

Consulting Recommendations

Prepared for



Green
Mountain
Energy®

Created by



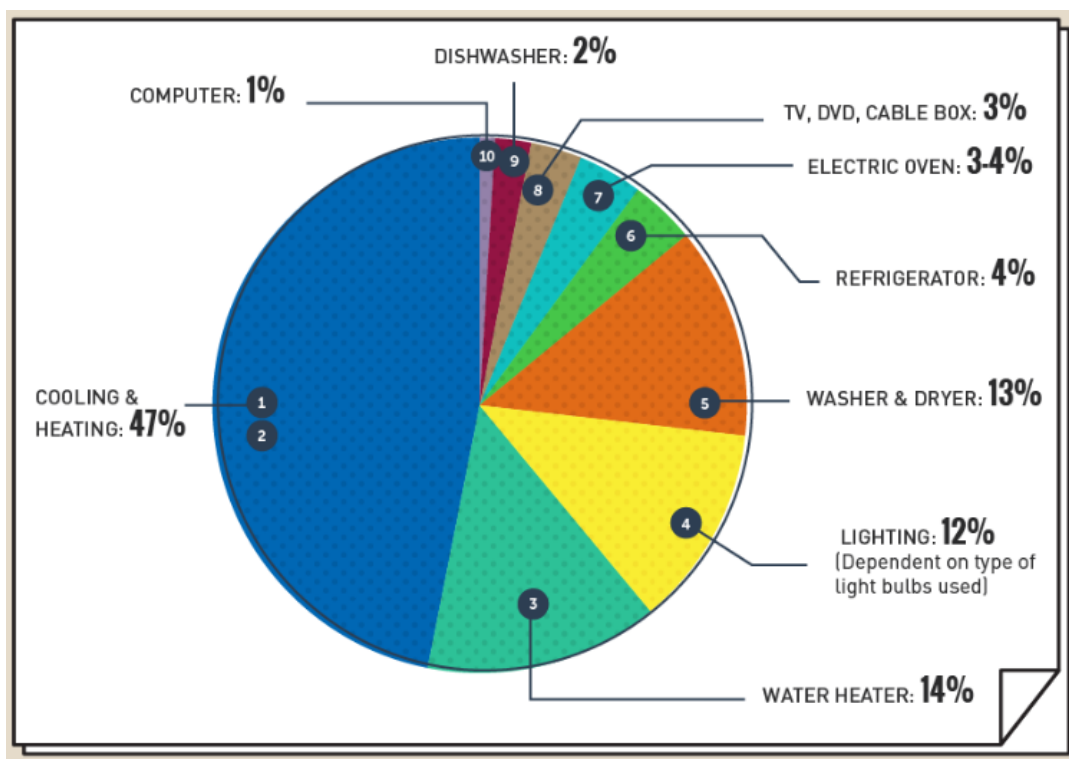
1. Introduction

As the premier and first green energy company to serve Texas, Green Mountain Energy understands accurately predicting residential energy consumption is crucial for optimizing resource allocation and promoting the adoption of sustainable solutions. This recommendation document outlines a solution that utilizes both linear regression and deep neural network models to predict residential energy consumption. By leveraging these advanced analytical techniques, Green Mountain can make informed decisions, improve operational efficiency, and support a greener future.

2. Data Collection

A comprehensive data collection process is crucial to understanding residential energy consumption. One of the main factors that influences energy usage is the weather. When it's hot, people tend to turn on their air conditioners, and when it's cold, they use heating systems. In fact, it is estimated that about 50% of residential energy consumption is due to heating and cooling, as I'm sure you are aware (see Figure 1 below). Therefore, our team decided to focus on weather conditions for our study.

