumidification, where more attention should be paid. In addition, most of the studies focused on mechanical HVAC systems while there is less focus on natural ventilation and other passive cooling/heating strategies.

Several studies have coupled different ML/DL methods to enhance building performance, for example, a detection model is used to obtain occupancy information from a building space, and then the data is fed into a prediction model to estimate the IAQ and thermal comfort in the future. While IAQ and thermal comfort are important indicators of the quality of conditions inside a building, there are also other factors which might influence the indoor environment quality. This includes the daylighting, visual comfort and acoustic conditions – less attention has been paid to these factors in the literature.

6. Conclusion and Recommendations for Future Works

This review provided a critical summary of the existing literature on the machine and deep learning methods for the built environment over the past decade, with special reference to holistic approaches. Different artificial intelligence (AI)-based techniques employed to resolve interconnected problems related to HVAC systems and enhance building performances were reviewed, including energy forecasting and management, indoor air quality and occupancy comfort/satisfaction prediction, occupancy detection and recognition, and fault detection and diagnosis. The present study explored existing AI-based techniques focusing on the framework, methodology, and performance.

Machine and deep learning methods have been successfully applied to building energy prediction. Many studies have shown the advantage of AI methods for predicting energy loads compared to conventional building energy simulation models. It requires fewer details and information about the building, which reduces the time of developing the model and, at the same time, AI-based models are significantly faster. However, previous works stressed the importance of selecting suitable input data and learning algorithms. Many studies recommended the integration of occupancy behaviour pattern recognition with the energy load forecasting model to enhance prediction performance. Due to the reliance on historical building data, AI-based models application in the design stage are limited. One cannot extrapolate the prediction results once changes are made to the design and operation of a building.

The literature showed that newer techniques such as Wi-fi, wireless sensors and cameras are increasingly being employed in research studies for occupancy studies and, at the same time, integrated with AI techniques. The advancement of deep learning-based techniques, such as convolutional neural network (CNN) and computer vision have made camera-based techniques one of the most popular sensing techniques for indoor environments and human recognition. Although promising, most studies have not attempted to integrate the vision-based occupancy detection approach with the HVAC control systems. Furthermore, the impact on energy demand and thermal comfort has not been well studied.

The applications of machine and deep learning in thermal comfort studies have been growing, such as in comfort prediction and management. Studies have shown that machine learning outperformed conventional and modified PMV models. Its prediction's higher accuracy and speed make it suitable for integration with demand-driven or occupancy-responsive HVAC controls, providing real-time feedback. Although not as developed as AI methods for thermal comfort optimisation, the applications of machine and deep learning in IAQ studies have been recently growing. This is probably driven by the COVID-19 pandemic and increased awareness of IAQ.

Different machine learning models are suitable for several applications related to optimising building performance. Supervised learning models are useful for applications such as building energy demand and IAQ forecasting. While unsupervised learning models are helpful for applications such as load profiling, detection and diagnostics of problems occurring in buildings and occupancy detection. In some cases, a combination of both learning techniques is employed to take advantage of the benefits of both methods.

The literature highlighted that it could be challenging to select the most suitable machine learning model for solving a problem. The recent explosive growth experienced by the research area has led to hundreds of machine learning algorithms being applied to building performance-

Table 7
A graphical/visual summary of some of the ML/DL-based studies covered in the literature review

Reference	Type					Building Energy		Indoor Environment Quality (IEQ)			Occupancy		Benefit Towards — Building		Benefit Towards — HVAC				Occupancy
	Heating	Cooling	Air-conditioning	Ventilation	Thermal Comfort	Prediction	Management	Comfort Prediction	IEQ Management	Air Quality Prediction	Detection	Recognition	Design	Performance	Response Time	Efficiency	System Design	Energy Saving	Satisfaction
Fan et al. (2019) [42]		X	X	X			X										X		
Terroso-Saenz et al. (2019) [46]	X	X	X	X			X							X				X	
Chammas et al. (2019) [45]	X			X		X										X			
Kwok and Lee (2011) [51]		X				X										X		X	
Ding et al. (2018) [52]	X					X										X		X	
Singaravel et al. (2018) [53]	X					X		X					X	X					
Xu et al. (2019) [56]	X		X	X		X								X					
Tsanas and Zifara (2012) [58]	X					X										X			
Dong and Lam (2014) [78]	X							X			X					X			
Zhao and Liu (2018) [49]	X	X	X			X										X			
Kumar et al. (2018) [54]	X					X	X						X				X		
Pham et al. (2020) [55]	X		X	X		X										X			
Chai et al. 2020 [102]	X			X				X	X										X
Peng et al. (2018) [105]	X			X		X	X				X					X		X	X
Yang et al. (2020) [106]	X			X		X		X								X			X
Gao et al. (2019) [109]						X	X	X	X							X			
Cho and Moon et al. 2022 [112]				X				X	X	X						X	X		X
Kim et al. (2021) [113]								X	X	X						X			
Yu et al. (2021) [114]								X	X	X						X			
Mateo et al. (2013) [127]				X		X				X						X			
Li et al. (2010) [128]		X				X										X		X	
Habib et al. (2016) [130]	X					X										X		X	
Carreira et al. (2018) [74]	X		X	X							X			X		X			
Gull et al. (2021) [118]	X	X														X	X		
Ciulla and D'Amico (2019) [134]	X	X				X								X		X	X		
Bilous et al. (2018) [135]	X			X		X								X		X			
Zhang et al. (2018) [136]	X	X	X	X				X											X
Wang et al. (2018) [137]							X				X	X							X
Wu et al. (2018) [138]	X			X				X								X			
Ryu and Moon (2016) [139]	X			X				X			X			X				X	
Wang et al. (2018) [140]		X		X												X			
Song et al. (2017) [142]	X	X				X			X							X			
Peng et al. (2017) [143]	X		X	X												X			
Suryanarayana et al. (2018) [156]	X					X								X				X	
Marino et al. (2016) [157]																			
Shaikh and Nor (2018) [11]	X	X	X	X												X			
Deng and Chen (2018) [160]	X			X				X	X							X			X
Wei et al. (2020) [161]	X	X				X		X						X	X	X		X	
Tien et al. (2020) [162]	X	X				X		X			X	X				X	X	X	
Tien et al. (2021) [164]	X			X		X	X	X			X	X		X		X	X	X	
Wei et al. (2022) [165]	X			X		X	X	X						X		X	X	X	
Ahn et al. (2017) [171]	X	X						X										X	X
Mutis et al. (2020) [172]		X		X				X								X	X		X

