Research Repository UCD

Title	Application Of Intelligent Algorithms For Residential Building Energy Performance Rating Prediction
Authors(s)	Ali, Usman, Shamsi, Mohammad Haris, Alshehri, Fawaz, Mangina, Eleni, O'Donnell, James
Publication date	2019-09-04
Publication information	Ali, Usman, Mohammad Haris Shamsi, Fawaz Alshehri, Eleni Mangina, and James O'Donnell. "Application Of Intelligent Algorithms For Residential Building Energy Performance Rating Prediction." IBPSA, September 4, 2019. https://doi.org/10.26868/25222708.2019.210232.
Conference details	The 16th International Building Simulation Association, Rome, Italy, 2-4 September 2019
Publisher	IBPSA
Item record/more information	http://hdl.handle.net/10197/12263
Publisher's version (DOI)	10.26868/25222708.2019.210232

Downloaded 2026-05-01 23:36:27

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

Application of Intelligent Algorithms for Residential Building Energy Performance Rating Prediction

Usman Ali[1], Mohammad Haris Shamsi[1], Fawaz Alshehri[1],
Eleni Mangina[2] and James O'Donnell[1]
[1]School of Mechanical and Materials Engineering and Energy Institute, UCD, Ireland
[2]School of Computer Science and Informatics, University College Dublin (UCD), Ireland

Abstract

Energy Performance Certificates (EPC) provide an indication of buildings' energy use. The creation of an EPC for individual building requires information surveys. Hence, these ratings are typically non-existent for entire building stock. This paper addresses these information gaps using machine-learning models. Developed models were evaluated with Irish EPC data that included approximately 650,000 residential buildings with 199 inputs variables. Results indicate that the deep learning algorithm produces results with highest accuracy level of 88% when only 82 input variables are available. This identified approach will allow stakeholders such as authorities, policymakers and urban-planners to determine the EPC rating for rest of the building stock using limited data.

Introduction

Over the past decades, energy use worldwide has increased significantly, and building sector is one of the largest energy consumers. This sector is responsible for approximately 40% of the total energy demand and one-third of CO_2 emissions (EU-Energy, 2018; EESI, 2018). One of the main reasons behind this is the below average energy performance exhibited by the current building stock. European member states have devised mandatory policies for building energy ratings to examine buildings' energy performance. Building energy ratings encourages urban planners and energy policymakers to formulate sustainable and energy conservation measures (Jones et al., 2001). The European Union (EU) has mandated the Energy Performance of Buildings Directive (EPBD) that aims to reduce CO_2 emissions (EU, 2018). The member states of EPBD must develop minimum energy performance standards and Energy Performance Certificates (EPCs) for both commercial and residential buildings. Their goal is to set minimum energy performance standards for new buildings and large existing buildings subjected to major renovation (Fokaides et al., 2017).

Building EPC rating requires extensive data collected through surveys in order to perform rating calculations. The main parameters required for performance calculations are daylighting, internal heat gains, thermal insulation, heat gains through glazed openings, ventilation, heating system, domestic hot water system, indoor climate, renewables, etc. Collecting this information for each building is a time-consuming and challenging task (Collins and Curtis, 2018).

In recent years, numerous approaches have been used for predicting energy demand of the building stock (Zhao and Magoulès, 2012; Amasyali and El-Gohary, 2018; Wei et al., 2018). These approaches are divided between as physics-based and data-driven approaches (Rahman et al., 2018). Physics-based approaches estimate energy use by using dynamic equations; the most common tools are EnergyPlus (Crawley et al., 2001) and TRNSYS (Trnsys, 2000). However, these tools require a large number of input parameters to estimate the energy demand , consequently, often do not scale to complex scenarios. The data-driven approaches use statistical and machine learning (ML) techniques to predict energy demand (Kapetanakis D., 2017). These approaches provide an accurate prediction of energy demand when the models are treated with an enriched training dataset (Li and Wen, 2014).

Generally, machine learning algorithms are divided into two main categories such as supervised and unsupervised Learning (Amasyali and El-Gohary, 2018). Supervised learning can be further grouped into regression and classification algorithms. A classification algorithm is used when the output variable is a label, such as energy rating and building type (Kontokosta and Tull, 2017). Regression algorithms are used when the output variable is a real value such as energy consumption (Deb et al., 2017; Yildiz et al., 2017). Some common supervised learning algorithms include the nearest neighbor, naive Bayes, rule induction, deep learning, Support Vector Machines (SVM) and neural networks, etc. Unsupervised learning is used when there are no corresponding output variables for the input data (Ali et al., 2018). Some common unsupervised learning algorithms include k-means clustering, association rules, etc.

