

Zero Shot Learning for Predicting Energy Usage of Buildings in Sustainable Design

Arun Zachariah¹, Praveen Rao¹, Brian Corn², Dominique Davison²

¹ University of Missouri-Columbia, Columbia, Missouri, USA

² PlanIT Impact, Kansas City, Missouri, USA

azachariah@mail.missouri.edu, praveen.rao@missouri.edu, bcorn@planitimpact.com, ddavison@planitimpact.com

Abstract

The 2030 Challenge is aimed at making all new buildings and major renovations carbon neutral by 2030. One of the potential solutions to meet this challenge is through innovative sustainable design strategies. For developing such strategies it is important to understand how the various building factors contribute to energy usage of a building, right at design time. The growth of artificial intelligence (AI) in recent years provides an unprecedented opportunity to advance sustainable design by learning complex relationships between building factors from available data. However, rich training datasets are needed for AI-based solutions to achieve good prediction accuracy. Unfortunately, obtaining training datasets are time consuming and expensive in many real-world applications. Motivated by these reasons, we address the problem of accurately predicting the energy usage of new or unknown building types, i.e., those building types that do not have any training data. We propose a novel approach based on zero-shot learning (ZSL) to solve this problem. Our approach uses side information from building energy modeling experts to predict the closest building types for a given new/unknown building type. We then obtain the predicted energy usage for the k -closest building types using the models learned during training and combine the predicted values using a weighted averaging function. We evaluated our approach on a dataset containing five building types generated using BuildSimHub, a popular platform for building energy modeling. Our approach achieved better average accuracy than a regression model (based on XGBoost) trained on the entire dataset of known building types.

Introduction

In 2020, the U.S. Energy Information Administration reported that buildings amounted to 40% of the total U.S. energy consumption.¹ In addition, residential and non-residential buildings contribute significantly to global energy-related CO₂ emissions.² Sustainable housing design can eliminate green house gas emissions. For this reason, The 2030 Challenge³ was issued by Architecture 2030. The goal of this challenge is to ensure that new buildings and major renovations shall be carbon neutral by 2030. Sustainable

building design is set to become a standard for the 21st century home⁴. Predicting the energy consumption of a building is necessary for energy planning, management, and conservation. Modeling the energy usage of a building is a complex problem since there are hundreds of factors/features (e.g. area, location, elevation, occupancy, orientation) that must be considered. Moreover, the structural characteristics of a building also impacts the energy usage further adding to its complexity.

The growth of AI in recent years provides an unprecedented opportunity to advance sustainable design by learning complex relationships between building factors from available real-world data. However, rich training data is needed for AI-based solutions to achieve good prediction accuracy. Unfortunately, obtaining training data is time consuming and expensive in many real-world applications. Motivated by these reasons, we address *the problem of predicting the energy usage of new building types*, for which no training data exists. We build on the success of ZSL for image classification and adapt it for our purpose. The main contributions of our work are as follows:

- We propose a new approach using ZSL to predict the energy usage of an unknown building type by first identifying the k -closest building types.
- We then compute a weighted average of the predictions from the known types that were used during training. Our approach can leverage side information required for ZSL from either a building energy modeling expert or using dimensionality reduction techniques.
- We performed an evaluation of our approach and observed that it achieves better accuracy than a single regression XGBoost model that was trained on the data of known building types. (This model ignores the building type during training as it is unknown during prediction.)

Related Work and Motivation

Building Energy Usage Prediction

Techniques for predicting building energy usage can be broadly categorized as engineering methods and data-driven approaches (Zhao and Magoulès 2012). Traditional engineering methods primarily use physics-based principles to

Copyright © 2022, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹ <https://www.eia.gov/totalenergy/data/monthly/>

² <https://globalabc.org> ³ https://architecture2030.org/2030_challenges/2030-challenge

⁴ <https://fontanarchitecture.com/sustainable-house-design-21-ideas>

predict energy consumption. They take into account building construction, operation, equipment usage, and environmental information to calculate precise energy consumption. Zhao et al. (Zhao and Magoulès 2012) provide a nice survey of common techniques for energy prediction of buildings. Numerous software tools have been developed over the years for building energy prediction such as EnergyPlus⁵, OpenStudio⁶, and BuildSim⁷.