related studies. The literature showed that existing research studies considered a wide range of scope/scales (from an HVAC component to urban areas) and time scales (minute to year). This makes it difficult to find an optimal algorithm for a specific task or case. Hence, further developments and more specific guidelines are required for the built environment field to encourage best practices in the evaluation and selection of models.

While the literature showed that machine and deep learning had been successfully applied in building energy efficiency research, most of the studies are still at the experimental or testing stage, and there are limited studies which implemented machine and deep learning strategies in actual buildings and conducted the post-occupancy evaluation. Most of the studies are focused on individual buildings or few building spaces while there are limited large-scale applications and implementation in other types of buildings such as industrial and retail buildings. This is probably due to the constraints and challenges of data acquisition. Future works should consider implementing their proposed machine and deep learning-based strategies in real buildings and systems.

Based on the reviewed literature, there is also a wide range of evaluation metrics employed by the studies, which further add to the challenge of evaluating the different models. The evaluations are carried out using data from simulations, historical data and experiments. Other performance indicators such as computational efficiency, data requirements and complexity are also equally important but have not been assessed or discussed in many of the reviewed studies. Future works should provide this information as it would be useful for researchers and model developers that have certain requirements for their projects. In addition, the challenges faced in practice related to the pre-processing and uncertainty of data should be discussed in future works.

While machine learning has shown good performance in various applications such as energy forecasting, thermal comfort prediction and occupancy detection. The literature has shown that these applications are increasingly making use of deep learning techniques. The increasing amount of data generated from buildings, the advances in computational power, and the developments of algorithms have led to the increased popularity of deep learning. Although most deep learning architectures are applicable to a range of prediction or classification tasks, some studies have combined deep learning techniques for better performance. In some cases, traditional machine learning techniques are combined with deep learning techniques

As observed, each machine and deep learning algorithm has its advantages and limitations, and it is imperative that several factors such as the task, data availability, practicability and computational cost are considered when selecting a suitable model for a project. With regards to building performance-related projects, although knowledge of artificial intelligence/machine learning is key, it is also paramount for model developers to have knowledge in building engineering/energy systems. For example, this is important when selecting an appropriate prediction target and predicting parameters for the machine and deep learning model.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

ACKNOWLEDGEMENTS

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.egyai.2022.100198.

REFERENCES

[1] Met Office. UK Climate Projections. 2019 [accessed on 13 April 2020]. Available from, https://www.metoffice.gov.uk/research/collaboration/ukcp.

[2] GOV.UK. Energy Consumption in the UK 2019. 2019 [accessed on 13 April 2020] Available from: https://www.gov.uk/government/collections/energy-consumption-in-the-uk.

[3] Chen C-W, Lee C-W, Lin Y-W. Air Conditioning — Optimizing Performance by Reducing Energy Consumption. Energy & Environment 2014;25(5):1019–24.

[4] Omar O. Intelligent building, definitions, factors and evaluation criteria of selection. Alexandria Engineering Journal 2018;57(4):2903–10.

[5] Clements-Croome D. Intelligent buildings: design, management and operation. Second edition, xx. London: ICE Publishing344; 2013. pages.

[6] Sinopoli J. Smart buildings systems for architects, owners and builders. 2010.

[7] Sarwat AI, Sundararajan A, Parvez I, Moghaddami M, Moghadasi A, Amini M, Boroojeni K, Iyengar S, Pardalos P, Blaabjerg F, Madni A. Toward a Smart City of Interdependent Critical Infrastructure Networks. Sustainable Interdependent Networks. Studies in Systems, Decision and Control, 145. Cham: Springer; 2018.

[8] ACEEE American Council for an Energy-Efficient Economy. Smart Buildings: A Deeper Dive into Market Segments. 2017. Available from, https://www.aceee.org/research-report/a1703.