Figure 1. General proportion of energy consumption by appliance in the home

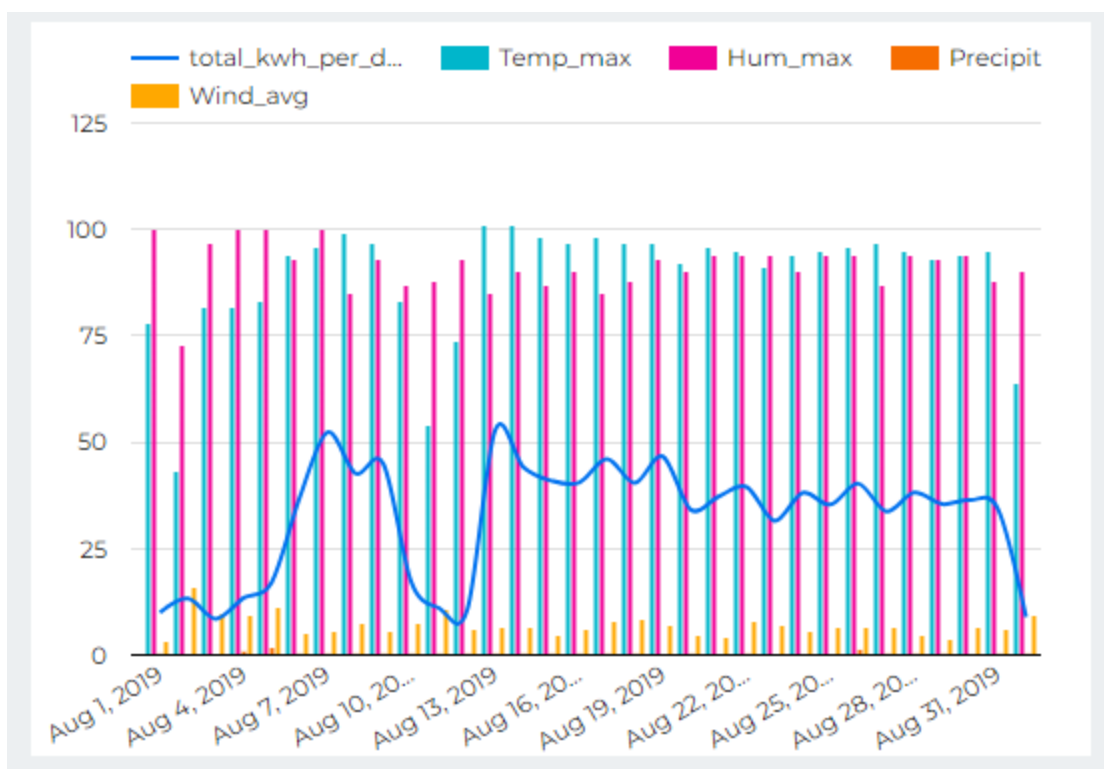


(source: "What Uses the Most Energy in Your Home?" by Jeff Desjardins of the Visual Capitalist)

To gather the necessary data, we collected historical weather readings including temperature, wind, and precipitation. We also obtained real-time data and forecasts to match with our collected sample of residential meter readings. By comparing weather patterns with energy usage, we aimed to identify any correlations or patterns between the two.

This approach allowed us to analyze the impact of weather on residential energy consumption more accurately. By considering weather factors alongside other variables, we can gain a better understanding of how different weather conditions affect energy usage in homes. In Figure 2 below we see the primary dimensions we used in developing our model, overlaid by the actual energy consumption we gathered over the course of a month.