Data-driven approaches can be further classified into statistical techniques and AI-based methods. Statistical techniques try to correlate the energy consumption with the parameters involved in building design using historic data. Bauer et al. (Bauer and Scartezzini 1998) used correlation for understanding both heating and cooling loads simultaneously. Dhar et al. (Dhar, Reddy, and Claridge 1998) used Fourier series to model hourly energy usage in commercial buildings. With advances in AI, artificial neural networks (ANNs) have been used to model non-linear relationships between variables for predicting energy consumption (Li et al. 2019). Support vector machines (SVMs) were more accurate when limited training data was available (Zhao and Magoulès 2012; Li et al. 2009). Gaussian process regression was used due to its ability to quantify and handle different sources of uncertainty in the data (Seyedzadeh et al. 2018). Data clustering, aimed at discovering natural grouping in data, was used in energy consumption analysis to sample representative building types (Lara et al. 2015; Pieri, Tzouvardakis, and Santamouris 2015). More recently, XGBoost has been used in numerous applications for predicting and modeling building energy performance (Chakraborty and Elzarka 2019b). It performs ensemble learning, combining the predictions from multiple base learners, enabling it to model complex relationships between attributes in the data. Chakraborty et al. (Chakraborty and Elzarka 2019a) showed that XGBoost performed better than ANN-based energy models. Thus, there is growing interest in using AI-based methods for building energy prediction.

Zero-Shot Learning

In recent years, ZSL (Xian et al. 2018) has attracted much attention in image classification for predicting classes that were not seen during training. ZSL requires some form of side information, which is used to share information between classes so that the knowledge from the known classes (seen during training) is transferred to unseen classes. Side information can be represented by attributes of classes or embeddings (i.e., low-dimensional representation of high-dimensional feature vectors) in a continuous space. ZSL is very useful in real-world applications when there is lack of training data for certain classes.

Romera-Paredes et al. (Romera-Paredes and Torr 2015) developed a simple ZSL approach, which is of particular interest to us. Essentially, a linear model is learned on the training instances to compute a matrix W . Using W and a matrix of signatures S of known classes, a matrix V is computed s.t. $W = V \times S$. At inference time, a new matrix of signatures S' (from the unknown/test classes) is used to

compute $W' = V \times S'$. This new linear model W' is used to obtain the predicted classes on the test data.

Motivation

The aforementioned techniques proposed to predict building energy usage perform well when provided with precise and detailed inputs describing the numerous factors involved in designing the building. Unfortunately, many of these factors are hard to obtain and audit. Furthermore, the success of AI-based techniques depends on high quality training data, which may be time consuming/expensive to produce. For example, running an energy model for a specific building (e.g., using BuildSimHub) can cost thousands of dollars. We hypothesize that ZSL is a promising approach for predicting the energy usage of new/unknown building types. However, modeling the problem of energy prediction using ZSL is non-trivial. Appropriate side information is needed for ZSL to be successful.

Our Approach

In this section, we describe our ZSL approach for predicting the energy usage of new or unknown building types, i.e., those building types that do not have any training data.

Our approach draws inspiration from the ZSL approach proposed by Romera-Paredes et al. (Romera-Paredes and Torr 2015). For side information, let us assume there exists a matrix \mathbb{S} , which is handcrafted by domain experts containing information for all the building types $B = \{b_1, b_2, \dots\}$ based on several building energy factors/parameters $P = \{p_1, p_2, \dots\}$. Note that \mathbb{S} is a $|P| \times |B|$ matrix. We will assume that some of the building types in B are treated as new/unknown. Let $B_u \subseteq B$ denote these unknown building types. Let X denote the feature vector of the training data instances of known building types $B - B_u$, and Y denote the target vector that represents the building type for each training data instance.