[9] Neri E, Coppola F, Miele V, Bibbolino B, Grassi R. Artificial intelligence: Who is responsible for the diagnosis? Radiol med 2020;125:517–21.

[10] Hernández-Muñoz JM, Vercher JB, Muñoz L, Galache JA, Presser M, Hernández Gómez LA, et al. Smart Cities at the Forefront of the Future Internet. The Future Internet, 2011. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011.

[11] Shaikh PH, Nor NBM, Nallagownden P, Elamvazuthi I. A review on optimized control systems for building energy and comfort management of smart sustainable buildings, 33; 2014. p. 409–29.

[12] Apanaviciene R, Vanagas A, Fokaides PA. Smart Building Integration into a Smart City (SBISC): Development of a New Evaluation Framework, 13; 2020. p. 2190.

[13] Google Trends. Google Trends comparison of: Machine learning, deep learning, IoT and artificial intelligence (May 2014 - May 2019). 2019 [accessed on 19 June 2019]. Available from: https://trends.google.com/trends/explore?date=2014-10-05%202019-05-15&q=Machine%20Learning,Deep%20Learning,IoT,Artificial%20intelligence.

[14] Cheng C-C, Lee D. Artificial Intelligence-Assisted Heating Ventilation and Air Conditioning Control and the Unmet Demand for Sensors: Part 1. Problem Formulation and the Hypothesis. 2019.

[15] Ngarambe J, Yun GY, Santamouris M. The use of artificial intelligence (AI) methods in the prediction of thermal comfort in buildings: energy implications of AI-based thermal comfort controls. Energy and Buildings 2020;211:109807.

[16] Lee S, Choi D-H. Energy Management of Smart Home with Home Appliances, Energy Storage System and Electric Vehicle: A Hierarchical Deep Reinforcement Learning Approach. Sensors 2020;20(7):2157.

[17] Asdrubali F, Desideri U. Chapter 1 - Introduction. In: Asdrubali F, Desideri U, editors. Handbook of Energy Efficiency in Buildings. Butterworth-Heinemann; 2019. p. 1–3.

[18] Mc Kinsey & Company. McKinsey Global Institute: Artificial Intelligence The Next Digital Frontier? Discussion Paper 2017. 2017.

[19] Glossary Hutson MAI. Artificial intelligence, in so many words. Science 2017;357 (6346):19.

[20] Dimiduk DM, Holm EA, Niezgoda SR. Perspectives on the Impact of Machine Learning, Deep Learning, and Artificial Intelligence on Materials, Processes, and Structures Engineering. Integrating Materials and Manufacturing Innovation. 2018;7(3):157–72.

[21] Bourdeau M, Xq Zhai, Nefzaoui E, Guo X, Chatellier P. Modeling and forecasting building energy consumption: A review of data-driven techniques. Sustainable Cities and Society 2019;48:101533.

[22] Rätz M, Javadi AP, Baranski M, Finkbeiner K, Müller D. Automated data-driven modeling of building energy systems via machine learning algorithms. Energy and Buildings 2019;202:109384.

[23] Amasyali K, El-Gohary NM. A review of data-driven building energy consumption prediction studies. Renewable and Sustainable Energy Reviews 2018;81: 1192–205.

[24] Sha H, Xu P, Yang Z, Chen Y, Tang J. Overview of computational intelligence for building energy system design. Renewable and Sustainable Energy Reviews 2019; 108:76–90.

[25] Mason K, Grijalva S. A review of reinforcement learning for autonomous building energy management. Computers & Electrical Engineering 2019;78:300–12.

[26] Machairas V, Tsangrassoulis A, Axarli K. Algorithms for optimization of building design: A review. Renewable and Sustainable Energy Reviews 2014;31:101–12.

[27] Jeon Y, Cho C, Seo J, Kwon K, Park H, Oh S, et al. IoT-based occupancy detection system in indoor residential environments. Building and Environment 2018;132: 181–204.

[28] Saha H, Florita AR, Henze GP, Sarkar S. Occupancy sensing in buildings: A review of data analytics approaches. Energy and Buildings 2019;188-189:278–85.

[29] Chen Z, Jiang C, Xie L. Building occupancy estimation and detection: A review. Energy and Buildings 2018;169:260–70.

[30] Mehmood MU, Chun D, Zeeshan Han H, Jeon G, Chen K. A review of the applications of artificial intelligence and big data to buildings for energy-efficiency and a comfortable indoor living environment. Energy and Buildings 2019;202:109383.

[31] Zhao H-x, Magoulès F. A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews 2012;16(6):3586–92.

[32] Wang Z, Srinivasan RS. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. Renewable and Sustainable Energy Reviews 2017;75:796–808.

[33] Ahmad T, Chen H, Guo Y, Wang J. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. Energy and Buildings 2018;165:301–20.

[34] Vázquez-Canteli JR, Nagy Z. Reinforcement learning for demand response: A review of algorithms and modeling techniques. Applied Energy 2019;235: 1072–89.

[35] Schwab K. The Fourth Industrial Revolution. Portfolio Penguin. 1st edition 2016. 5 Jan. 2017.

