**Author:- Sunaina Jain**



**Objective 2: One-Week-Ahead occupancy Prediction Using Seasonal-Trend Decomposition**

2025

**1. Introduction** This study presents a predictive and optimization-based strategy for reducing HVAC electricity consumption using an occupancy-aware control approach. Inspired by the work in the original research paper titled "Occupancy-based one-year-ahead heating, ventilation, and air-conditioning electricity consumption optimization using machine learning", this project adapts the methodology to a shorter horizon: a one-week forecast for the period of January 1–7, 2006. **This control logic reduced total HVAC energy from 8,635.90 kWh to 8,629.09 kWh for the week, achieving 0.08% savings.**

**2. Methodology Overview**

* **Occupancy Forecasting**: The number of occupants was forecasted at 10-minute intervals using time-series and machine learning techniques.
* **Feature Engineering**: Time-based features, lagged occupancy values, and weather-related parameters were engineered.
* **Machine Learning Models**: Two models were used: Support Vector Regression (SVR) and Feedforward Neural Network (FNN).
* **HVAC Optimization**: The predicted occupancy was utilized to design a rule-based control strategy that adjusts HVAC operation based on occupancy and weather.

**3. Feature Table (Customized for Current Work)**

Three categories of features were considered in the construction of the HVAC electricity consumption prediction models: **occupancy-based**, **time schedule-based**, and **weather-based** features.



#### ****3.0.1 Occupancy-Based Features****

Given that HVAC systems require a delay period to adjust environmental conditions, we considered the predicted number of occupants at the current timestamp and with time shifts of ±0.5, ±1, ±1.5, and ±2 hours (9 features total). This time-lagging helps capture occupancy patterns leading up to and following a given prediction time point, improving forecasting performance.

#### ****3.0.2 Time Schedule-Based Features****

To represent weekly temporal trends, we included features such as the hour and minute of the day, the day of the week, a binary weekend flag, and an indicator for working hours (9AM–5PM). These allow the model to distinguish between typical office hours and off-hours, as well as between weekdays and weekends.

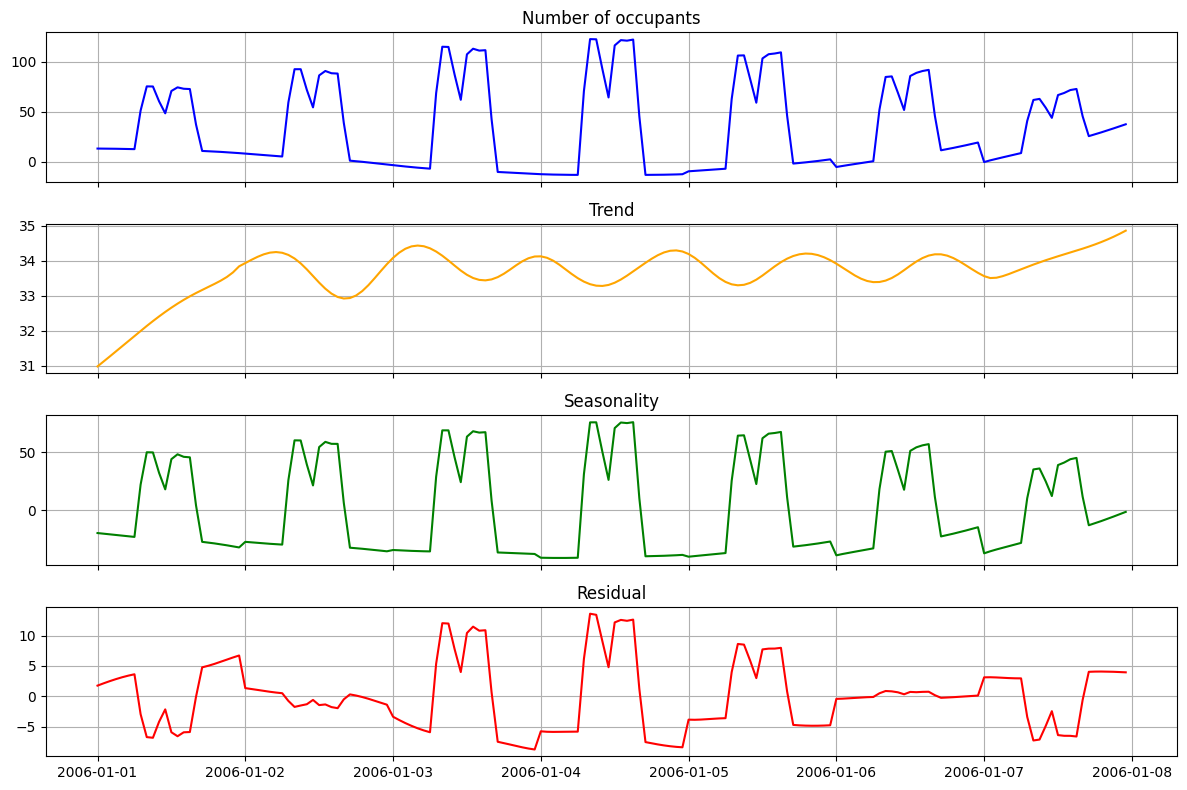
#### ****3.0.3 Weather-Based Features****

