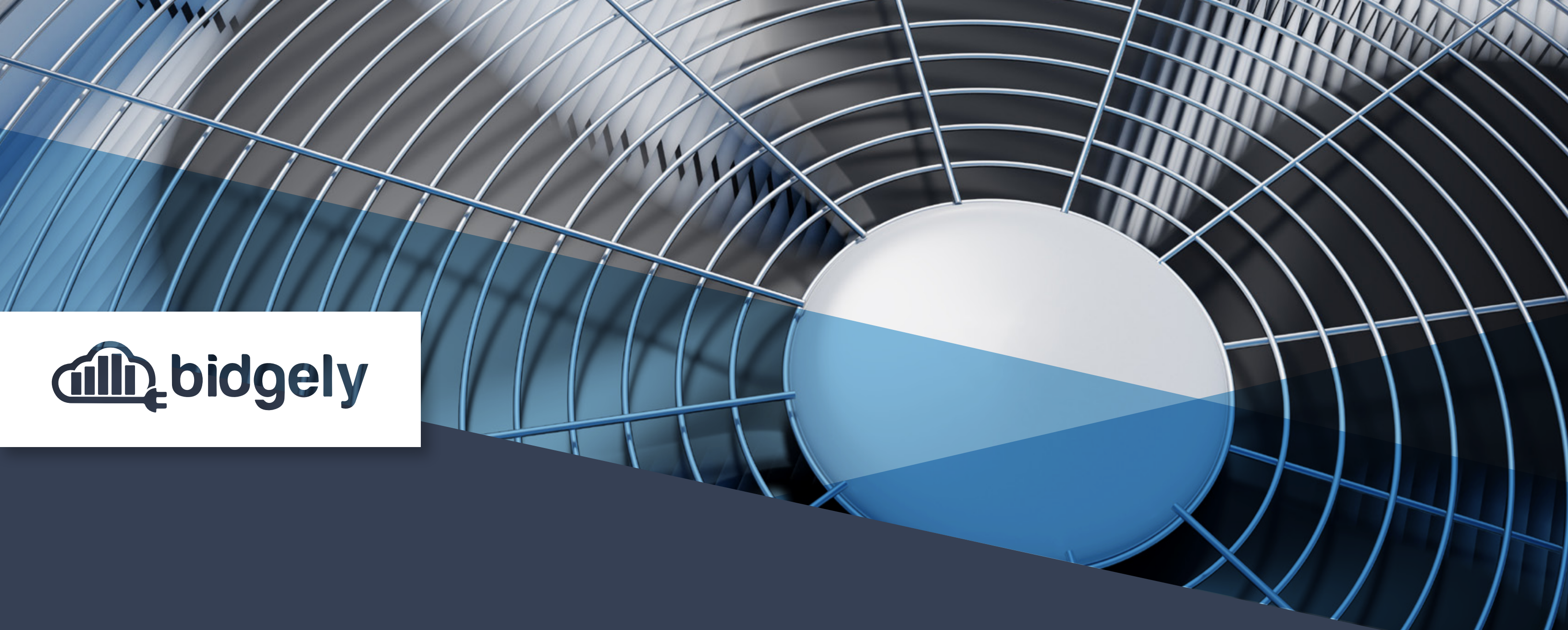

HVAC TYPE AND FUEL TYPE CLASSIFICATION WITH AMI DATA

A TECHNICAL BRIEF BY THE BIDGELY DATA SCIENCE TEAM

1 BACKGROUND

HVAC appliances are amongst the highest energy consuming appliances at residential sites. Ensuring that customers have access to and ultimately choose energy efficient HVAC appliances is critical to improving long-term energy efficiency in utility service territories. If a customer chooses a less-than-optimal HVAC type, the customer may incur higher electricity bills while the power utility may experience greater energy demand management challenges. Hence, utilities have a vested interest in identifying the type of HVAC and the primary fuel type for appliances that are installed in a residence. Utilities can leverage this information to efficiently and cost effectively conduct hyper-targeted HVAC programs to influence only those customers who have an opportunity to optimize their HVAC usage.