[36] Rics.org.. Artical intelligence: what it means for the built environment. 2017. Available from: https://www.rics.org/uk/news-insight/research/insights/artifi cial-intelligence-what-it-means-for-the-built-environment/ (Accessed: 1st August 2022).

[37] Alanne K, Sierla S. An overview of machine learning applications for smart buildings. Sustainable Cities and Society 2022;76:103445.

[38] Iddianozie C, Palmes P. Towards smart sustainable cities: Addressing semantic heterogeneity in Building Management Systems using discriminative models. Sustainable Cities and Society 2020;l62:102367.

[39] Alanne K. A novel performance indicator for the assessment of the learning ability of smart buildings. Sustainable Cities and Society 2021;72:103054.

[40] Azuatalam D, Lee W, Nijs F, Liebman A. Reinforcement learning for whole-building HVAC control and demand response. Energy and AI 2020;2:100020.

[41] European Commission. Energy Performance of Buildings. 2019 [accessed on 19 June 2019]. Available from: https://ec.europa.eu/energy/en/topics/energy-eff iciency/energy-performance-of-buildings/overview (Accessed: 1st August 2022).

[42] Fan C, Wang J, Gang W, Li S. Assessment of deep recurrent neural network-based strategies for short-term building energy predictions. Applied Energy 2019;236: 700–10.

[43] Ahmad MW, Mourshed M, Mundow D, Sisinni M, Rezgui Y. Building energy metering and environmental monitoring – A state-of-the-art review and directions for future research. Energy and Buildings 2016;120:85–102.

[44] Avancini DB, Rodrigues JJPC, Martins SGB, Rabêlo RAL, Al-Muhtadi J, Solic P. Energy meters evolution in smart grids: A review. Journal of Cleaner Production 2019;217:702–15.

[45] Chammas M, Makhoul A, Demerjian J. An efficient data model for energy prediction using wireless sensors. Computers & Electrical Engineering 2019;76: 249–57.

[46] Terroso-Saenz F, González-Vidal A, Ramallo-González AP, Skarmeta AF. An open IoT platform for the management and analysis of energy data. Future Generation Computer Systems 2019;92:1066–79.

[47] Marinakis V. Big Data for Energy Management and Energy-Efficient Buildings. Energies 2020;13:1555.

[48] Din IU, Guizani M, Rodrigues JJPC, Hassan S, Korotaev VV. Machine learning in the Internet of Things: Designed techniques for smart cities. Future Generation Computer Systems 2019.

[49] Zhao J, Liu X. A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis. Energy and Buildings 2018;174:293–308.

[50] Swamidass PM. MAPE (mean absolute percentage error)MEAN ABSOLUTE PERCENTAGE ERROR (MAPE). editor. Encyclopedia of Production and Manufacturing Management. Boston, MA: Springer US; 2000. p. 462. p..

[51] Kwok SSK, Lee EWM. A study of the importance of occupancy to building cooling load in prediction by intelligent approach. Energy Conversion and Management 2011;52(7):2555–64.

[52] Ding Y, Zhang Q, Yuan T, Yang K. Model input selection for building heating load prediction: A case study for an office building in Tianjin. Energy and Buildings 2018;159:254–70.

[53] Singaravel S, Suykens J, Geyer P. Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction. Advanced Engineering Informatics 2018;38:81–90.

[54] Kumar S, Pal SK, Singh RP. A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes. Energy and Buildings 2018;176:275–86.

[55] Pham A-D, Ngo N-T, Truong TTH, Huynh N-T, Truong N-S. Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability. Journal of Cleaner Production 2020;260:121082.

[56] Xu X, Wang W, Hong T, Chen J. Incorporating machine learning with building network analysis to predict multi-building energy use. Energy and Buildings 2019;186:80–97.

[57] Chou J-S, Bui D-K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. Energy and Buildings 2014;82:437–46.

[58] Tsanas A, Xifara A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. Energy and Buildings 2012;49:560–7.

[59] Zhou Y, Zheng S. Machine-learning based hybrid demand-side controller for high-rise office buildings with high energy flexibilities. Applied Energy 2020;262: 114416.

[60] Fan L, Li J, Zhang X. Load prediction methods using machine learning for home energy management systems based on human behavior patterns recognition. CSEE Journal of Power and Energy Systems 2020;6(3):563–71.

[61] Chen S, Yang W, Yoshino H, Levine MD, Newhouse K, Hinge A. Definition of occupant behavior in residential buildings and its application to behavior analysis in case studies. Energy and Buildings 2015;104:1–13.

[62] Tien PW, Wei S, Calautit JK, Darkwa J, Wood C. A vision-based deep learning approach for the detection and prediction of occupancy heat emissions for demand-driven control solutions. Energy and Buildings 2020;226:110386.

[63] Wei S, Tien PW, Wu Y, Calautit JK. A coupled deep learning-based internal heat gains detection and prediction method for energy-efficient office building operation. Journal of Building Engineering 2022;47:103778.

[64] D'Oca S, Hong T, Langevin J. The human dimensions of energy use in buildings: A review. Renewable and Sustainable Energy Reviews 2018;81:731–42.

[65] Barthelmes VM, Becchio C, Fabi V, Corgnati SP. Occupant behaviour lifestyles and effects on building energy use: Investigation on high and low performing building features. Energy Procedia 2017;140:93–101.