This work introduces a methodology to facilitate the prediction of EPCs using different supervised machine learning algorithms. The main objective of this paper is to formulate an intelligent machine learning model that can be used to predict building energy performance. As there exists a multitude of machine learning algorithms, this research also compares these different algorithms in terms of prediction accuracy when applied to predict building energy ratings using existing EPCs for the building stock. Furthermore, the research also investigates the importance of classification and identification of key variables that influence building energy performance in the feature selection process.

The paper is organized as follows: Section 2 provides detailed discussion on methodology for residential building energy performance rating prediction. Section 3 discusses the case study of Irish building stock and compares different machine learning algorithms in term of prediction performance. Finally, conclusions are discussed in Section 4.

Methodology

The prediction of building energy performance using intelligent machine learning algorithms requires the steps described in Figure 1. The methodology starts with data pre-processing followed by feature selection, data splitting for training and testing purposes, implementation of analysis learning algorithms and finally the results validation.

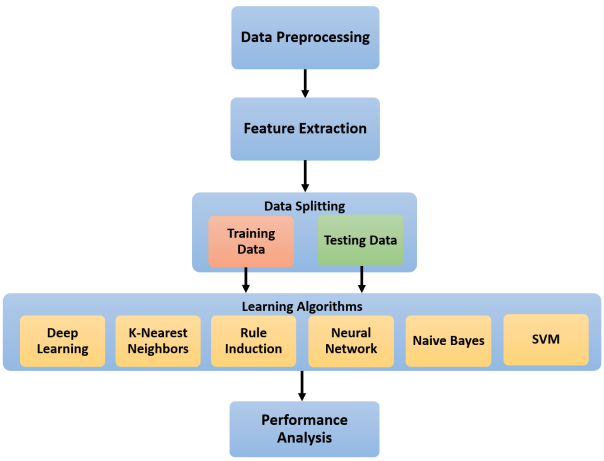

Figure 1: Methodology for residential building energy performance rating prediction using machine learning algorithm.

Data Pre-processing

The first step of methodology is data pre-processing of the EPC data; this is used as input to the learning algorithms. The building EPC data obtained through extensive surveys, and as a result, contains many incomplete, missing or inconsistent information including essential variables. Therefore, the data needs to be cleaned to remove noise, errors or outliers before it is suitable as input to machine learning algorithms. Data pre-processing enhances the accuracy of prediction results (Molina-Solana et al., 2017). Some of the important pre-processing techniques are data cleaning, data transformation, outlier detection etc. (Ali et al., 2016).

Data transformation is the process of converting data types, for example, nominals to numerics. Outliers are data points that posses exceptionally different distribution information and deviate from the majority of the data. Outlier detection is an essential step before executing a learning algorithm (Molina-Solana et al., 2017). The most common outlier detection techniques are distance-based, density-based and Local Outlier Factor (LOF). In this paper, the LOF algorithm is used for detecting the outliers from the EPC dataset because this algorithm is viable for large datasets (Tang et al., 2001). LOF measures the density of objects between each another using the nearest neighbors distance formula (Breunig et al., 2000).

Feature Extraction

Feature extraction is an essential step before training the learning model. The primary goal is removal of irrelevant or redundant variables and determination of features which are most important for predicting the model performance. The feature selection procedure reduces the dimensionality of model inputs that in turn reduces the model complexity and computational load, and also improves learning accuracy. Feature selection is usually performed using statistical and engineering methods. Statistical methods use various statistical or data mining algorithms such as ANOVA, Chi-Square, SVM, genetic algorithms, etc (Gao and Malkawi, 2014; Fan et al., 2017; Feng et al., 2017). Engineering methods are informed by engineering judgment and existing practices in the literature (Famuyibo et al., 2012; Egan et al., 2018; Kapetanakis et al., 2015). In this paper, engineering methods are used for feature selection.