Bidgely is one of the only companies that has developed a classification suite of algorithms that identifies the type of HVAC and primary fuel type in use within a residence.

2 USE CASES

Insight into fuel and heating/cooling appliance type enables utilities to efficiently implement targeted customer programs, including:

1. **Electrification programs:** Target residences in which HVAC is powered by gas, and where there is incentive for residents to shift to an electric based system
2. **Efficient HVAC Appliance Retrofits:** Target customers with incentives/rebates to purchase and install more efficient HVAC systems, such as heat pumps
3. **Behavioral Energy Efficiency:** Provide more relevant similar home comparisons of HVAC load (based on detected appliances and fuel types) to more effectively promote energy efficiency behaviors

3 THE BIDGELY SOLUTION

Our AI-based HVAC classification relies on characteristic features extracted from our patented HVAC disaggregation algorithm. The algorithm receives energy consumption information from smart meters and ambient temperature and extracts 18 defining features related to cooling and heating appliances in a house. These features primarily capture appliance amplitude and usage behavior while also measuring sensitivity to temperature. In addition, the algorithms extract other supporting parameters that capture the range of HVAC amplitude variation and associated set points.

These features form the characteristic base for each residence's HVAC-type classification. Using the smart meter data and the ground truth about the type of HVAC appliance for a small set of homes of a region, we are able to train a machine learning model to identify complex classification boundaries based on the 18 cooling and heating features extracted through disaggregation.

Figure 1 below illustrates the feature extraction process and model creation. This model can be applied to an entire regional home population to predict their HVAC types.

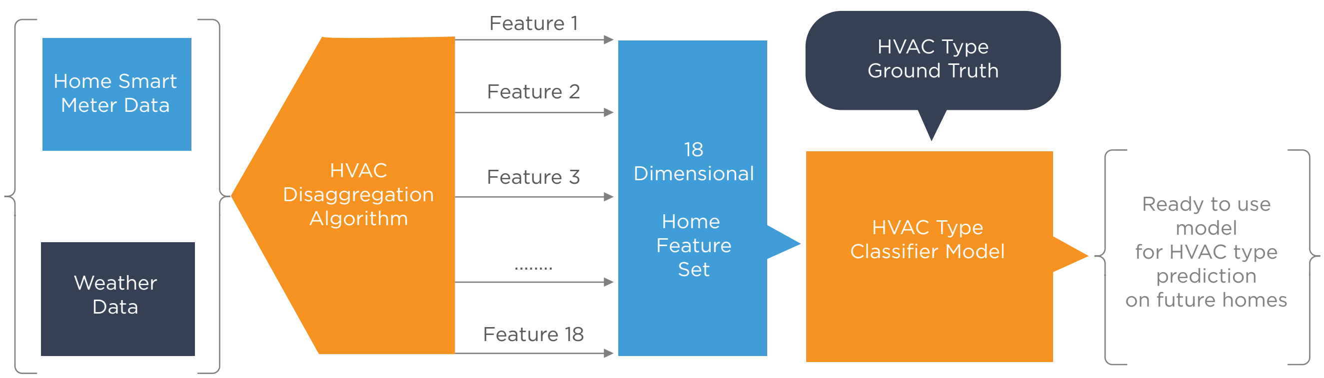

Figure 1: Flow chart showing Feature extraction process and model creation for HVAC Type Classification

The disaggregation algorithm also reveals information about on-demand HVAC and always-on HVAC usage. Bidgely's technology compares these components and their amplitude characteristics for each house in order to extract the information about primary fuel type in use.