[66] Delzendeh E, Wu S, Lee A, Zhou Y. The impact of occupants' behaviours on building energy analysis: A research review. Renewable and Sustainable Energy Reviews 2017;80:1061–71.

[67] Lawrence Berkeley National Laboratory. Occupancy Simulator. 2017 [accessed on 19 June 2019]Available from: http://occupancysimulator.lbl.gov/pages/intro (Accessed: 1st August 2022).

[68] Pereira PF, Ramos NMM, Almeida RMSF. Simões ML. Methodology for detection of occupant actions in residential buildings using indoor environment monitoring systems. Building and Environment 2018;146:107–18.

[69] Zou H, Zhou Y, Yang J, Spanos CJ. Towards occupant activity driven smart buildings via WiFi-enabled IoT devices and deep learning. Energy and Buildings 2018;177:12–22.

[70] Zou H, Jiang H, Yang J, Xie L, Spanos C. Non-intrusive occupancy sensing in commercial buildings. Energy and Buildings 2017;154:633–43.

[71] Chetty K, Tang C;, Li W, Julier S, Woodbridge K. Occupancy Detection and People Counting Using WiFi Passive Radar. In: Proceedings of the 2020 IEEE Radar Conference - Renaissance meets advancing technology. The Institute of Electrical and Electronics Engineers (IEEE); 2020. In press.

[72] Wang W, Chen J, Hong T, Zhu N. Occupancy prediction through Markov based feedback recurrent neural network (M-FRNN) algorithm with WiFi probe technology. Building and Environment 2018;138:160–70.

[73] Wang W, Hong T, Li N, Wang RQ, Chen J. Linking energy-cyber-physical systems with occupancy prediction and interpretation through WiFi probe-based ensemble classification. Applied Energy 2019;236:55–69.

[74] Carreira P, Costa AA, Mansur V, Arsénio A. Can HVAC really learn from users? A simulation-based study on the effectiveness of voting for comfort and energy use optimization. Sustainable Cities and Society 2018;41:275–85.

[75] Jiang C, Masood M, Soh Y, Li H. Indoor occupancy estimation from carbon dioxide concentration. Energy and Buildings 2016;131:132–41.

[76] Zou J, Zhao Q, Yang W, Wang F. Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation. Energy and Buildings 2017;152:385–98.

[77] Diraco G, Leone A, Siciliano P. People occupancy detection and profiling with 3D depth sensors for building energy management. Energy and Buildings 2015;92: 246–66.

[78] Dong B, Lam KP. A real-time model predictive control for building heating and cooling systems based on the occupancy behavior pattern detection and local weather forecasting. Building Simulation 2014;7(1):89–106.

[79] Candanedo LM, Feldheim V. Accurate occupancy detection of an office room from light, temperature, humidity and CO2 measurements using statistical learning models. Energy and Buildings 2016;112:28–39.

[80] Pigg S, Reed J, Stacey T. Behavioral Aspects of Lighting and Occupancy Sensors in Private Offices: A Case Study of a University Office Building. 1996.

[81] Agarwal Y, Balaji B, Gupta R, Lyles J, Wei M, Weng T. Occupancy-driven energy management for smart building automation. In: Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building; Zurich, Switzerland. ACM; 2010. p. 1–6. 1878433.

[82] Mikkilineni AK, Dong J, Kuruganti T, Fugate D. A novel occupancy detection solution using low-power IR-FPA based wireless occupancy sensor. Energy and Buildings 2019;192:63–74.

[83] Yavari E, Song C, Lubecke V, Boric-Lubecke O. Is There Anybody in There?: Intelligent Radar Occupancy Sensors. IEEE Microwave Magazine 2014;15(2): 57–64.

[84] Gabrielson TB. Mechanical-thermal noise in micromachined acoustic and vibration sensors. IEEE Transactions on Electron Devices 1993;40(5):903–9.

[85] Hamilton JM, Joyce BS, Kasarda ME, Tarazaga PA. Characterization of Human Motion Through Floor Vibration. Dynamics of Civil Structures, 4. Cham: Springer International Publishing; 2014. 2014//.

[86] Millar M, Baloglu S. Hotel Guests' Preferences for Green Guest Room Attributes. Cornell Hospitality Quarterly 2011;52(3):302–11.

[87] Gul MS, Patidar S. Understanding the energy consumption and occupancy of a multi-purpose academic building. Energy and Buildings 2015;87:155–65.

[88] Zhao Y, Zeiler W, Boxem G, Labeodan T. Virtual occupancy sensors for real-time occupancy information in buildings. Building and Environment 2015;93:9–20.

[89] Patel SN, Reynolds MS, Abowd GD. Detecting Human Movement by Differential Air Pressure Sensing in HVAC System Ductwork: An Exploration in Infrastructure Mediated Sensing. Pervasive Computing; 2008. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. //.

[90] Chen D, Barker S, Subbaswamy A, Irwin D, Shenoy P. Non-Intrusive Occupancy Monitoring using Smart Meters. In: Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings; Roma, Italy. 2528294. ACM; 2013. p. 1–8.

