Predicting Consumption of The Buildings using Machine Learning Technique (Survey)

Akhila C.V

VI semester, M.Tech, Department of cse
GSSS, Institute of Engineering & Technology for Women. Mysuru, Karnataka, India

Abstract- Ever growing inhabitants and progressive municipal business demands for constructing new buildings are known as the primary contributor to greenhouse gasses. Therefore, improvement of energy efficiency of the building segment has become a vital target to decrease the amount of gas production as well as fossil fuel consumption. One most efficient approach to reducing CO2 emission and energy consumption with regard to new buildings is to think about energy competence at a very early design stage. On the other hand, efficient energy administration and smart refurbishments can improve energy recital of the existing stock. All these elucidation entail accurate energy prediction for finest decision making. In recent years, artificial intelligence (AI) in general and machine learning (ML) techniques in explicit terms has been projected for forecasting of building energy consumption and performance. The existing models have achieved roughly 74% of the accuracy on the test set. Received results can be further concentrating on with artificial intelligence methods.

Index Terms - Prediction, finest decision, Multi-Task Learning, Artificial Neural Network, Long Short-Term Memory.

1. INTRODUCTION

An examination of the NYC convention dataset which measures 60 variables correlated to energy use for over 11,000 buildings. Of main concern is the Energy Star Score, which is often used as cumulative measure of the overall effectiveness of a building. The Energy Star Score is a percentile measure of a building's energy performance calculated from self-reported energy usage. Therefore, it is very important to scrutinize this data in order to extract some useful in sequence and to enlarge an algorithm based on this analysis. This can be accomplished through data mining and machine learning. Machine learning is an essential part of artificial intelligence, which is used to plan algorithms based on the data trends and chronological relationships between data. Machine learning is used in assorted fields such as bioinformatics, intrusion detection, Information retrieval, game playing, marketing, malware detection, image intricacy and so forth. There are a load of machine learning models to pick from and fix on where to start can be intimidating.

Walk through solving a entire machine learning problem via a real-world dataset. Machine learning is still a turf driven primarily by pragmatic (experimental) rather than theoretical results, and it's almost impractical to know ahead of time which model will do the best. The building energy data can predict a building's Energy Star score, and then construe the results to find the variables that are most projecting of the score. This is a supervised, regression machine learning task: given a set of data with objectives (in this case the score) included, we want to train a model that can learn to record the features (also known as the explanatory variables) to the objective. AI centers around the advancement of PC programs that can get to information and use it learn for themselves. Here is an outline that demonstrates the various groupings of AI:
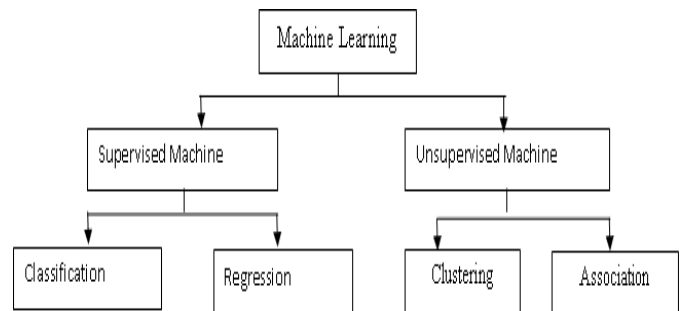

Figure 1.1 chart that shows the different groupings of machine learning

Some machine learning methods Machine learning algorithms are often categorized as supervised or unsupervised.

•Supervised machine learning algorithms can relate what has been educated in the history to new data using tagged examples to predict future events. Preliminary from the analysis of a recognized training dataset, the learning algorithm fabricates an inferred function to formulate predictions about the yielded values. The system is able to present targets for any new input after adequate training. The learning algorithm can also contrast its output with the correct, intended output and find errors in order to amend the model accordingly.

•In contrast, unsupervised machine learning algorithms are used at what time the information used to train is neither classified nor labeled. Unsupervised learning revises how systems can gather a function to portray a concealed structure from unlabeled data. The system doesn't stature out the right output, but it explores the data and can draw

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCRACES - 2019 Conference Proceedings

reckoning from datasets to describe hidden structures from unlabeled data.

•Semi-supervised machine learning algorithms plunge somewhere in between supervised and unsupervised learning, since they employ both labeled and unlabeled data for training – typically a small quantity of labeled data and a huge amount of unlabeled data. The systems that use this technique are able to significantly improve learning exactness. Usually, semi-supervised learning is chosen when the acquired labeled data requires expert and relevant resources in order to train it / learn from it. If not, gaining unlabeled data generally doesn't require supplementary resources.

•Reinforcement machine learning algorithms is learning methods that interact with its environment by producing events and discover errors or rewards. Assessment and error search and deferred reward are the mainly important characteristics of reinforcement learning. This method permits machines and software agents to robotically determine the superlative behavior within a specific context in order to exploit its performance.