Environmental conditions heavily influence HVAC demand. The following weather features were extracted from Vancouver International Airport’s climate data for January 1–7, 2006:

* Mean temperature [°C]
* Heating degree days [°C]
* Cooling degree days [°C]
* Total precipitation [mm]
* Total snowfall [cm]
* Snow on ground [cm]

These parameters serve as proxies for heating and cooling needs, with temperature-related variables capturing thermal comfort conditions and precipitation/snow contributing to HVAC system load.

**Figure 1**



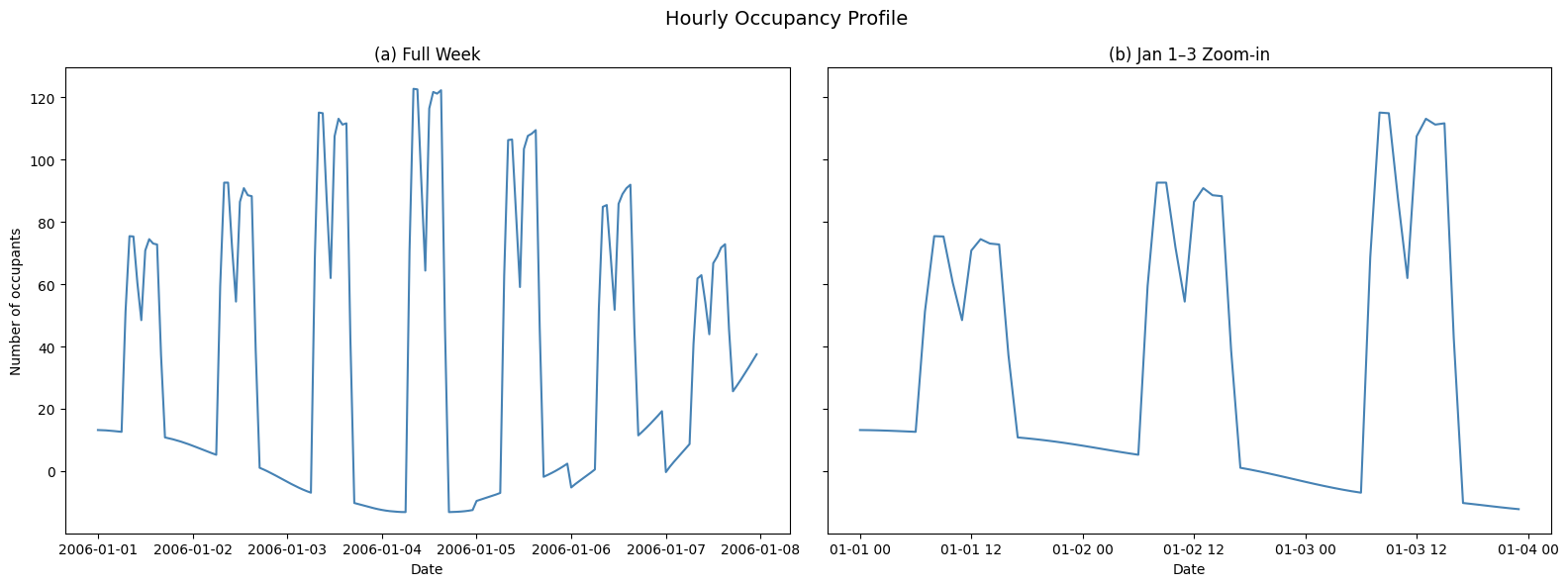
### ****3.1 Data Preprocessing and Decomposition****

To better understand the occupancy pattern prior to training the HVAC consumption models, we applied **Seasonal-Trend decomposition using Loess (STL)** to the predicted number of occupants over the week of January 1–7, 2006.

