Energy Consumption Prediction of Residential Buildings using Machine Learning: A Study on Energy Benchmarking Datasets of Selected Cities across the United States

Milad Parvaneh[1, 2], Abolfazl Seyrfar[1], Ali Movahedi[1], Hossein Ataei[1], Khuong Le Nguyen[3], and Sybil Derrible[1]

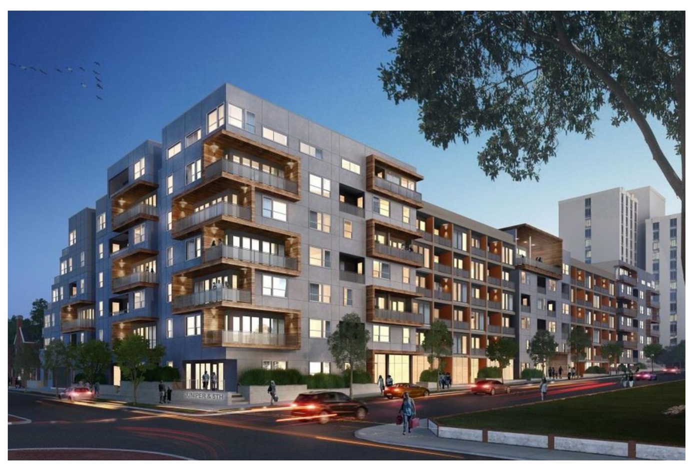
1 University of Illinois at Chicago, Chicago, IL, USA
2 Stanford Center for Professional Development, Stanford, CA, USA
3 University of Transport Technology, Hanoi, Vietnam

Abstract

- A significant portion of energy consumption around the globe occurs in ***residential buildings***.
- Many cities across the U.S. have mandatory energy ***benchmarking programs*** requiring large buildings to track and report their energy use.
- Our goal is to evaluate and compare the ***datasets*** provided by the benchmarking programs.
- We employ ***Extreme Gradient Boosting***, ***Random Forest***, and ***Artificial Neural Network*** as three common Machine Learning methods to predict building energy use in eight U.S. metropolitan areas.
- Suggestions are provided to enhance the datasets and further improve building energy use research.

Introduction

- Residential buildings accounted for 16% of the total energy use in the United States in 2019 [1].
- Many cities across the U.S. has started energy benchmarking programs requiring owners of non-residential and large multifamily buildings to track their energy consumption [2].
- Most of these data are openly available to the public in the form of energy use datasets.
- Our goal is to utilize these publicly available datasets to analyze and compare well-established ML algorithms' performance in predicting energy consumption.
- The ML algorithms are ***Extreme Gradient Boosting (XGBoost)***, ***Random Forest (RF)***, and ***Artificial Neural Network (ANN)***.
- We then investigate how accurate a prediction model we can train using those data.
- Our learning algorithms' primary inputs are ***Gross Floor Area***, ***Year Built***, and ***Energy Star Score***. We also use other features where available.
- We use three ML methods to output ***Energy Use Intensity (EUI)***.



Methodology

- Three steps are taken to clean the dataset:
  1. First, unwanted variables are filtered out so we are only left with selected features and target variable.
  2. Second, data points corresponding to buildings of *Multifamily Housing* type are selected.
  3. Finally, data points with EUI greater than 3154.5 KWh/m^2 (equivalent to 1000 kBtu/ft^2) are considered as outliers and eliminated.
- Dataset for each city is split into two sets:
  1. One training/validation set consisting of 80% of the data points used for training the models.
  2. A test set with the remaining 20% used for model evaluation.
- Three ML methods are employed to train prediction models: ***XGBoost***, ***RF***, and ***ANN***.
- Hyper-parameter tuning for each model is performed using grid search.
- For model evaluation, we use ***Mean Absolute Percentage Error (MAPE)***, ***Mean Squared Error (MSE)***, and ***R^2***.
- As an example, Table 1 shows the number of data points for residential buildings and the available features within the dataset for the City of Chicago.

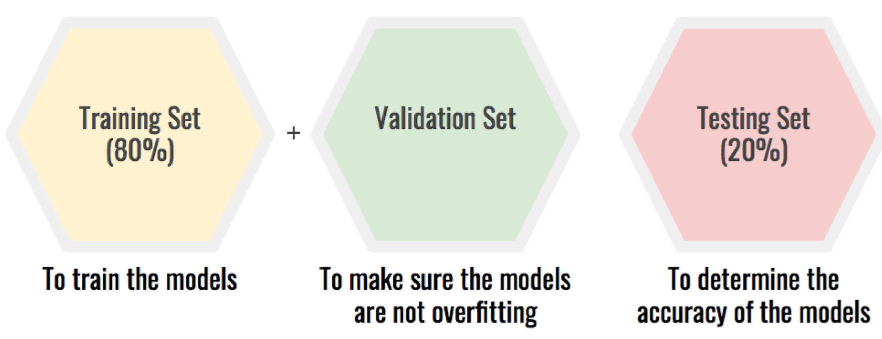


Table 1. Energy benchmarking datasets for residential buildings.