•Regression and classification are classified under the identical umbrella of supervised machine learning. Both share the similar perception of utilizing known datasets (referred to as training datasets) to make predictions. In supervised learning, an algorithm is engaged to learn the mapping function from the input variable (x) to the output variable (y); that is $y = f(X)$. The purpose of such a problem is to estimate the mapping function (f) as precisely as possible such that at any time there is a new input data (x), the output variable (y) for the dataset can be predicted. The main difference among them is that the output variable in regression is numerical (or continuous) while that for classification is categorical (or discrete).

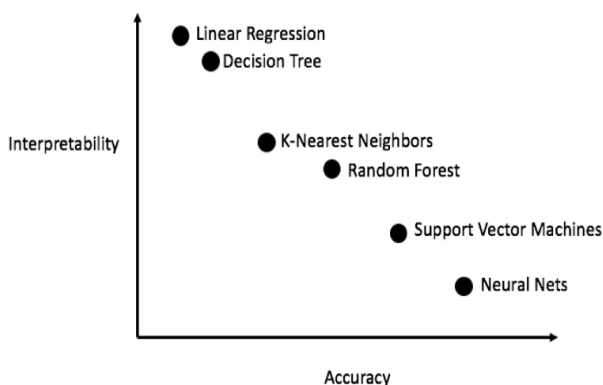The model executes the five unique models covering the multifaceted nature range:

Figure 1.2 some of the machine learning algorithm implemented in the model

- •   Linear Regression
- •   K-Nearest Neighbors Regression
- •   Random Forest Regression
- •   Gradient Boosted Regression

- •   Support Vector Machine Regression

1.2. How the calculation works

Part one covers data cleaning, exploratory data analysis, feature engineering, and feature selection. Part two implements imputing missing values, implementing and comparing machine learning models, hyper parameter tuning using random search with cross validation, and evaluating a model. In the third portion of the model, the first two methods are specific to ensembles of trees, while the third applies machine learning model. LIME is a relatively new package and symbolizes an exciting step in the ongoing effort to explain machine learning predictions.

Based on the information entered in the catalog about building, such as its location, size, number of occupants, number of PCs, etc.. The algorithm built to approximate score's show the amount energy the building would use if it based on the recital like best performing, the worst performing, and every level in between. The estimation contrasts genuine energy data entered to the estimated to conclude whether building ranks comparative to its peers. All of the calculations are based on basis energy and description for the contact of weather variations, as well as modification in key property use details. To approximate how much energy building would use at each performance level, EPA demeanors statistical examination on the review data. For all type of building for which there is an ENERGY STAR score, EPA goes through a scrupulous progression that involves:

•Ensuring the superiority and amount of the information will support an ENERGY STAR score

•Creating a statistical worsening model that correlates the energy information to the asset use details to classify the key drivers of energy use

•Testing the model subsequently to thousands of buildings in Portfolio Manager

1.3. Feature Importance:

Feature importance's effort to prove the significance of each characteristic, to the task of predicting the goal. The technical details of feature importance's are intricate. In Scikit-Learn, we can dig out the characteristic, importance from any company of tree-based learners. With model as our trained model, we can find the feature importance's using model.

•Feature engineering: The process of taking unprocessed data and extort or creating new features. Taking out conversion of variables, such as a natural log and square root, or one-hot training categorical variables so they can be used in a model. One-hot encode categorical variables (borough and property use type).

•Feature selection: The process of selecting the most significant features in the data. In feature selection, we

eliminate features to help the model simplify better to new data and generate a more interpretable model..

For example, here is a plot of Site EUI vs Weather Normalized Site EUI which have a correlation coefficient of 0.997.

➢ Objectives

•Identify predictors contained by the dataset for the Energy Star Score

•Build Regression/Classification models that can forecast the Energy Star Score of a building given the building's energy data

• Deduce the results of the model and utilize the trained model to infer Energy Star Scores of new buildings

➢ Aim

The aims of this module are:

•To pioneer the theory and benefits of energy effectiveness in buildings.

•To give a summary of the methodology used to establish the energy efficiency of buildings.

•To present the dissimilar opportunity and events for reducing energy use in buildings without forfeiting comfort levels.

•To give a summary of legislative and policy utensils that has been flourishing in promoting energy efficiency in buildings.

In this project, we will work with publicly obtainable building energy data from New York City. The aim is to use the energy data to construct a model that can predict the Energy Star Score of a building and infer the results to find the factors which sway the score. The data includes the Energy Star Score, which formulates the model as a supervised regression machine learning task:

- Supervised: we have access to both the features and the target and our goal is to train a model that can learn a mapping among the two

- Regression: The Energy Star score is a uninterrupted variable

The model is expected to be both **accurate** it can predict the Energy Star Score close to the true value and interpretable. The prediction of the model must be understandable. Since, the goal is known; the user can dig the data and build the models.