Figure 2. Weather parameters vs. daily average kwh consumed, August 2019



3. Data Preprocessing

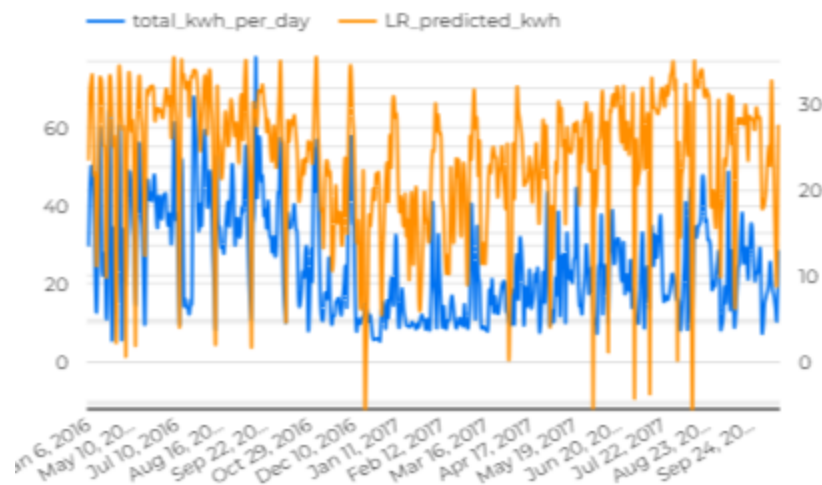
Once the data was collected, thorough preprocessing was required to ensure data quality and compatibility across different sources. This involved cleaning the data, to include identifying and addressing outliers and anomalies; handling missing values; and data relational joining for a comprehensive dataset.

4. Linear Regression Model

Linear regression can serve as an initial predictive model for residential energy consumption due to its simplicity and interpretability. By our gathered data on weather conditions we developed a linear regression model which could estimate energy consumption accurately. Below is an example of how the model

predicted consumption versus our historical actuals (Figure 3).

Figure 3. Linear regression predicted consumption vs. actual January 2016 - September 2017

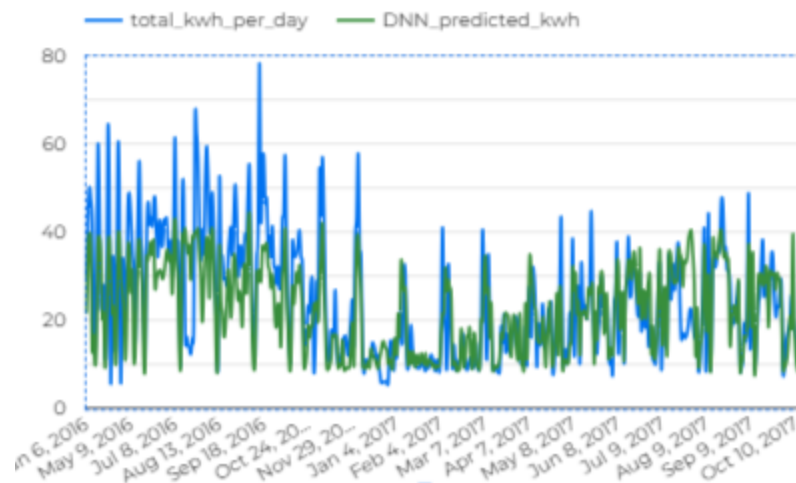


Overall, the model was a moderately good fit, although it was a big conservative when it came to the highs of energy consumption, a scalar adding to the constant for the model, may be beneficial for some greater cushion in action.

5. Deep Neural Network Model

To capture more complex relationships and nonlinear patterns, a deep neural network (DNN) model was developed. DNNs excel at handling large amounts of data and can automatically learn intricate features from the input variables. By utilizing multiple hidden layers, activation functions, and appropriate optimization algorithms, DNNs can effectively predict residential energy consumption with significant accuracy. Below (Figure 4) was the result for our DNN model, over the same time period as our linear regression sample.

Figure 4. DNN predicted consumption vs. actual January 2016 - September 2017



6. Model Training and Evaluation

The dataset was randomly split into training and testing sets to evaluate the performance of both the linear regression and DNN models. The models were trained using various techniques such as cross-validation, regularization, and hyperparameter tuning to optimize their performance. Below aEvaluation metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared can be used to compare and assess the models' predictive capabilities.