Data splitting

Data splitting is the process of dividing the dataset into two subsets; a training set (a subset to train a model) and a test set (a subset to test the trained model) (Picard and Berk, 1990). Generally, data splitting applies one of two techniques, random data splitting and cross-validation. In random data splitting, the random data is split into training and test sets, according to a 80/20 split respectively. Cross-validation is the most common technique to achieve balance between minimal bias and variance of the training model. In cross-validation, the data is first divided into k of subsets and then the data splitting process is applied to each sub-sets. In each kth iteration, a different subset is used for testing while the other k-1 are used for training.

Learning Algorithms

Learning algorithms are used to predict the class of given set of data points; classes are also known as labels or categories. In this paper, six different learning algorithms are used for rating prediction. These six algorithms offer excellent performance when used for energy forecasting as evident from previous studies (Wei et al., 2018; Amasyali and El-Gohary, 2018). The following sub-sections describe each of these in detail.

kNN

k-Nearest-Neighbours (kNN) is a non-parametric classification method that classifies objects based on closest training data in the feature space. The classification depends on the value of k, and the type of distance algorithm. This algorithm is robust to noisy and large training datasets. However, computational time is quite high, and it is considered to be a lazy learner.

Rule Induction

Rule induction is based on the set of formal rules extracted from a set of training data. The algorithm iteratively prunes rules until the error rate is greater than 50% of data and there are no positive examples left. This algorithm is simple to implement. Furthermore, the algorithm is quite efficient when dealing with large datasets. However, the algorithm doesn't respond well to noisy datasets.

Neural Network

Neural networks, also known as Artificial Neural Networks (ANN), is a learning model based on the structure and functional aspects of biological neural networks. Neural networks are widely recognized as a powerful learning algorithm. These networks can be trained with any number of layers and inputs, as well as a hidden layer consisting of units that transform the input into the output layer. This algorithm functions efficiently with nonlinear training data with a large number of inputs. Often, this algorithm is computationally expensive when used to train models.

Naive Bayes

Naive Bayes learning algorithms are based on Bayes' theorem with the strong (naive) conditional independence assumption between the given features. This algorithm is a low-variance classifier with high-bias. It is relatively simple to implement and has good performance with small datasets. Its major limitation is the assumption that every feature in the training dataset is independent, which isn't always the case.

SVM

A Support Vector Machine (SVM) is a discriminative classifier. In other words, given a set of labeled training data, the algorithm outputs a model that assigns existing categories to the new examples, thus, making a non-probabilistic binary linear classifier. It uses a subset of training points and is a highly accurate technique. Its major limitation is that the training time can be high for larger datasets. It can also become less effective on overlapping classes with noisier datasets.

Deep Learning

Deep learning, also known as deep structured learning, uses neural network architectures. It is based on a multi-layer artificial neural network, and the network can contain a large number of hidden layers consisting of neurons, output labels, and epochs that are trained with a set of propagation formulae. The major difference between the neural network and deep learning techniques is that deep learning includes an unlimited number of layers and neurons. The deep leaning technique admits higher levels of abstraction and gives higher prediction accuracy. The major limitation is that the complexity of the model is increased when applied to large datasets.

Performance Analysis

To examine the effectiveness of the learning prediction models, adopted performance indices such as the Root Mean Squared Error (RMSE), ACCuracy (ACC), and Classification Error (CE) are used (Wei et al., 2018). These are computed as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(c_i - \bar{c}_i)^2} \tag{1}$$

$$ACC = \frac{TP + TN}{P + N} \tag{2}$$

$$CE = 1 - \frac{TP + TN}{P + N} \tag{3}$$

The most common way to evaluate the performance of the prediction algorithm is by using accuracy. Accuracy is a percentage of the correct number of predictions from all results. RMSE shows the differences between actual and predictions outcomes. CE is the percentage of the incorrect predictions out of all results.