2.   LITERATURE SURVEY

Literature Survey is one of the grave stages in project life process. It is the depiction of understand scholar papers to comprehend the techniques, methodologies and concepts related to the project. It also concise the inclusion and exclusion of the methodology in the project and their restrictions to discover out problem more clearly and precisely for the future work.

2.1 Related works

In building sector, ANN models have been practical for quick estimation of heating and cooling masses [1][2][3][4] and space heating[5]. Several victorious submission of ANN for Automated Fault Detection and Diagnostics (AFDD) in building energy preservation [5][6], solar water furnace[7] and HVA have been reported. ANN is also implemented in building organization structure to provide mechanical energy utilization control [1][8], Inflation of heating system and placate management.

In 1995, an early study on the functioning of ANN in forecasting of energy utilization using uncomplicated FFN model was executed to forecast electric energy handling of a building in tropical climate based on the residence and temperature data. ANN for short-term assessment of building electricity requirement. Targeting the bio-climatic stockpile, it is exposed that outside temperature and solar waves have a notable force on electricity consumption.

Mihalakakou et al. 2002[16] used FFN and RNN for forecast of frequent electricity energy consumption in a inhabited building located in Athens. The models believed meteorological variables including air, warmth and solar radiation data gathered over six years. Gonzales & Zamarreno 2005[10] predictable short-term voltage energy consumption using a opinion ANN. Effect of the quantity of neurons in hidden layers, the most excellent size of data windows and the ANN parameters on the exactness of the model is explored. Li et al. 2015[15] projected an optimized ANN for prediction of timely electricity consumption using fractional swarm optimization (PSO) algorithm. PCA is used to eliminate redundant input variables obtained from datasets.

Platon et al. 2015[18] applied primary component analysis (PCA) to inspect the pre-input variables of ANN in the prophecy of hourly electricity utilization of an institutional building. Results from contrast of ANN and case-based reasoning (CBR), divulges that the ANN is finer in term of accuracy. However, as CBR offers more lucidity than the ANN and the potentiality to learn from small data, it can be a alternate approach for composite systems reliant on more variables. Li et al. 2015[15] proposed an optimized ANN for forecast of opportune electricity burning up using partial swarm optimization (PSO) algorithm. PCA is used to eliminate redundant attempt variables obtained from two datasets.

Wong et al.[19] used ANN for levying the dynamic energy performance of a commercial building with day-lighting in Hong Kong. Energy Plus software along with algorithms for computation of interior manifestation is applied to produce the building daily energy practice. Nash–Sutcliffe Efficiency Coefficient (NSEC) is used as the most

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCRACES - 2019 Conference Proceedings

important dimension to investigate ANN accuracy in predicting cooling, heating, electric lighting and total electricity consumption.

ANN can be used for determination of restrictions for energy performance assessment of buildings. Buratti et al. 2014 [6] occupy ANN as a tool for estimating energy certificates exactness using 6500 energy labels in Italy. The study investigates a different amalgamation of input variables to diminish the number of training features. Using the result of the ANN, a new catalog is proposed to check the accuracy of affirmed data for energy certificates with a low error of 3.6%.

Khayatian et al. 2016[13] predicts energy recital certificates for residential building using an ANN model and Italian CENED record as training records. A blend set of direct and calculated features are implemented as inputs and heat demand indicators (derived using CENED software) as the production target of ANN.

Ascinoe et al. 2017[2][3] proposed an ANN for evaluation of energy expenditure and inhabitants' thermal console to predict energy performance of the building. Energy assessment of the buildings is achieved using Energy Plus software, and a simulation-based compassion/ ambiguity analysis is proposed for further perfection of network parameters. New buildings and retrofitted stock in presence of energy retrofit measures are considered separately. For the latter case, ANN is employed for optimization of modify parameters. For the first one, three single output ANN is developed to forecast main energy consumption of space heating and cooling and the ratio of yearly uneasiness hours by setting whole-building parameters as association inputs (i.e. geometry, envelope, operation and HVAC). At the same time, Beccali et al. 2017 suggested the use of ANN fast forecasting as a decision support tool for optimizing the retrofit actions of buildings located in Italy.

Kalogirou & Bojic [13] applies RNN to envisage timely energy requirement of a submissive solar building. ZID software has been engaged to calculate the output target. Although results reveal high exactness of estimation, the number of input characteristics (and total training records (40 simulated cases) are not enough. Later in 2001, Kalogitrou (Kalogirou et al. 2001) applies ANN for estimating the daily heat loads of model house buildings with different aspersion of the wall and roof types using a distinctive meteorological data for Cyprus. In this study, TRNSYS software was worn as an energy evaluation engine for all cases and the data validated by contrast of building energy expenditure with the actual measurement. Karatasou et al. 2006 [14] develops an FFN model for timely prediction of energy loads in inhabited buildings. The impact of a diversity of parameters on the precision of a trained model is also investigated, and it is shown that restriction such as humidity and wind speed are less significant and can be exterminated from training features. Furthermore, the function of statistical analysis for augmentation of ANN model and 24 hours in advance

prediction of energy consumption is demonstrated. These methods consist of proposition testing, information criteria and cross-validation in pre-processing and model development. However, there is less enlightenment about the main peculiarity of applied FFN models.