City, Dataset Reference	Features	Mean $\pm$ SD	# of Data Points
Chicago, IL	Gross Floor Area (m^2) Energy Star Score Year Built	19333.6 $\pm$ 19375.6 58.5 $\pm$ 29.4 1965.1 $\pm$ 33.7	1332

Results

- Figure 1 shows the scatter plots of the measured target variable versus the predicted values for the three ML techniques used in this study.
- The solid pink lines at 45 degrees depict the zero error line, where measured and predicted EUIs would be equal.
- The dashed pink lines represent $\pm$25% error threshold.

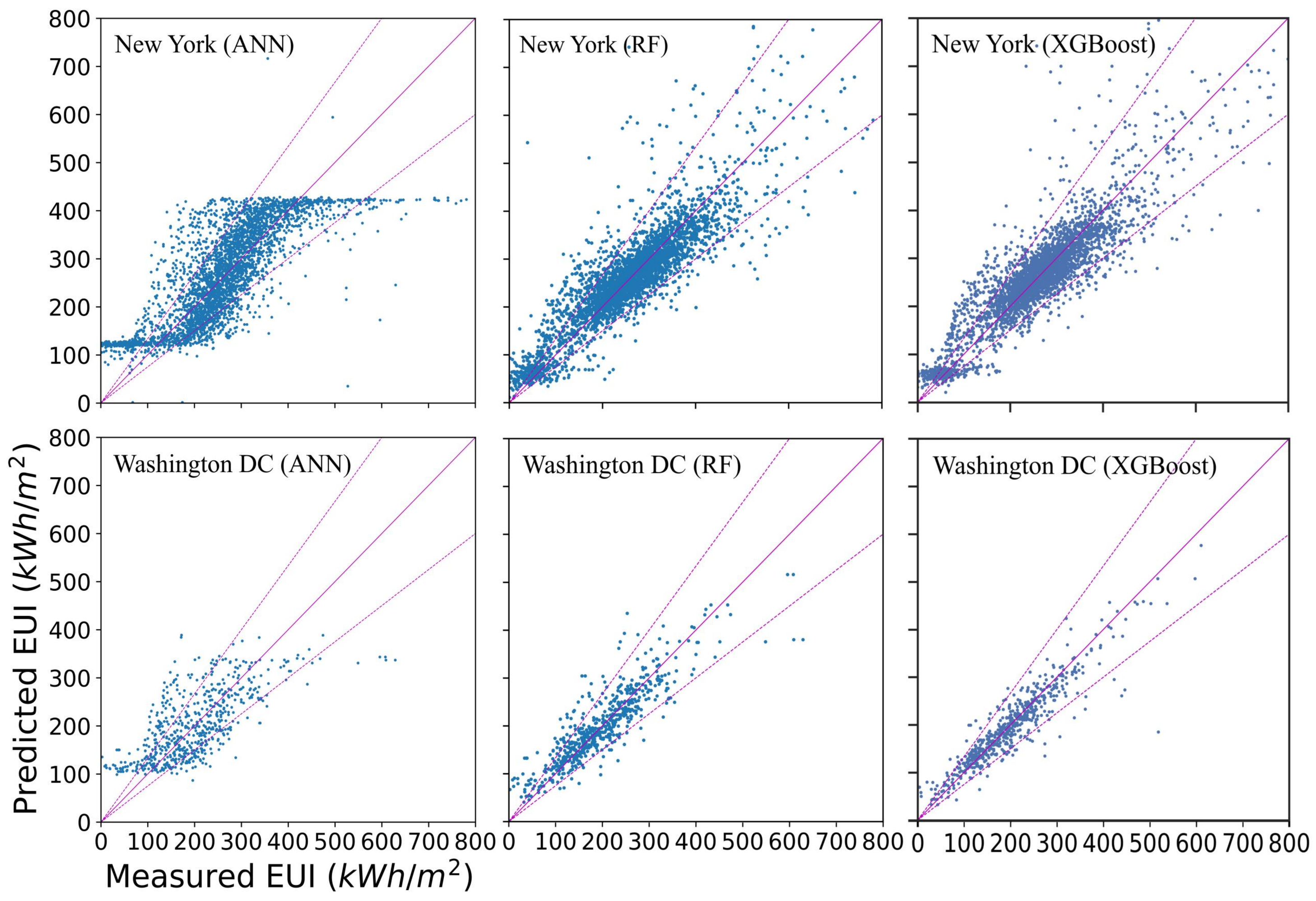


Figure 1. Scatter plot of measured vs. predicted EUI for New York and Washington D.C.