Results and Discussion

The main objective of this paper is to develop an energy rating prediction system for energy policymakers. The proposed methodology is applied to the publicly available Irish residential Building Energy Per-

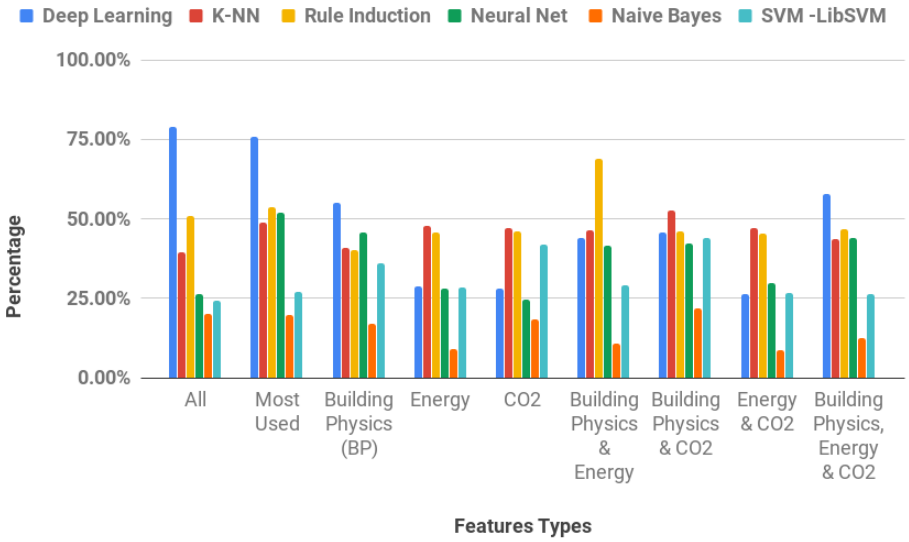

Figure 2: Learning algorithms' accuracy comparison of the detail building EPC rating prediction.

Table 1: The Irish building stock's EPC simple and detailed rating labels with primary energy consumption (kWh/m^2/year) for classification algorithms.

Detailed Rating	Energy KWh/m^2/yr	Simple Rating	Energy KWh/m^2/yr
A1	25 or less	A	75 or less
A2	26-50		
A3	51-75		
B1	76-100	B	76-150
B2	101-125		
B3	126-150		
C1	151-175	C	151-225
C2	176-200		
C3	201-225		
D1	226-260	D	226-300
D2	261-300		
E1	301-340	E	301-380
E2	341-380		
F	381-450	F	381-450
G	451 or more	G	451 or more

formance Certificate (EPC) dataset. An EPC rating is given to a building based on the overall energy building performance measured in terms of energy and carbon dioxide emissions. The building energy rating varies on a scale of A1 to G, with A1 and G being the most and least energy efficient ratings respectively. A building's energy rating is calculated using Ireland's official Dwelling Energy Assessment Procedure (DEAP) software. The publicly available EPC dataset contains more than 695,000 Irish residential buildings' data with 199 inputs variables such as building physics, energy, and CO_2 information. There are more than 1,983,715 residential building in Ireland, however, the EPC data is available for only $\approx$ 39% of residential building stock (SEAI, 2018). This paper focuses on predicting the energy rating of the rest 70% of the stock with limited variables using machine learning algorithms. For the case study, the learning algorithms are trained and tested on Dublin city's EPC data that represents 30% of the EPC building stock in Ireland.

After EPC data collection, the next step is data pre-processing. The EPC data is based on surveys and questionnaires. Therefore there are a lot of inconsistent, missing, irrelevant and incomplete values in the dataset. During pre-processing, missing or zero values are removed or replaced with their averages and irrelevant values are filtered out from the data. Machine learning algorithms generally work with numeric values, hence, categorical or nominal values must be converted into numerical values. To remove inconsistent values or outliers from the data, the LOF algorithm is used to clear out these outliers (by using Euclidean distance). The energy rating is the label or output variable that is used for the classification. In this paper, energy rating is divided into two categories: simple rating and detailed rating. The simple rating includes the 7 rating labels, such as A, B, C, D, E, F and G. The detailed rating includes 15-rating labels, such as A1, A2, A3 etc. A summary of implemented ratings is shown in Table 1. Irish EPC rating is based on a detailed rating chart with different band and sub-band levels. To generate the simple rating, we further aggregated the same letter sub-band ratings (A1, A2 and A3) into a single letter band rating (A). This kind of classification would help to identify the effects of aggregated and detailed ratings on model accuracy.