In 2010, Dombayci (Dombayci 2010)[11] used ANN to prediction timely energy consumption of a simple model house based on Turkish standards. The degree-hour scheme is applied to derive the hourly energy utilization to be used in ANN training. The models are suitable for single building energy supervision of simple residential buildings as it does not take many features into account. Kialashaki & Reilsel 2013[14][18] compared an ANN with MLR for opinion of the US family buildings energy demand. Seven sovereign variables (population, gross domestic product, house size, median household income, cost of residential electricity, natural gas and oil) are selected from different data sources (1984–2010) to represent the building characteristics.

Neto & Fiorelli 2008[17][12] compared predicted energy demand of a building in Brazil using ANN model and replication software, EnergyPlus. The research explored the influence using hidden layer showing an insignificant disparity in accuracy of the models. Furthermore, it reveals that exterior temperature is more essential than humidity and solar radiation in estimating energy consumption of the study case. The authors show that ANN is more precise that detailed simulation model, especially in short-term prediction. They conclude that indecent assessment of lighting and tenancy would be the main reason for hesitation in engineering models. Popesco et al.[18] 2009 developed an innovative simulation and ANN-based models for predicting hourly heating energy demand of buildings associated to district heating system. Climate and gathering flow rate variables of prior 24h are used as inputs. Deb et al. 2016[11] also used five previous day's data as ANN model participation to anticipate daily cooling demand of three institutional buildings in Singapore. Paudel et al. 2014[18] used dynamic ANN to anticipate heating energy utilization focal point on building possession profile and equipped short-term heating power level characteristics.

Ben-Nakhi[16] 2014 used a general RNN for forecast of public buildings profile of the successive days using hourly energy consumption data, mean to optimise HVAC thermal energy storage. Data from a community office building in Kuwait constructed from 1997 to 2001 is used for training and testing the ANN model. Energy consumption assessment of buildings is calculated using ESP-r simulation software and bearing in mind climate information, various densities of habitation and orientation uniqueness. The results show that ANN only needs external temperature for precise prediction of cooling loads, whereas imitation software demand for complicated climate detail.

Yokoyama et al. 2009[21] used back-propagation ANN to predict cooling load demand by commenceing a global

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCRACES - 2019 Conference Proceedings

optimisation method for the improvement of network parameters. The effect of the number of hidden layers and the number of neurons in each layer is investigated to optimise the accuracy of the proposed ANN. Azadeh et al.[2] demonstrate the application of ANN based electricity utilization prediction model in the manufacturing industry. The model is used to foresee the annual long-term consumption of industries in Iran using a multilayer perception model. The results compare with the traditional regression model using ANOVA and show superiority for the application

➤ Knowing Your Task and Knowing Your Data

The most important part in the machine learning procedure is indulgent the data we are working with and how it transmits to the task you want to solve. It will not be successful to randomly choose an algorithm and throw the data at it. It is essential to recognize dataset before its processed to begin building a model. Every algorithm is differs in terms of requirement of the data and setting of the goal. While building a machine learning solution certain questions needs to be answered:

- The problem which needs to be addressed. The data collected is relevant to produce appropriate question?
- What is the best technique to phrase my question(s) as a machine learning problem?
- Have I collected adequate data to represent the problem I want to solve?
- What characteristic of the data do I require, and will these enable the right predictions?
- How will I assessment successes in my application?
- How will the machine learning resolution interact with other parts of my research or business product?

In a larger context, the algorithms and methods in machine learning are only one part of a greater procedure to solve a particular problem, and it is good to keep the big picture in mind at all times. Many people spend a lot of time building compound machine learning solutions.

3. SYSTEM DESIGN

The phase of system design involves comprehensive architectural and data flow diagrams for a detailed examination of planned and formulated system. This chapter gives the overview of the proposed system, its design and goal.