---

[1] Ground truth means information that is known to be real or true, provided by direct observation and measurement, it is used to check the results of machine learning for accuracy against the real world

3.1. Data Requirements

It is possible to make a prediction of HVAC type and primary fuel type for an unknown residence using only its smart meter energy data and weather data, including:

1. **Smart meter data:** Home energy data acquired by smart meters at 5, 15, 30 or 60 minutes sampling.
2. **Weather data:** The ambient temperature at the zipcode area of the house.
3. **Ground truth data**: HVAC type data is required for a small set of houses to let the model learn the classification pattern.

3.2. HVAC Types and Fuel Types

The classification is dynamic enough to accommodate a wide variety of HVAC types. It is only limited by the number of classes present in the heating and cooling ground truth data provided by the utility. The most popular class prediction is as follows:

1. **Heating:**
    a. Central furnace,
    b. Baseboard,
    c. Packaged terminal air conditioner (PTAC),
    d. Air source heat pump,
    e. Mini split,
    f. Boiler,
    g. Wood furnace,
    h. Ground-coupled heat pump (GCHP)
2. **Cooling:**
    a. Packaged DX,
    b. Mini split,
    c. Split DX,
    d. PTAC,
    e. Air source heat pump,
    f. GCHP,

For the primary fuel type, we have capability to tag a residence into two categories:

1. **Gas-based**
2. **Electricity based**

3.3. Solution for HVAC Class

3.3.1. Model Training

The homes with known HVAC types are divided into two sub groups. The first group of homes (training set) with known HVAC types are taken for model training. The second group of homes (evaluation set) is used to evaluate the accuracy of the trained model. The model ingests 18 disaggregation features for each home to learn the pattern associated with each HVAC class present in the training set. Our latest model is based on a random forest classifier that is tuned for most accurate predictions for both heating and cooling types.

The random forest classifier learns complex decision boundaries in 18 dimensional space and makes predictions for the most probable HVAC class for a home.

3.3.2. Model Evaluation

The trained model is used to make predictions for HVAC classes on unseen evaluation datasets. The tuned random forest classifier evaluates 18 dimension features for each home in order to predict heating and cooling classes. These predicted classes are compared with actual known classes for each home in the evaluation set and evaluated using an accuracy score.

Our latest model demonstrates an overall accuracy of 91% for predicting heating classes and 80% for predicting cooling classes.

Figure 2 below illustrates the process of model evaluation on the evaluation set.

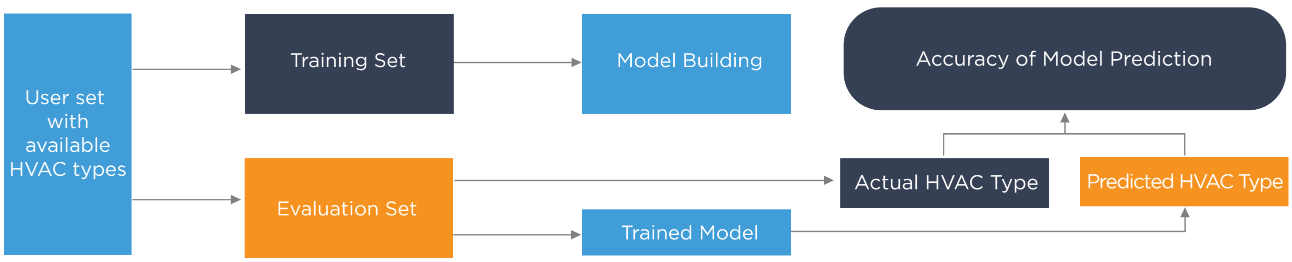

Figure 2: Flow chart showing Model evaluation process

---

[2] Random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time.

	central_furnace	central_furnace_electric	baseboard_electric	heat_pump_electric	mini_split_electric	boiler
central_furnace	0.967000	0.002000	0.01200	0.13000	0.004000	0.002000
central_furnace_electric	0.310000	0.381000	0.214000	0.071000	0.024000	0.000000
baseboard_electric	0.136000	0.012000	0.8230000	0.016000	0.008000	0.004000
heat_pump_electric	0.224000	0.041000	0.0820000	0.606000	0.035000	0.012000
mini_split_electric	0.337000	0.010000	0.2480000	0.089000	0.287000	0.030000
boiler	0.260000	0.000000	0.000000	0.020000	0.000000	0.720000

Table 1: Table shows accuracy numbers on evaluation dataset for heating classes

The Table 1 above shows class-wise accuracy scores for heating type classification of homes from North Western, United States. Each row represents true labels for HVAC class and each column carries predicted classes.