The EPC data has 199 inputs variables, which are

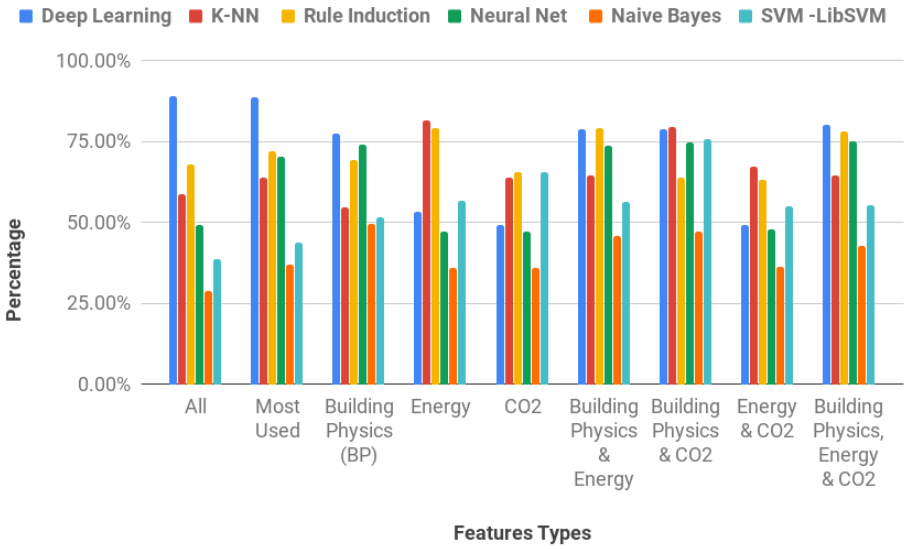

Figure 3: Learning algorithms' accuracy comparison of the simple building EPC rating prediction.

used to calculate the energy rating of a building. Feature selection processes groups similar types of variables based on the engineering method approach. The groups includes the variables that are most used, building physics parameters, energy consumption values, CO_2 emissions and a combination of all these variables. The 199 input variables constitute of 65 detailed floor level variables (33%), 52 variables (26%) with minimum variations or missing values and the remaining 82 most likely used variables (41%). We ignored the first batch of the detailed 65 floor level variables and used average value of each variable per building. Similarly, We also ignored the second batch of 52 variables with missing values. The rest 82 variables were identified to be the most likely used in energy simulations. These 82 variables were further segregated into groups according to building features, namely, building physics parameters, energy consumption values, CO_2 emissions and a combination of all these variables

The most likely used variables were identified based on existing literature. Studies by Famuibo at al. and Egan et al. have identified variables that are relevant and would influence the building's energy performance of the Irish building stock (Famuyibo et al., 2012; Egan et al., 2018). Building physics variables include U-values, area, fuel type, etc. CO_2 emissions are analyzed based on emissions from different sources that include CO_2 emissions from lighting use, equivalent emissions from water usage, etc. Energy variables include consumption of lighting, main space, water, etc.

The data is split into two parts after the feature selection process to create training or testing data. The cross-validation algorithm is used for data splitting. The EPC data is partitioned into 10 subsets of equal size.

The EPC data is trained and compared using six different algorithms and a robust model input is chosen to evaluate the sensitivity of the learning algorithms. For the deep learning algorithm, two hidden layers are used for the size of 50. The dataset is iterated 10 times to achieve the best results. The layer size for the neural network algorithm is calculated from the number of input attributes for the training data. 500 training cycles are used for neural network training with 0.3 weights at each step. For the K-NN algorithm, the Euclidean distance measure is used for distance calculation. For the Navie Bayes algorithm, a value of 0.9 is used as the sample ratio to train the dataset (for growing and pruning). To examine the performance of all of the learning models, the results are calculated for the simple and detailed ratings as shown in Figure 2 and Figure 3.

In the simple building EPC rating prediction (Figure 3), all the variable groups except CO_2 and Energy-CO_2 groups yield accuracy levels of more than 75%. On the other hand, the prediction accuracy of detailed building EPC rating is quite lower for all variable groups (Figure 2) when compared to simple building EPC rating mainly because of detailed label classification. Furthermore, it can also be inferred that the building physics variable group has the strongest influence on the prediction accuracy of simple rating when compared to other variable groups, namely, Energy, CO_2, and Energy and CO_2. The accuracy is further improved when the building physics variable group is combined with CO_2 and Energy variable groups.

The best algorithm with the highest accuracy and minimum RMSE for each group of feature types is shown in Table 2. The results show that deep learn-

Table 2: Best learning algorithms performance analysis for energy rating prediction by using different feature groups types.