3.2 Methodology

➤ High Level Designs

High Level Design recognizes module and interactions between each module. It mainly comprises data flow diagrams and architectural design. Its key focus is to find out the different modules that are mandatory in the design the overall system, and the data flow in each module when

incorporated. The system architecture is shown in Fig 4.1 and according to the shown diagram; the working and components of the proposed system are as follows. The primary apprehension of the proposed system is to obtain input data samples and thus generate a raw data heap for processing. These data are acquired and are progressed with respect to overall system requirements and restrains. On storing these data, considering one of the most accurate and reliable data set, the system is computed and mathematically modeled.
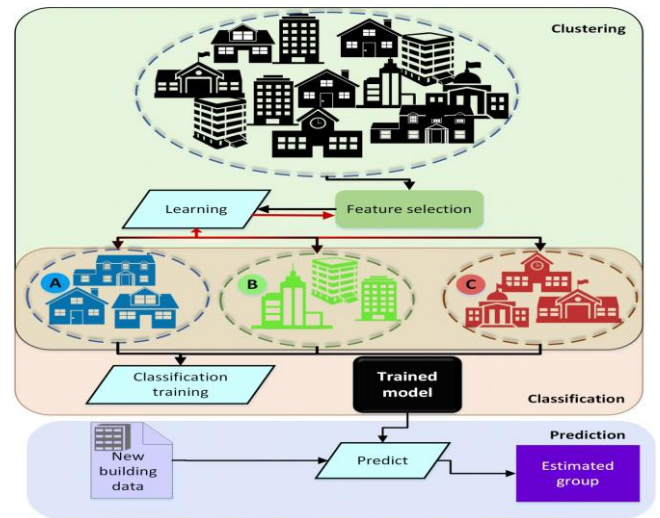

Fig 4.1: Overall System Architectural Design

On the other hand the system experiences training of data integration and system cracks the data into training and testing data. The huddling algorithms provide supplementary accurate tools for grouping various building in evaluation with traditional method where primarily relay on building usage type. It should be noted that via the clustering algorithm for outlining groups it is not possible to approximate clusters for new buildings. Hence, for the rationale of determining calculation three major methods that have widely used in the building sector for supervised learning are ANN, SVM and Gaussian distribution regression models. K-means and hierarchical clustering techniques have also utilized for unsupervised learning purposes. These methods are conversed in detail in the following sections, and a outline of other ML techniques is presented subsequently.

➤ Low Level Designs

1. Data cleaning and formatting

Dataset is an exact group of observations with no missing values or glitchs (looking at you mtcars and iris datasets). Real-world data is muddled which means we need to clean and wrangle it into an adequate format before we initiate the analysis.
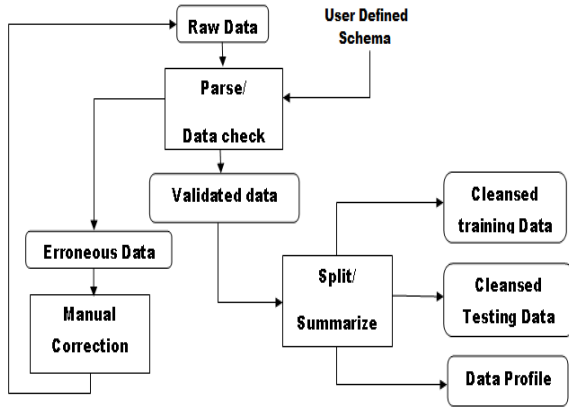
Fig 4.2: Data cleaning and formatting

Data onslaught is an un-glamorous, but essential part of most genuine data science problems. System is simulated and designed according to the necessities of data cleaning and monitoring as shown in Fig 4.2. Machine learning is all about training and feeding information to algorithms to perform diverse compute intensive tasks. However, businesses naturally face challenges in feeding the right data to machine learning algorithms or cleaning of inappropriate and error-prone data. In other words, when it comes to exploiting ML data, most of the time is spent on cleaning data sets or generating a dataset that is free of errors. Setting up a superiority plan, substantial missing values, removing rows, dropping data size are some of the best applications used for data cleaning in Machine Learning. Enterprises nowadays are gradually more utilizing machine learning for obtaining, storing, and analyzing data in arrange to assist better decision making models.

2. Exploratory data analysis:

To commence the EDA, we will spotlight on a single variable, the Energy Star Score, because this is the goal for our machine learning models. We can rename the feature to score for effortlessness and then start exploring this value.
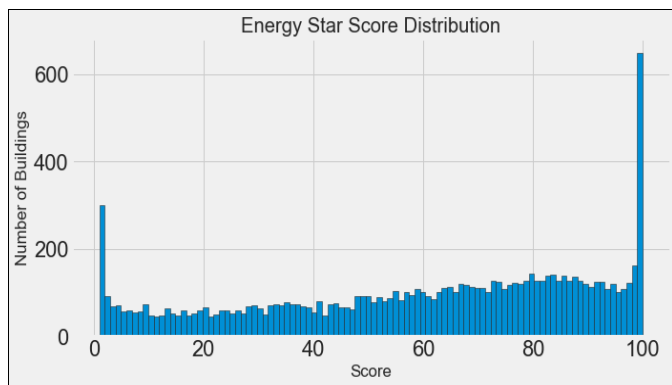