Since, the class level accuracies improve with the increase in number of class instances available for training, including your specific geography's ground truth data will improve accuracy.

3.3.3. Proof Point of Class Separation

The core of the classification logic is class separation through features, where good feature patterns differ from class-to-class. This difference in feature patterns across classes serves as a base for the model to learn decision boundaries for classification. When multiple features are present, a feature may separate only a few classes, but aggregated features enable a better separation in higher dimensions.

In the case of our HVAC type classification, the 18 features evidently induce enough class separation for accurate class prediction. Figure 3 and Figure 4 below show class separation based on two of the features produced by the HVAC disaggregation algorithm.

In Figure 3, we see some of the heating classes separated using only the feature heating_1m_ mean. This heating_1m_mean feature indicates one of the heating amplitudes revealed by our disaggregation suite.

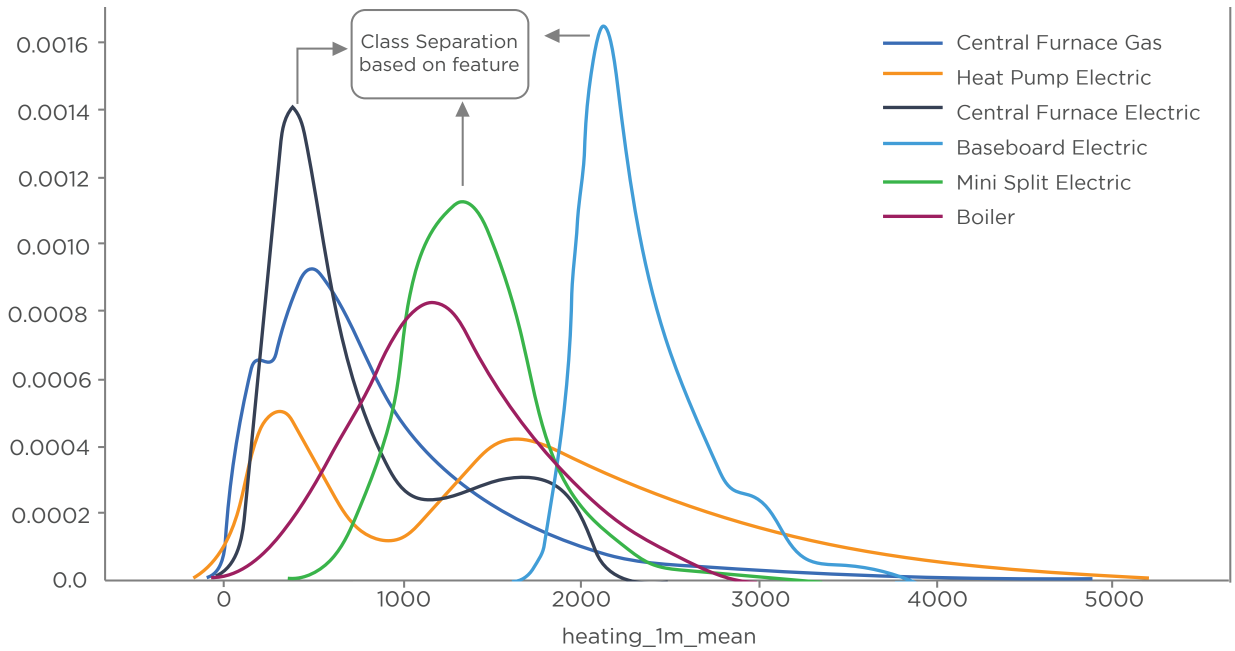

Figure 3: Separation of heating classes through disaggregation feature heating_1m_mean for classification

Figure 4 illustrates a different separation of heating classes by considering only the feature heating_2m_coefficient. This heating_2m_coefficient feature captures the consumption sensitivity of a heating device when there is a change in ambient temperature.