Features Types	Input Variables	Algorithm	Simple Rating		Detail Rating	
			ACC	RMSE	ACC	RMSE
All	199	Deep Learning	89.11%	0.284	78.84%	0.399
Most Used	82	Deep Learning	88.48%	0.292	75.79%	0.427
Building Physics	39	Deep Learning	77.30%	0.411	54.91%	0.612
Energy	18	K-NN	81.49%	0.417	47.79%	0.417
CO_2	16	K-NN	63.88%	0.601	47.05%	0.728
Building Physics & Energy	54	Rule Induction	79.26%	0.39	68.77%	0.545
Building Physics & CO_2	52	K-NN	79.59%	0.388	52.66%	0.688
Energy & CO_2	24	K-NN	65.18%	0.59	47.19%	0.727
Building Physics, Energy & CO_2	67	Deep Learning	80.14%	0.383	57.75%	0.581

ing, K-NN and rule induction algorithms performed best for the different type of features. The highest accuracy achieved for the simple rating was 89% with 0.28 RMSE while that for the detailed rating was 78% with 0.4 RMSE. However, these accuracy values are achieved when all the variables are used for training and testing. The best results with minimum number of variables (82) for the simple rating is 88% with 0.29 RMSE while that for the detailed rating is 76% with 0.42 RMSE shown in Figure 4. Results indicate that intelligent machine learning algorithms require only 82 variables to predict a building's energy rating. Furthermore, the results establish that the model accuracy obtained using the simple rating is higher than the one obtained using the detailed rating.

Figure 4: Comparison of performance analysis of simple vs detail rating using deep learning algorithm with most used features.

Conclusions and Future Work

EPC provides an overview of the building stock energy performance that helps the stakeholders to formulate policy measures for reducing the energy consumption and CO_2 emissions. EPC rating calculation for a building is a time consuming, and complex task and a number of important variables are required for the calculation. The number of resources, required to perform the analytical, extensive, and time-consuming calculation, are significantly increased at an urban scale. The devised methodology in this research implements machine learning algorithms to identify the building energy rating for an entire urban building stock. The proposed methodology is able to calculate a building's current energy performance even with a limited knowledge of the building dynamics. The feature selection method identifies key variables influencing the building energy performance. For instance, the Irish EPC data uses 199 variables to classify buildings based on their energy rating. However, the proposed methodology requires only 82 variables to predict a building's energy rating. The values obtained for accuracy are 88% and 79% for simple and detailed ratings respectively by using the deep learning algorithm.

Currently, EPC data covers ≈ 30- 50% of entire building stock. This identified approach will allow stakeholders such as the local authorities, energy policymakers and urban planners to determine the EPC rating for the rest of the building stock by using limited data. As such, the stakeholders will be able to make informed decisions when planning retrofit measures at a large scale.

As the proposed methodology is only tested with DEAP energy performance calculation software, variations that might occur with other software. The performance of the proposed best algorithm might turn out to be different for other energy performance calculation software, for instance, EnergyPlus. However, the nature of features required for all energy performance calculation software is consistent.

Currently, the methodology implements the feature selection procedure based on engineering judgment or existing literature to predict the rating. Future work will consider the hybrid approach by using sensitivity analysis and data-driven feature selection methods

with the current proposed methodology to achieve more robust prediction results.

Acknowledgment

This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under the SFI Strategic Partnership Programme Grant number SFI/15/SPP/E3125. The opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Science Foundation Ireland.

Nomenclature

TP	number of true positives
FP	number of false positives
FN	number of false negatives
P	number of positives in ground truth
N	number of negatives in ground truth
K	training or testing samples
c_i	predicted value
$\bar{c}_i$	target value

References

Ali, U., C. Buccella, and C. Cecati (2016). Households electricity consumption analysis with data mining techniques. In *Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE*, pp. 3966–3971. IEEE.

Ali, U., M. H. Shamsi, F. Alshehri, E. Mangina, and J. ODonnell (2018). Comparative analysis of machine learning algorithms for building archetypes development inurban building energy modeling. In *Building Performance Modeling Conference and SimBuild*.

Amasyali, K. and N. M. El-Gohary (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews 81*, 1192–1205.

Breunig, M. M., H.-P. Kriegel, R. T. Ng, and J. Sander (2000). Lof: identifying density-based local outliers. In *ACM sigmod record*, Volume 29, pp. 93–104. ACM.