Fig 4.3: Histogram to show energy star score percentile

Exploratory Data Analysis is shown in Fig 4.3. As the Energy Star Score is a percentile grade, we would expect to see a completely flat divisions with each achieve making up 1% of the distribution (about 90 buildings). However, this is definitely not the case as we can see that the two most common gain, 1 and 100, make up a inconsistent figure of the overall scores. To disparity the Energy Star Score, we can gaze at the Energy Use Intensity (EUI), which is the entire energy use divided by the square footage of the building. Here the energy procedure is not self-reported, so this could be a more intention calculate of the energy efficiency of a building. Moreover, this is not a percentile grade, so the complete values are important and we would anticipate them to be approximately usually distributed with perhaps a few outliers on the low or high end.

3. **A baseline and compare several machine learning models on a performance metric:** Scheming numerous metrics and trying to conclude how important each one is. In this case, we are using regression; the mean absolute error is a suitable metric. This is also interpretable because it symbolize the standard amount our estimate if off by in the same units as the target value
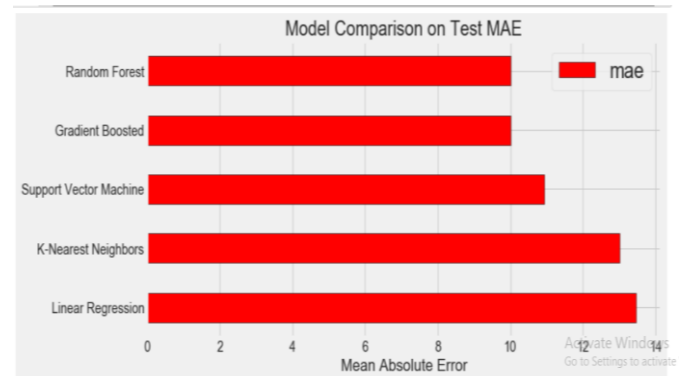

Fig 4.4: comparing several machine learning models

On fetching input, the system is designed to perform a transformation and data cypher as shown in Fig 4.4. To evaluate the models, the Scikit-Learn defaults for the model hyperparameters are used. Scikit should be optimized before actually using a model. At first, we determine the baseline performance of each model, and then we select the best performing model for further optimization using hyperparameter tuning.

4. Evaluate the best model on the testing set:

We will use the paramount model from hyperparameter tuning to make prognosis on the testing set. The performance of the test set should be a good indicator of how the model would perform if deployed in the real world.

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
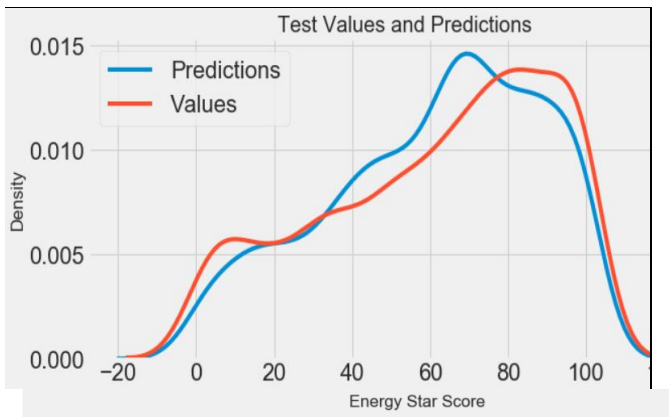NCRACES - 2019 Conference Proceedings

Fig 4.4: density of the predicted value

Fig 4.4 shows the predicted as well as the training values. The distribution appears to be nearly the same even though the concentration of the predicted values is closer to the median of the test values relatively to the actual peak at 100. It materialize the model might be less precise at predicting the tremendous values and instead predicts values closer to the median

5. **Interpret the model results to the extent possible:** One of the essential ways we can understand an ensemble of decision trees is through the feature importance. These can be interpreted as the variables which are most extrapolative of the target.
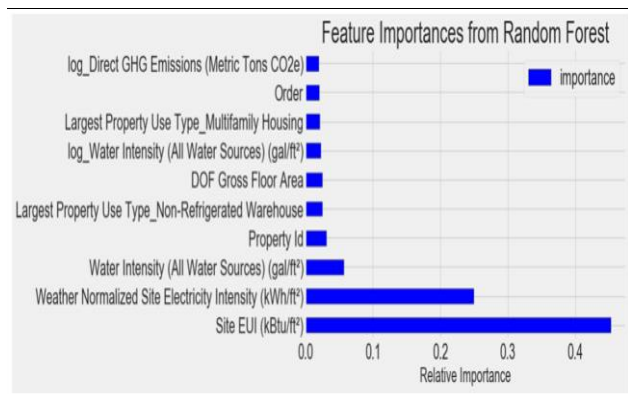