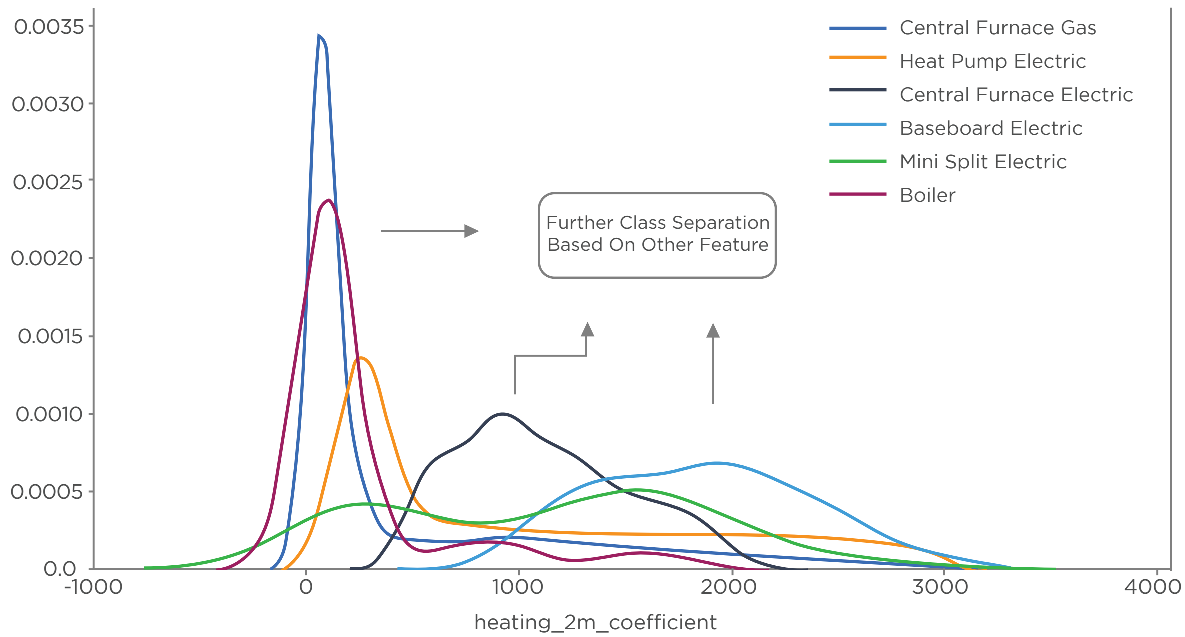

Figure 4: Separation of heating classes through disaggregation feature heating_2m_coefficient for classification

Likewise, each of 18 features induces useful class separation. When combined, classification is more accurate than when only individual features are considered.

3.4. Solution for Fuel Type

Homes with gas as primary heating fuel are characterized by low amplitude of electricity consumption. Additionally, this low amplitude corresponds to a fan component that is operational throughout the day. On the other hand, high amplitude electricity-fueled on-demand HVAC consumption is detected with a rise in ambient temperature. This critical information is extracted out of our disaggregation suite to predict primary fuel type in a house.

3.4.1. Methodology

The HVAC disaggregation algorithm gives information of "always-on" HVAC usage (low amplitude) and "on-demand" HVAC usage (relatively higher amplitude). The amplitude characteristics of these components along with their contribution to total HVAC consumption makes a feature set for a home. This feature set is taken through a decision tree algorithm to extract information of primary fuel type used at a house.