During training, we first learn W (e.g., using logistic regression) given X and Y . Given a set of known building types $B - B_u$, we construct S from \mathbb{S} by dropping the columns corresponding to building types in B_u . We then compute V using S and W . For each known building type, we use the training data instances for that building type and construct a predictive model for the building energy usage metrics. This model is then used during inference/energy prediction. Algorithm 1 summarizes the training steps.

Algorithm 1: Training the models

Input: X : Feature vector of training data; Y target vector containing the building type; S : Side information for ZSL; B_u : Unknown building types

- 1: Learn W using logistic regression given X and Y
- 2: Construct S by dropping columns from \mathbb{S} using B_u
- 3: Compute V s.t. $W \leftarrow V \times S$
- 4: **for** $b_i \in B - B_u$ **do**
- 5: Learn predictive models using X_{b_i} for the building energy usage metrics
- 6: **end for**

⁵ <https://energyplus.net/>

⁶ <https://www.openstudio.net/>

⁷ <https://www.buildsim.io>

During inference, our goal is to predict the energy usage of B_u . (For simplicity, we will assume only one building type in B_u is used during inference.) Our key idea is to first find the k -closest known building types to the given unknown building type $b_i \in B_u$. For computing this efficiently, we propose the following steps, summarized in Algorithm 2: Construct S' from \mathbb{S} by dropping all columns in \mathbb{S} for unknown building types $B_u - \{b_i\}$ (Line 1). Compute W' using S' and V (Line 2). Let X' denote the feature vector of the test data instances for b_i . We then compute Y' (Line 3). Note that Y' is a matrix of real-valued numbers, wherein each row of Y' corresponds to a test data instance. The columns of Y' correspond to the known building types and b_i . Thus, we have a score that associates each test data instance with a building type. We ignore the score for b_i and sort the remaining scores (high to low) to obtain the k -closest building types.

Algorithm 2: Energy usage prediction using ZSL

Input: X' : Feature vector of test data; Y building type for each training data instance; S : Side information from ZSL; b_i : Unknown building type; k : number of closest building types to consider

Output: P : Predicted energy metrics for b_i

- 1: Construct S' by dropping columns from \mathbb{S} corresponding to $B_u - \{b_i\}$
 - 2: $W' \leftarrow V \times S'$
 - 3: $Y' \leftarrow X' \times W'$
 - 4: **for** t in X' **do**
 - 5: Let (s_1, s_2, \dots, s_k) denote the sorted scores from Y' for t after ignoring the score for b_i
 - 6: Let (e_1, e_2, \dots, e_k) denote the predicted values (given t) for the k closest building types using the trained regression models
 - 7: Let P denote weighted average of (e_1, \dots, e_k) using (s_1, \dots, s_k)
 - 8: **end for**
 - 9: **return** P
-

For a test data instance with feature vector t , let (s_1, s_2, \dots, s_k) denote the scores for the k -closest building types, where $k > 0$. We use the predictive models for each of the k -closest building types to predict a building energy metric of interest. Let (e_1, e_2, \dots, e_k) denote the predicted values for the unknown building b_i . We compute the weighted average of the predicted values, where the weights are computed by either applying *softmax* or any normalization technique on the scores. These steps are shown in Lines 5-8 of Algorithm 2.

If \mathbb{S} is not available from a domain expert, we could employ singular value decomposition (SVD) on simulation data for a given building type. (Tools like BuildSimHub could be used to generate simulation data.) The singular values obtained from SVD can now be used to represent a column of \mathbb{S} for that building type. In summary, robust side information is required for our ZSL approach to succeed.

Experimental Evaluation

Dataset and Predicted Energy Metrics

For our evaluation, we used BuildSimHub⁸ to generate a dataset by using the Monte Carlo method through repeated sampling of the various factors/parameters involved in the design of a building type. The simulations were done for five building types, namely, *Educational* (ED), *MixedUse* (MU), *Office* (OF), *Retail-Strandalone* (RS), and *Retail-Stripmall* (RL). The dataset had 69 features that contained both categorical and continuous variables. The total instances/records for each building type are shown in Table 1.