Fig 4.6: Describes feature is important for finding the score

Extracting the attribute importance from a trained group of trees is quite easy in scikit-learn. We will store the feature importance in a data frame to scrutinize and envision them. The Site Energy Use Intensity, Site EUI (kBtu/ft²), and the Weather Normalized Site Electricity Intensity, Weather Normalized Site Electricity Intensity (kWh/ft²) are the two majority important features by quite a large margin. After that, the comparative importance drops off considerably which designate that we might not need to retain all of the features to create a model with nearly the same performance.
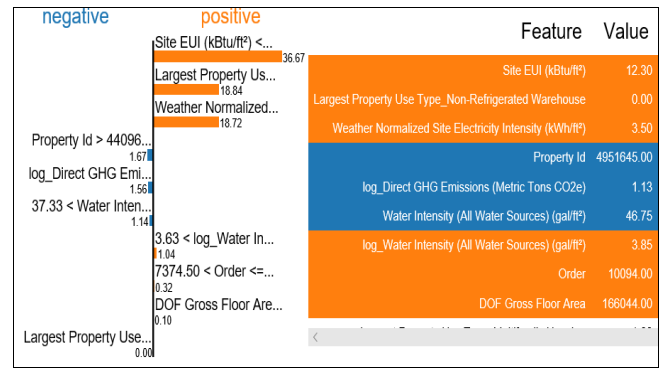

Fig 7.5: Model gradient boosted model predicted a score of 12.86 and the actual value was 100

Observing break down plots like these allow us to get an design of how the model makes a forecast. This is probably most precious for cases where the model is off by a large amount as we can scrutinize the errors and perhaps engineer enhanced features or regulate the hyperparameters of the model to recover predictions for next time. The examples where the model is off the most could also be interesting edge cases to look at physically. The models drastically misjudge the Energy Star Score for the preliminary building because of the elevated Site EUI. We might therefore want to ask why the building has such a elevated Energy Star Score even though it has such a high EUI. A development such as this where we try to work with the machine learning algorithm to expand understanding of a problem seems much better than simply letting the model make predictions and completely trusting them.

The plot from LIME again shows the contribution to the prediciton of each of feature variables for the example. For instance, because the Site EUI was less than 62.70, that contributed significantly to a higher estimate of the score. Likewise, the year built being less than 1927 also positively contributed to the final prediction. concluding model is composed of 800 decision trees, and looking at a lone one is not indicative of the entire model, it still allows us to see the general idea of how a decision tree works. We will first pull out a tree from the forest and then put aside it using sklearn.tree.export_graph viz. This keep the tree as a .dot file which can be changed to a png using command line instructions

4.CONCLUSION

We can make a model to exactly induce the Energy Star Score of a structure and we have undaunted that the Site Energy Use Intensity, the Electricity Intensity, the floor zone, the petroleum gas use, and the Building Type are the most valuable measures for deciding the vitality star score. Vitality Star Scores are inclined high with a conflicting worldwide greatest at 100 and an auxiliary most extreme at 1. Living arrangement corridors, non-refrigerated product houses and workplaces have higher vitality star gain than senior consideration networks and lodgings with scores of multi-family lodging declining in the middle. Petroleum

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCRACES - 2019 Conference Proceedings

gas, the Site Energy Use Intensity and the Electricity Intensity are the use is all adversely connected with the Energy Star Score.   An irregular timberland regressor qualified on the preparation information had the option to achieve a normal supreme deficiency of 10 points on a hold-out testing set, which was significantly superior to anything the pattern measure. Whenever gave information for another structure, a prepared model can precisely construe the Energy Star Score. This report additionally identified regions for outline grasp finding a target estimation of overall structure vitality execution and formalizing why the Energy Star Score circulations change between structure types.

FUTURE ENHANCEMENT

The existing model has achieved roughly 74% of the accuracy on the test set. Received results can be further concentrating on with artificial intelligence methods.

REFERENCE

[1] Ahmad, A.S., Hassan, M.Y., Abdullah, M.P., Rahman, H.A., Hussin, F, Abdullah, H.,
Saidur, R. (2014). A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renewable and Sustainable Energy Reviews, 33, 102–109. https://doi.org/10.1016/j.rser.2014.01.069.

[2] Alam, A.G., Baek, C.I., Han, H. (2016). Prediction and Analysis of Building Energy
Efficiency Using Artificial Neural Network and Design of Experiments. Applied Mechanics and Materials, 819, 541–545. https://doi.org/10.4028/ www.scientific.net/AMM.819.541.

[3] Aydinalp, M., Ugursal, V.I., Fung, A.S. (2004). Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks. Applied Energy, 79(2), 159–178. https://doi.org/10. 1016/j.apenergy.2003.12.006.