[91] Corna A, Fontana L, Nacci AA, Sciuto D. Occupancy detection via iBeacon on Android devices for smart building management. In: Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition; Grenoble, France. 2755896: EDA Consortium; 2015. p. 629–32.

[92] Mahmoud S, Lotfi A, Langensiepen C. Behavioural pattern identification and prediction in intelligent environments. Applied Soft Computing 2013;13(4): 1813–22.

[93] Wheeler A. Commercial Applications of Wireless Sensor Networks Using ZigBee. IEEE Communications Magazine 2007;45(4):70–7.

[94] Dodier RH, Henze GP, Tiller DK, Guo X. Building occupancy detection through sensor belief networks. Energy and Buildings 2006;38(9):1033–43.

[95] İçoğlu O, Mahdavi A. VIOLAS: A vision-based sensing system for sentient building models. Automation in Construction 2007;16(5):685–712.

[96] Sangogboye FC, Arendt K, Singh A, Veje CT, Kjærgaard MB, Jørgensen BN. Performance comparison of occupancy count estimation and prediction with common versus dedicated sensors for building model predictive control. Building Simulation 2017;10(6):829–43.

[97] Trivedi MM, Shinko Yuanhsien C, Childers EMC, Krotosky SJ. Occupant posture analysis with stereo and thermal infrared video: algorithms and experimental evaluation. IEEE Transactions on Vehicular Technology 2004;53(6):1698–712.

[98] Ekwevugbe T, Brown N, Fan D. A design model for building occupancy detection using sensor fusion. In: 2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST); 2012. 18-20 June 2012.

[99] Ijjina EP, Chalavadi KM. Human action recognition in RGB-D videos using motion sequence information and deep learning. Pattern Recognition 2017;72:504–16.

[100] Acharya D, Yan W, Khoshelham K. Real-time image-based parking occupancy detection using deep learning 2018.

[101] Castro D, Hickson S, Bettadapura V, Thomaz E, Abowd G, Christensen H, et al. Predicting Daily Activities From Egocentric Images Using Deep Learning. Proc Int Symp Wearable Comput 2015;2015:75–82.

[102] Chai Q, Wang H, Zhai Y, Yang L. Using machine learning algorithms to predict occupants' thermal comfort in naturally ventilated residential buildings. Energy Build 2020;17:109937.

[103] Hu W, Wen Y, Guan K, Jin G, iTCM Tseng K. Toward Learning-Based Thermal Comfort Modeling via Pervasive Sensing for Smart Buildings. IEEE Internet of Things Journal 2018;5:4164–77.

[104] Chaudhuri T, Soh Y, Li H, Xie L. Machine learning based prediction of thermal comfort in buildings of equatorial Singapore. In: IEEE International Conference on Smart Grid and Smart Cities (ICSGSC); 2017. p. 72–7.

[105] Peng Y, Rysanek A, Nagy Z, Schlüter A. Using machine learning techniques for occupancy-prediction-based cooling control in office buildings. Applied Energy 2018;211:1343–58.

[106] Yang S, Wan M, Chen W, Ng B, Dubey S. Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization. Appl. Energy. 2020;271:115147.

[107] Lu S, Wang W, Lin C, Hameen EC. Data-driven simulation of a thermal comfort-based temperature set-point control with ASHRAE RP884. Building and Environment 2019;156:137–46.

[108] Valladares W, Galindo M, Gutiérrez J, Wu WC, Liao KK, Liao JC, Lu KC, Wang CC. Energy optimization associated with thermal comfort and indoor air control via a deep reinforcement learning algorithm. Build. Environ. 2019;155:105–17.

[109] Gao G, Li J, Wen Y. Energy-Efficient Thermal Comfort Control in Smart Buildings via Deep Reinforcement Learning. Computing Research Repository (CoRR) 2019. abs/1901.04693.11.

[110] Fard ZQ, Zomorodian ZS, Korsavi SS. Application of machine learning in thermal comfort studies: A review of methods, performance and challenges. Energy and Buildings 2022;256:111771.

[111] CIBSE Chartered Institute of Building Services Engineers. CIBSE Knowledge Series: KS17 Indoor air quality and ventilation October 2011.

[112] Cho JH, Moon JW. Integrated artificial neural network prediction model of indoor environmental quality in a school building. Journal of Cleaner Production 2022; 344:131083.

[113] Kim J, Hong Y, Seong N, Kim DD. Assessment of ANN Algorithms for the Concentration Prediction of Indoor Air Pollutants in Child Daycare Centers. Energies 2022;15(7):2654.

[114] Yu K-H, Chen Y-A, Jaimes E, Wu W-C, Liao K-K, Liao J-C, Lu K-C, Sheu W-J, Wang C-C. Optimization of thermal comfort, indoor quality, and energy-saving in campus classroom through deep Q learning. Case Studies in Thermal Engineering 2021;24:100842.

[115] Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. Electron Markets 2021;31:685–95.

[116] Ahmad T, Chen H. A review on machine learning forecasting growth trends and their real-time applications in different energy systems. Sustainable Cities and Society 2020;54:102010.

[117] Miller C, Nagy Z, Schlueter A. A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings. Renewable and Sustainable Energy Reviews 2018;81(1):1365–77.