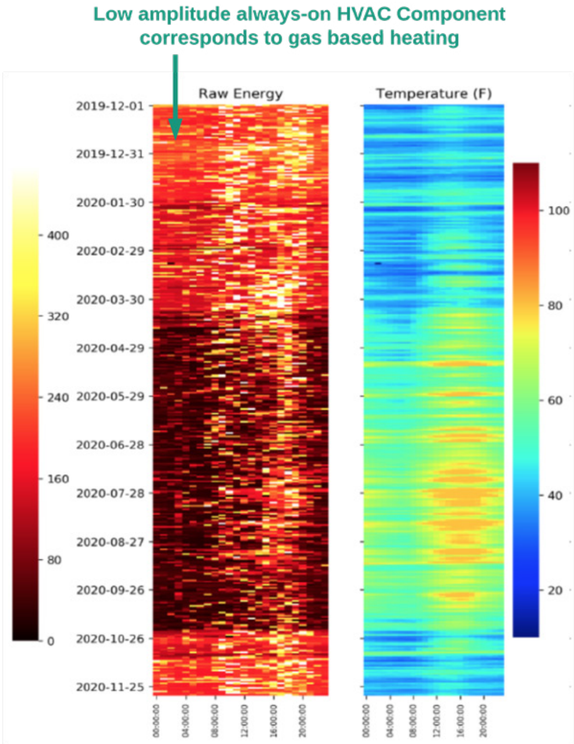

Figure 5: Heatmap showing consumption pattern of a house in a year, with gas as primary heating fuel

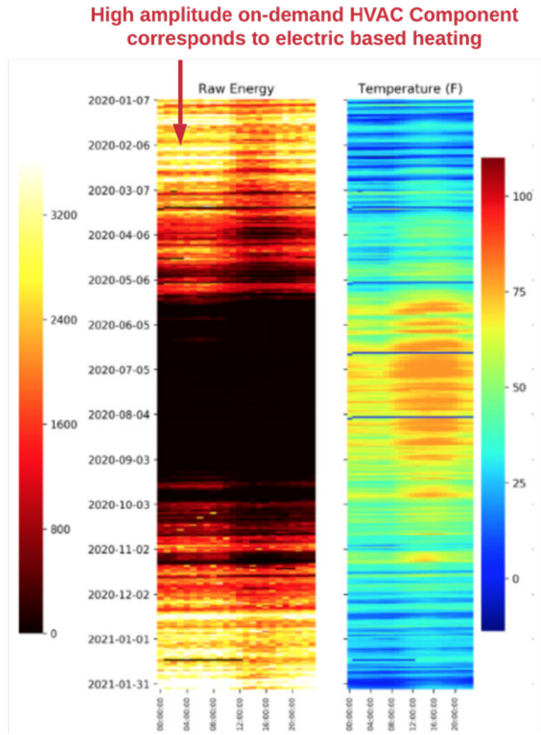

Figure 6: Heatmap showing consumption pattern of a house in a year, with electricity as primary heating fuel

3.4.2. Illustration

Figure 5 above shows the annual consumption pattern for a residence with gas as the primary fuel type. The low amplitude always-on HVAC component runs throughout the day and does not vary a lot with changes in ambient temperature. This consistent HVAC usage corresponds to an electric component of fan that is used to circulate air around the house. Figure 6 shows the annual consumption pattern of a home with electricity as the primary fuel type. In this example, we see the appliance runs at high amplitude and exhibits variation in energy consumption associated with changes in temperature. Our flexible fuel type decision tree algorithm accurately detects a range of energy consumption levels across geographies.

4 CONCLUSION

The HVAC type and fuel type classification suite in association with our disaggregation suite has capability to produce useful information that empowers utilities to run more efficient customer-targeted programs. These AI capabilities have the capacity to learn new data patterns with the introduction of new classes for HVAC or fuel types. This feature makes our classification suite a strong candidate for classification of homes around the globe.