We predicted three metrics of interest for a given building type. These included *Total Site Gas Energy Usage Intensity* (TGAS), *Cooling Electricity Demand* (COOL) and *Facility Peak Electricity Demand* (PFAC). TGAS denotes the total gas usage of a building per square foot per year. COOL denotes the amount of energy required for cooling in a given duration of time. PFAC denotes the highest amount of electricity consumed in a given duration of time.

Setup and Implementation

All the experiments were conducted on a machine in Cloud-Lab (Duplyakin et al. 2019). We used Python (v3.6), NumPy (v1.19.5), and Pandas (v1.1.5). We also used the Python package for XGBoost⁹ (v1.0.1) and scikit-learn (v0.24.2) for implementing logistic regression (in order to compute the regression coefficients required by our ZSL approach). For normalization, we used the softmax function, included with SciPy (v1.5.4), as it gave us the best results.

Evaluation

We compared our ZSL-based approach, referred hereinafter as *ZSL*, with a baseline model (Baseline) that used XGBoost. The dataset for each building type was split randomly into train and test instances. (The train-test ratio was set at 9:1.) In order to compare the accuracy of *ZSL* with Baseline, we performed the following steps: (1) We selected one building type as unknown (e.g., ED) and assumed the remaining were known building types for training (e.g., MU, OF, RS, and RL). (2) A predictive model was trained on the training sets of the known building types. (3) For the unknown building type, we used each data record in the test set that denotes specific building parameter values, and predicted the three metrics, namely, TGAS, COOL, and PFAC. (4) Finally, we computed the accuracy of the predictions for the three metrics of the unknown building type. We report the average accuracy (%) across all the test instances of the unknown building type. We repeated the above steps for all the five building types.

For Baseline, the predictive model was trained using XGBoost. During training, the building types of the known buildings were ignored as the building type of the test data was assumed to be unknown. We performed hyperparameter tuning (on maximum tree depth and rate of learning) and k fold cross validation by splitting the training instances into 5 folds. (We observed that our XGBoost models

⁸ <https://www.buildsim.io> ⁹ <https://github.com/dmlc/xgboost>

Unknown building type	Total # of records	TGAS			COOL			PFAC			Avg. Accuracy		
		Base line (%)	ZSL _d (%)	ZSL _s (%)	Base line (%)	ZSL _d (%)	ZSL _s (%)	Base line (%)	ZSL _d (%)	ZSL _s (%)	Base line (%)	ZSL _d (%)	ZSL _s (%)
ED	13,645	56.05	68.68	68.11	85.71	81.41	81.41	83.61	88.00	88.00	75.13	79.37	79.17
MU	25,801	71.06	78.24	78.25	55.58	59.55	59.55	74.70	78.06	78.46	67.11	71.95	72.09
OF	13,519	84.88	88.12	88.12	63.69	84.55	84.55	82.18	78.87	78.87	76.92	83.85	83.85
RS	13,428	91.31	92.74	92.74	85.82	87.66	87.66	87.85	94.30	94.30	88.33	91.57	91.57
RL	13,816	87.56	85.56	85.56	77.15	74.76	77.14	80.51	78.73	83.99	81.74	79.69	82.23

Table 1: Evaluation results (best results shown in bold)

outperformed neural networks for regression on our dataset. Hence, we do not use neural networks for comparison.)

For ZSL , the predictive model composed of 4 XGBoost models trained separately on each known building type as we first identify the closest building types during prediction on an unknown building type. We evaluated two variations of ZSL based on how the signature matrix was constructed. The first one (denoted by ZSL_d) used the signature matrix provided by a domain expert. The second one (denoted by ZSL_s) used the signature matrix generated by applying SVD on the simulated data of a building type.

The evaluation results are shown in Table 1. Out of 15 cases, ZSL_d or ZSL_s achieved the best accuracy for 12 cases compared to Baseline. (The best results for ZSL_d and ZSL_s were obtained for $k = 4$.) We also computed the average of the accuracy (%) for the three metrics for each unknown building type. Once again, our ZSL-based approaches was *the winner in all cases*. These results demonstrate the benefit of using ZSL for predicting the energy usage of unknown building types. Interestingly, the signature matrices used by ZSL_d and ZSL_s , though independently generated, were both very useful side-information for prediction.