[118] Gull CQ, Aguilar J, R-Moreno MD. A semi-supervised learning approach to study the energy consumption in smart buildings. In: 2021 IEEE Symposium Series on Computational Intelligence (SSCI); 2021. p. 1–7.

[119] Fu Q, Han Z, Chen J, Lu Y, Wu H, Wang Y. Applications of reinforcement learning for building energy efficiency control: A review. Journal of Building Engineering 2022;50:104165.

[120] Wang P, Fan E, Wang P. Comparative Analysis of Image Classification Algorithms Based on Traditional Machine Learning and Deep Learning. Pattern Recognition Letters 2021;141:61–7.

[121] Google. AI & Machine Learning Products: AI Platform Documentation - Machine Learning Workflow. 2019 [accessed on 19 June 2019]. Available from: https ://cloud.google.com/ml-engine/docs/tensorflow/ml-solutions-overview.

[122] Amazon Web Services A. Amazon SageMaker. 2019 [accessed on 19 June 2019]. Available from: https://aws.amazon.com/sagemaker/.

[123] Freitag D. Information extraction from HTML: application of a general machine learning approach. In: Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence; Madison, Wisconsin, USA. 295726. American Association for Artificial Intelligence; 1998. p. 517–23.

[124] Raschka S. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning. Computing Research Repository (CoRR) 2018. abs/1811.12808.

[125] Domingos P. A few useful things to know about machine learning. Commun ACM 2012;55(10):78–87.

[126] Wang Z, Liu J, Zhang Y, Yuan H, Zhang R, Srinivasan RS. Practical issues in implementing machine-learning models for building energy efficiency: Moving beyond obstacles. Renewable and Sustainable Energy Reviews 2021;143:110929.

[127] Mateo F, Carrasco JJ, Sellami A, Millán-Giraldo M, Domínguez M, Soria-Olivas E. Machine learning methods to forecast temperature in buildings. Expert Systems with Applications 2013;40(4):1061–8.

[128] Li X, Deng Y, Ding L, Jiang L. Building cooling load forecasting using fuzzy support vector machine and fuzzy C-mean clustering. In: 2010 International Conference on Computer and Communication Technologies in Agriculture Engineering; 2010. 12-13 June 2010.

[129] Xue L, Yong D, Hao T, Feng X. A data mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data. Energy and Buildings 2021;231:110601.

[130] Habib U, Hayat K, Zucker G. Complex building's energy system operation patterns analysis using bag of words representation with hierarchical clustering. Complex Adaptive Systems Modeling 2016;4(1):8.

[131] Hong D, Wang H, Whitehouse K. Clustering-based Active Learning on Sensor Type Classification in Buildings. In: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management; Melbourne, Australia. 2806574. ACM; 2015. p. 363–72.

[132] Guo Y, Wall JK, West S, Li J. DA-13C 018 Intelligent Model Based Fault Detection and Diagnosis for HVAC System Using Statistical Machine Learning Methods 2013.

[133] Trabelsi D, Mohammed S, Chamroukhi F, Oukhellou L, Amirat Y. An Unsupervised Approach for Automatic Activity Recognition based on Hidden Markov Model Regression. IEEE Transactions on Automation Science and Engineering 2013;10(3):829–35.

[134] Ciulla G, D'Amico A. Building energy performance forecasting: A multiple linear regression approach. Applied Energy 2019;253:113500.

[135] Bilous I, Deshko V, Sukhodub I. Parametric analysis of external and internal factors influence on building energy performance using non-linear multivariate regression models. Journal of Building Engineering 2018;20:327–36.

[136] Zhang W, Liu F, Fan R. Improved thermal comfort modeling for smart buildings: A data analytics study. International Journal of Electrical Power & Energy Systems 2018;103:634–43.

[137] Wang L, Lee EWM, Yuen RKK. Novel dynamic forecasting model for building cooling loads combining an artificial neural network and an ensemble approach. Applied Energy 2018;228:1740–53.

[138] Wu Z, Li N, Peng J, Cui H, Liu P, Li H, et al. Using an ensemble machine learning methodology-Bagging to predict occupants' thermal comfort in buildings. Energy and Buildings 2018;173:117–27.

[139] Ryu SH, Moon HJ. Development of an occupancy prediction model using indoor environmental data based on machine learning techniques. Building and Environment 2016;107:1–9.

[140] Wang W, Chen J, Hong T. Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings. Automation in Construction 2018;94:233–43.

[141] Johannesen NJ, Kolhe M, Goodwin M. Relative evaluation of regression tools for urban area electrical energy demand forecasting. Journal of Cleaner Production 2019;218:555–64.

[142] Song K, Kwon N, Anderson K, Park M, Lee H-S, Lee S. Predicting hourly energy consumption in buildings using occupancy-related characteristics of end-user groups. Energy and Buildings 2017;156:121–33.

[143] Peng Y, Rysanek A, Nagy Z, Schlüter A. Occupancy learning-based demand-driven cooling control for office spaces. Building and Environment 2017;122:145–60.

[144] Xiong L, Yao Y. Study on an adaptive thermal comfort model with K-nearest-neighbors (KNN) algorithm. Building and Environment 2021;202:108026.