Collins, M. and J. Curtis (2018). Bunching of residential building energy performance certificates at threshold values. *Applied Energy 211*, 662–676.

Crawley, D. B., L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, et al. (2001). Energyplus: creating a new-generation building energy simulation program. *Energy and buildings 33*(4), 319–331.

Deb, C., F. Zhang, J. Yang, S. E. Lee, and K. W. Shah (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews 74*, 902–924.

EESI (2018). Building and climate change by environmental and energy study institute. `http://www.eesi.org/`. [Online; accessed 01-February-2019].

Egan, J., D. Finn, P. H. D. Soares, V. A. R. Baumann, R. Aghamolaei, P. Beagon, O. Neu, F. Pallonetto, and J. ODonnell (2018). Definition of a useful minimal-set of accurately-specified input data for building energy performance simulation. *Energy and Buildings 165*, 172–183.

EU (2018). Directive (eu) 2018/844 of the european parliament and of the council of 30 may 2018 amending directive 2010/31/eu on the energy performance of buildings and directive 2012/27/eu on energy efficiency. *Official Journal of the European Union 61*.

EU-Energy (2018). Energy for europe by european commission. `https://ec.europa.eu/energy/en`. [Online; accessed 01-February-2019].

Famuyibo, A. A., A. Duffy, and P. Strachan (2012). Developing archetypes for domestic dwellingsan irish case study. *Energy and Buildings 50*, 150–157.

Fan, C., F. Xiao, and Y. Zhao (2017). A short-term building cooling load prediction method using deep learning algorithms. *Applied Energy 195*, 222–233.

Feng, C., M. Cui, B.-M. Hodge, and J. Zhang (2017). A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. *Applied Energy 190*, 1245–1257.

Fokaides, P. A., K. Polycarpou, and S. Kalogirou (2017). The impact of the implementation of the european energy performance of buildings directive on the european building stock: The case of the cyprus land development corporation. *Energy Policy 111*, 1–8.

Gao, X. and A. Malkawi (2014). A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm. *Energy and Buildings 84*, 607–616.

Jones, P., S. Lannon, and J. Williams (2001). Modelling building energy use at urban scale. In *7th International IBSPA conference, Rio de Janeiro, Brazil, August*, pp. 1683–1697.

Kapetanakis, D.-S., E. Mangina, E. H. Ridouane, K. Kouramas, and D. Finn (2015). Selection of input variables for a thermal load prediction model. *Energy Procedia 78*, 3001–3006.

Kapetanakis D., Mangina E., F. D. (2017). Input variable selection for thermal load predictive models of commercial buildings. *Energy and Buildings 137*, 13–26.

Kontokosta, C. E. and C. Tull (2017). A data-driven predictive model of city-scale energy use in buildings. *Applied energy 197*, 303–317.

Li, X. and J. Wen (2014). Review of building energy modeling for control and operation. *Renewable and Sustainable Energy Reviews 37*, 517–537.

Molina-Solana, M., M. Ros, M. D. Ruiz, J. Gómez-Romero, and M. J. Martín-Bautista (2017). Data science for building energy management: A review. *Renewable and Sustainable Energy Reviews 70*, 598–609.

Picard, R. R. and K. N. Berk (1990). Data splitting. *The American Statistician 44*(2), 140–147.

Rahman, A., V. Srikumar, and A. D. Smith (2018). Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks. *Applied Energy 212*, 372–385.

SEAI (2018). Energy in the residential sector report. `https://www.seai.ie`. [Online; accessed 01-February-2019].

Tang, J., Z. Chen, A. W.-c. Fu, and D. Cheung (2001). A robust outlier detection scheme for large data sets. In *In 6th Pacific-Asia Conf. on Knowledge Discovery and Data Mining*. Citeseer.

Trnsys, A. (2000). Transient system simulation program.

Wei, Y., X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, and X. Zhao (2018). A review of data-driven approaches for prediction and classification of building energy consumption. *Renewable and Sustainable Energy Reviews 82*, 1027–1047.

Yildiz, B., J. I. Bilbao, and A. B. Sproul (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews 73*, 1104–1122.

Zhao, H.-x. and F. Magoulès (2012). A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews 16*(6), 3586–3592.