[4] Beccali, M., Ciulla, G., Lo Brano, V., Galatioto, A., Bonomolo, M. (2017). Artificial
neural network decision support tool for assessment of the energy performance and the refurbishment actions for the nonresidential building stock in Southern Italy. Energy, 137, 1201–1218. https://doi.org/10. 1016/j.energy.2017.05.200.

[5] Benedetti, M., Cesarotti, V., Introna, V., Serranti, J. (2016). Energy consumption control automation using Artificial Neural Networks and adaptive algorithms: Proposal of a new methodology and case study. Applied Energy, 165, 60–71. https://doi.org/10.1016/j.apenergy.2015.12.066.

[6] Ben-Nakhi, A.E., & Mahmoud, M.A. (2004). Cooling load prediction for buildings
using general regression neural networks. Energy Conversion and Management, 45(13–14), 2127–2141. https://doi.org/10.1016/j.enconman. 2003.10.009. Biswas, M.R., Robinson, M.D., Fumo, M.D. (2016). Prediction of residential building energy consumption: A neural network approach. Energy, 117, 84–92. https://doi.org/10.1016/j.energy.2016.10.066.

[7] Buratti, C., Barbanera, M., Palladino, D. (2014). An original tool for checking energy performance and certification of buildings by means of Artificialn Neural Networks. Applied Energy, 120, 125–132. https://doi.org/10.1016/j.apenergy.2014.01.053.

[8] Burkhart, M.C., Heo, Y., Zavala, V.M. (2014). Measurement and verification of building systems under uncertain data: A Gaussian process modeling approach. Energy and Buildings, 75, 189–198. https://doi.org/10.1016/j. enbuild.2014.01.048.

[9] Cheng, M.-Y., & Cao, M.-T. (2014). Accurately predicting building energy performance using evolutionary multivariate adaptive regression splines. Applied Soft Computing, 22, 178–188. https://doi.org/10.1016/j.asoc.2014.05.015.

[10] Chung, W (2011). Review of building energy-use performance benchmarking

methodologies. Applied Energy, 88(5), 1470–1479. https://doi.org/10.1016/j. apenergy.2010.11.022.

[11] Deb, C., Eang, L.S., Yang, J., Santamouris, M. (2016). Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. Energy and Buildings, 121, 284–297. https://doi.org/10.1016/j.enbuild.2015.12.050.

[12] Edwards, R.E., New, J., Parker, L.E. (2012). Predicting future hourly residential electrical consumption: A machine learning case study. Energy and Buildings, 49, 591–603. https://doi.org/10.1016/j.enbuild.2012.03.010.

[13] Khayatian, F., Sarto, L., Dall'O', G. (2016). Application of neural networks for evaluating energy performance certificates of residential buildings. Energy and Buildings, 125, 45–54. https://doi.org/10.1016/j.enbuild.2016.04.067.CrossRefGoogle Scholar

[14] Kialashaki, A., & Reisel, J.R. (2013). Modeling of the energy demand of the residential sector in the United States using regression models and artificial neural networks. Applied Energy, 108, 271–280. https://doi.org/10.1016/j.apenergy.2013.03.034.

[15] Li, Q., Ren, P., Meng, Q. (2010). Prediction model of annual energy consumption of residential buildings. In 2010 international conference on advances in energy engineering. Retrieved from https://doi.org/10.1109/ICAEE.2010.5557576. IEEE, (pp. 223–226).

[16] Magoulès, F., Zhao, H.x., Elizondo, D. (2013). Development of an RDP neural network for building energy consumption fault detection and diagnosis. Energy and Buildings, 62, 133–138. https://doi.org/10.1016/j.enbuild.2013.02.050.

[17] Neto, A.H., & Fiorelli, F.A.S. (2008). Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption. Energy and Buildings, 40(12), 2169–2176. https://doi.org/10.1016/j.enbuild.2008.06.013.

[18] Paudel, S., Elmtiri, M., Kling, W.L., Corre, O.L., Lacarrière, B. (2014). Pseudo dynamic transitional modeling of building heating energy demand using artificial neural network. Energy and Buildings, 70, 81–93. https://doi.org/10.1016/j.enbuild.2013.11.051.

[19] Wong, S., Wan, K.K., Lam, T.N. (2010). Artificial neural networks for energy analysis of office buildings with daylighting. Applied Energy, 87(2), 551–557. https://doi.org/10.1016/j.apenergy.2009.06.028.CrossRefGoogle Scholar

[20] Xing-ping, Z., & Rui, G.U. (2007). Electrical Energy Consumption Forecasting Based on Cointegration and a Support Vector Machine in China. In Wseas transactions on mathematics,vol.6.Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.533.9017 &rep=rep1&type=pdf, (pp. 878–883).

[21] Yang, R., & Wang, L. (2013). Development of multi-agent system for building energy and comfort management based on occupant behaviors. Energy and Buildings, 56, 1–7. https://doi.org/10.1016/j.enbuild.2012.10.025.