[145] Qiong L, Peng R, Qinglin M. Prediction model of annual energy consumption of residential buildings. In: 2010 International Conference on Advances in Energy Engineering; 2010. 19-20 June 2010.

[146] Liu Y, Chen H, Zhang L, Feng Z. Enhancing building energy efficiency using a random forest model: A hybrid prediction approach. Energy Reports 2021;7: 5003–12.

[147] Farzaneh H, Malehmirchegini L, Bejan A, Afolabi T, Mulumba A, Daka PP. Artificial Intelligence Evolution in Smart Buildings for Energy Efficiency. Applied Sciences. 2021;11(2):763.

[148] Zhang W, Wu Y, Calautit JK. A review on occupancy prediction through machine learning for enhancing energy efficiency, air quality and thermal comfort in the built environment. Renewable and Sustainable Energy Reviews 2022;167: 112704.

[149] Drakos G. Towards Data Science. How to select the Right Evaluation Metric for Machine Learning Models: Part 1 Regression Metrics. 2018 [accessed on 19 June 2019]. Available from: https://towardsdatascience.com/how-to-select-the-right-evaluation-metric-for-machine-learning-models-part-1-regrression-metrics-360 6e25beae0.

[150] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436.

[151] Kim J, Moon J, Hwang E, Kang P. Recurrent inception convolution neural network for multi short-term load forecasting. Energy and Buildings 2019;194: 328–41.

[152] Cai M, Pipattanasomporn M, Rahman S. Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques. Applied Energy 2019; 236:1078–88.

[153] Almalaq A, Zhang JJ. Evolutionary Deep Learning-Based Energy Consumption Prediction for Buildings. IEEE Access 2019;7:1520–31.

[154] Zhang Y, Ling C. A strategy to apply machine learning to small datasets in materials science. npj Comput Mater 2018;4:25.

[155] Somu N, Raman GMR, Ramamritham K. A deep learning framework for building energy consumption forecast. Renewable and Sustainable Energy Reviews 2021; 137:110591.

[156] Suryanarayana G, Lago J, Geysen D, Aleksiejuk P, Johansson C. Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods. Energy 2018;157:141–9.

[157] Marino DL, Amarasinghe K, Manic M. Building energy load forecasting using Deep Neural Networks. In: IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society; 2016. p. 7046–51.

[158] Somu N, Sriram A, Kowli A, Ramamritham K. A hybrid deep transfer learning strategy for thermal comfort prediction in buildings. Building and Environment 2021;204:108133.

[159] Gao N, Shao W, Rahaman MS, Zhai J, David K, Salim FD. Transfer learning for thermal comfort prediction in multiple cities, 195. Building and Environment; 2021, 107725. Volume.

[160] Deng Z, Chen Q. Artificial neural network models using thermal sensations and occupants` behavior for predicting thermal comfort. Energy and Buildings 2018; 174:587–602.

[161] Wei S, Tien PW, Calautit JK, Wu Y, Boukhanouf R. Vision-based detection and prediction of equipment heat gains in commercial office buildings using a deep learning method. Applied Energy 2020;277:115506.

[162] Tien PW, Wei S, Calautit JK, Darkwa J, Wood C. Occupancy heat gain and prediction using deep learning approach for reducing building energy demand. Journal of Sustainable Development of Energy, Water and Environment Systems. 2020;3:1080378.

[163] Tien PW, Wei S, Calautit JK. A Computer Vision-Based Occupancy and Equipment Usage Detection Approach for Reducing Building Energy Demand. Energies 2020; 14:156.

[164] Tien PW, Wei S, Lui T, Calautit JK, Darkwa J, Wood C. A deep learning approach towards the detection and recognition of opening of windows for effective management of building ventilation heat losses and reducing space heating demand. Renewable Energy 2021;177:603–25.

[165] Wei S, Tien PW, Chow T, Wu Y, Calautit J. Deep learning and computer vision based occupancy CO2 level prediction for demand-controlled ventilation (DCV). Journal of Building Engineering 2022;56:104715.

[166] Pincott J, Tien PW, Wei S, Calautit J. Development and evaluation of a vision-based transfer learning approach for indoor fire and smoke detection. Building Services Engineering Research and Technology 2022;43(3):319–32.

[167] Ahn J, Shin D, Kim K, Yang J. Indoor Air Quality Analysis Using Deep Learning with Sensor Data. Sensors 2017;17(11):2476.

[168] Mutis I, Ambekar A, Joshi V. Real-time space occupancy sensing and human motion analysis using deep learning for indoor air quality control. Automation in Construction 2020;116:103237.

[169] Taheri S, Ahmadi A, Mohammadi-Ivatloo B, Asadi S. Fault detection diagnostic for HVAC systems via deep learning algorithms. Energy and Buildings 2021;250: 111275.

[170] Yabin G, Zehan T, Huanxin C, Guannan L, Jiangyu W, Ronggeng H, Jiangyan L, Tanveer A. Deep learning-based fault diagnosis of variable refrigerant flow air-conditioning system for building energy saving. Applied Energy 2018;225: 732–45.

[171] Fan C, Xiao F, Zhao Y, Wang J. Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. Applied Energy 2018;211:1123–35.