Figure 5. Evaluation metrics for Linear Regression vs. DNN models

Metric	Linear Regression	DNN
MAE (0 is best)	5.57	5.32
RMSE (0 is best)	8.11	7.77
R-squared (1 is best)	0.63	0.66

It should be noted that visually (Figures 3 and 4) and by our evaluation metrics, the DNN model performs marginally better, but still has difficulty predicting spikes in usage. The DNN model also required a significant amount more computing resources to develop, which would only increase with the potential introduction of increased complexity in the dataset.

7. Insights and Recommendations:

The predictions generated by the linear regression and DNN models can provide valuable insights into residential energy consumption patterns. By analyzing the model outputs, Green Mountain can identify peak energy demand periods, target energy-saving initiatives, optimize resource allocation, and develop tailored energy efficiency programs for Houston, Texas, as well as the other regions you are involved in to include New York and Pennsylvania.

As discussed in the previous section, the DNN model has proven to be more accurate than the Linear Regression model, but also requires a significant amount more time and computer power to process. These tradeoffs have to be weighed against your current goals as an organization and IT infrastructure to see what is feasible. Either way, you will be incorporating accurate models into your planning, which will only get more accurate over time with increased dimensions to consider.

That said, given the Linear Regression model is more agile and user friendly in terms of computing and adjusting, we recommend incorporating this model into Green Mountain Energy’s planning strategy at this time.

8. Ways Forward

We do not want to leave Green Mountain with only today's recommendation, but provide a vision for the future on what the assimilation and continued development of our models may look like. Thus, below, we have listed 10 steps forward we might take upon acceptance of our proposal.

1. Collaborate with other utility companies: Partnering with other utility companies can provide access to their data streams and metrics, which can further enhance the accuracy of our models. Sharing data and knowledge can lead to better insights and more comprehensive solutions.
2. Explore additional data sources: Apart from Green Mountain's data streams and metrics, consider incorporating data from other sources such as demographic information or general housing characteristics. This will help create a more diverse and extensive dataset, improving the reliability and effectiveness of the models.
3. Conduct pilot projects in different cities: While Houston can serve as a valuable test site, it is important to validate the model's effectiveness in different cities with varying characteristics. By conducting pilot projects in other cities, we can refine the model and tailor it to specific regional needs, ensuring its applicability nationwide.
4. Further engagement with stakeholders and decision-makers: To gain support and buy-in for incorporating data-driven decision making and analytics, it is crucial to engage with key stakeholders and decision-makers. Present the benefits and potential outcomes of utilizing our model, showcasing how it can improve operational efficiency, reduce costs, and enhance sustainability efforts.
5. Continuously update and improve the model: Technology and data are ever-evolving, and it is essential to continuously update and improve the model. We must stay abreast of the latest advancements in data analytics, machine learning, and AI, and incorporate new features and algorithms into the model to enhance its performance and accuracy.
6. Foster a culture of data-driven decision making: Implement training programs and workshops to educate employees about the benefits and importance of data-driven decision making. Encourage a culture that values data-driven insights and promotes the use of analytics for making informed decisions at all levels of the organization.
7. Establish partnerships with research institutions: Collaborating with research institutions can provide access to cutting-edge research, expertise, and resources. Partnering with universities such as utilizing Rice University's REINVENTS initiative or research centers specializing in energy analytics can help validate and improve our models while fostering innovation in the field.
8. Develop a roadmap for scaling up: Create a comprehensive roadmap outlining the steps and timeline for scaling up the implementation of our model nationwide. Identify potential challenges, resource requirements, and strategies for overcoming obstacles to ensure a smooth and successful transition.
9. Monitor and evaluate the performance: Continuously monitor and evaluate the performance of the model once implemented. Regularly analyze the outcomes and compare them against predefined metrics to assess the effectiveness and identify areas for improvement. This feedback loop will enable us to refine the model and maximize its benefits over time.
10. Share success stories and best practices: As the model proves its value and yields positive results, share success stories and best practices with other utility companies and industry forums. By showcasing the real-world impact of data-driven decision making, we can inspire and influence others to adopt similar approaches, leading to a broader transformation in the industry.