Conclusion

We proposed a novel approach based on ZSL for predicting the energy usage of an unknown building type. Our approach predicts the k -closest known building types by using side information from building energy modeling experts. The predicted energy usage of a new/unknown building type is the weighted average of the predictions from these k -closest building types. We evaluated our approach on a simulated dataset created using BuildSimHub for five building types to predict three different energy usage metrics. Our approach achieved better average accuracy compared to XGBoost-based models trained without taking into account the building type. Our work aims to advance the field of sustainable design using AI-based techniques and positively impact multiple engineering disciplines.

Acknowledgments: This work was supported by the National Science Foundation under Grant No. 1747751.

References

Bauer, M.; and Scartezzini, J.-L. 1998. A Simplified Correlation Method Accounting for Heating and Cooling Loads in Energy-

Efficient Buildings. *Energy and Buildings*, 27(2): 147–154.

Chakraborty, D.; and Elzarka, H. 2019a. Advanced Machine Learning Techniques for Building Performance Simulation: A Comparative Analysis. *Journal of Building Performance Simulation*, 12(2): 193–207.

Chakraborty, D.; and Elzarka, H. 2019b. Early Detection of Faults in HVAC Systems Using an XGBoost Model with a Dynamic Threshold. *Energy and Buildings*, 185.

Dhar, A.; Reddy, T. A.; and Claridge, D. E. 1998. Modeling Hourly Energy Use in Commercial Buildings with Fourier Series Functional Forms. *Journal of Solar Energy Engineering*, 120(3): 217–223.

Duplyakin, D.; Ricci, R.; Maricq, A.; Wong, G.; Duerig, J.; Eide, E.; Stoller, L.; Hibler, M.; Johnson, D.; Webb, K.; Akella, A.; Wang, K.; Ricart, G.; Landweber, L.; Elliott, C.; Zink, M.; Cecchet, E.; Kar, S.; and Mishra, P. 2019. The Design and Operation of CloudLab. In *2019 USENIX Annual Technical Conference*, 1–14.

Lara, R. A.; Pernigotto, G.; Cappelletti, F.; Romagnoni, P.; and Gasparella, A. 2015. Energy Audit of Schools by Means of Cluster Analysis. *Energy and Buildings*, 95: 160–171.

Li, Q.; Meng, Q.; Cai, J.; Yoshino, H.; and Mochida, A. 2009. Applying Support Vector Machine to Predict Hourly Cooling Load in the Building. *Applied Energy*, 86(10): 2249–2256.

Li, Z.; Dai, J.; Chen, H.; and Lin, B. 2019. An ANN-Based Fast Building Energy Consumption Prediction Method for Complex Architectural Form at the Early Design Stage. *Building Simulation*, 12: 665–681.

Pieri, S. P.; Tzouavadakis, I.; and Santamouris, M. 2015. Identifying Energy Consumption Patterns in the Attica Hotel Sector Using Cluster Analysis Techniques with the Aim of Reducing Hotels’ CO2 Footprint. *Energy and Buildings*, 94: 252–262.

Romera-Paredes, B.; and Torr, P. 2015. An Embarrassingly Simple Approach to Zero-Shot Learning. In *Proc. of the 32nd International Conference on International Conference on Machine Learning*, volume 37, 2152–2161. JMLR.org.

Seyedzadeh, S.; Rahimian, F. P.; Glesk, I.; and Roper, M. 2018. Machine Learning for Estimation of Building Energy Consumption and Performance: A Review. *Visualization in Engineering*, 6(86): 2579–2605.

Xian, Y.; Lampert, C. H.; Schiele, B.; and Akata, Z. 2018. Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly. *IEEE transactions on pattern analysis and machine intelligence*, 41(9): 2251–2265.

Zhao, H.-X.; and Magoulès, F. 2012. A Review on the Prediction of Building Energy Consumption. *Renewable and Sustainable Energy Reviews*, 16(6): 3586–3592.