Conclusion

- From a qualitative point of view, by comparing the scatter plots, we see that XGBoost and RF perform better (see Fig. 1 above).
- Figure 2 plots Mean CV R^2 (dashed lines) and R^2 (solid lines) for the three ML methods.
- The better performance of XGBoost is evident.
- Based on our observations, ***Gross Floor Area***, ***Energy Star Score***, and ***Year Built*** are the main contributors in building energy use prediction.
- These features are advised to be requested from building owners and provided in benchmarking datasets.
- As for other features, we could not make definitive remarks, although more data is generally preferable.

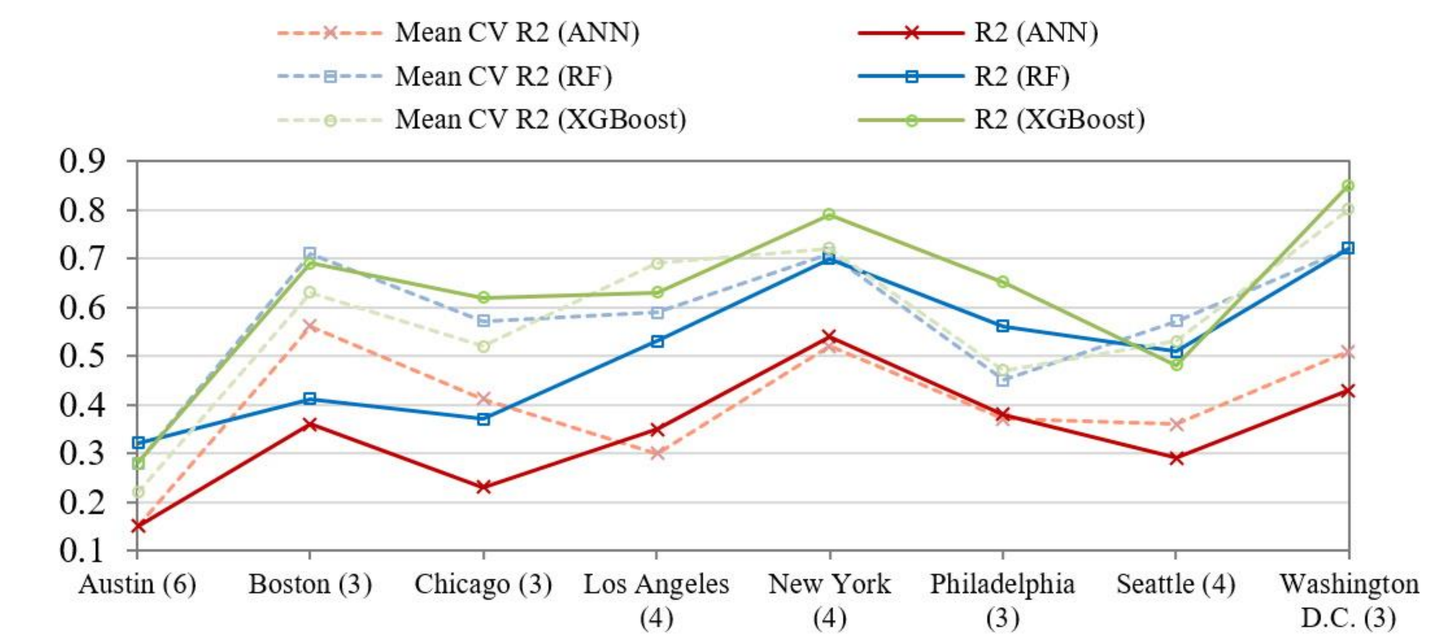


Figure 2. Mean CV R^2 (dashed lines) and test set R^2 (solid lines) for the three ML methods.

References

[1] The U.S. Energy Information Administration (EIA), https://www.eia.gov/totalenergy/data/monthly/pdf/flow/css_2019_energy.pdf, last accessed 2021/03/24.

[2] Seyrfar, A., Ataei, H., Derrible, S.: A Review of Building Energy Benchmarking Policies Across the US Cities. In: Energy Proceedings, 8 (2020).