As shown in **Figure 1**, the time series of predicted occupancy was decomposed into the following components:

* **Observed Series (Top Panel – Blue)**: This plot represents the original forecasted occupancy values at 10-minute intervals. It highlights clear periodic behavior, with noticeable peaks during working hours and dips during off-hours or weekends.
* **Trend Component (Second Panel – Orange)**: The underlying trend shows a gradual rise and stabilization, indicating a steady occupancy level throughout the week. This component captures long-term directional movement independent of daily fluctuations.
* **Seasonal Component (Third Panel – Green)**: The seasonal variation represents recurring daily patterns of occupancy. The periodic structure confirms that occupants follow a consistent routine across weekdays.
* **Residual Component (Bottom Panel – Red)**: The residual (or remainder) shows the irregular part of the signal that cannot be explained by trend or seasonality. Spikes and dips in this plot may indicate unusual occupancy activity or noise in the prediction.

This decomposition confirms the suitability of including time-lagged and season-aware features in the model. The clear daily patterns and structured trend validate the decision to use time-based and occupancy-shifted variables for HVAC consumption prediction.



**Figure 2**

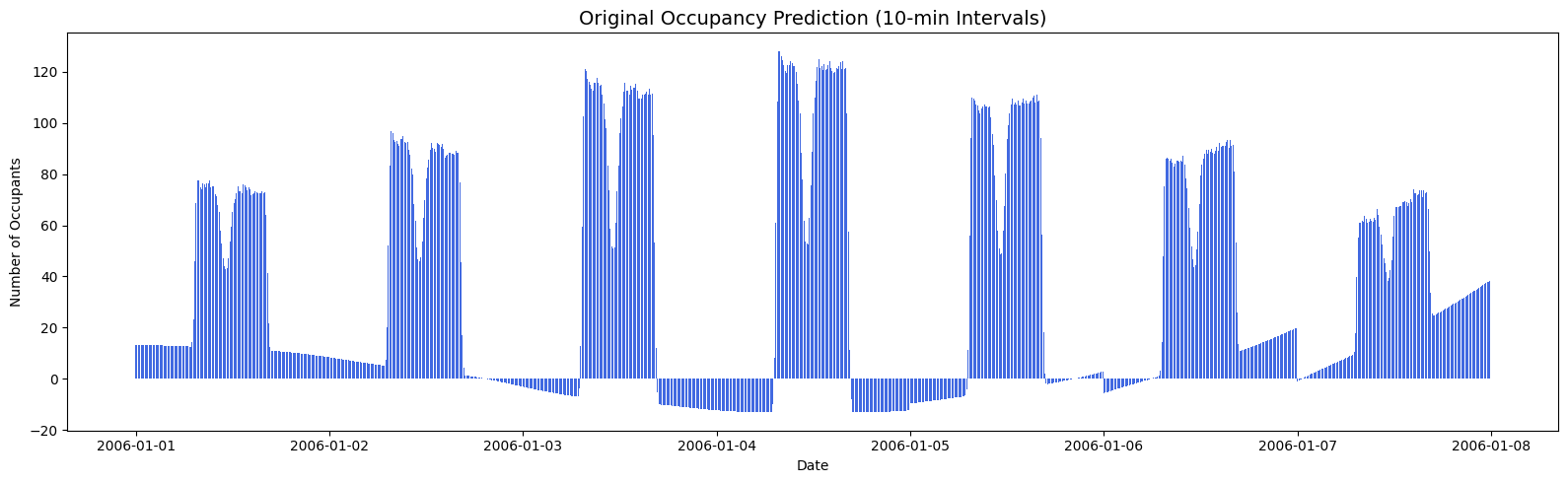
To understand occupancy trends that influence HVAC operations, we visualized the predicted number of occupants at an hourly resolution for the one-week period of **January 1–7, 2006**.

As shown in **Figure 2**, the left plot (a) displays the full-week occupancy profile, while the right plot (b) provides a zoomed-in view for the first three days:

* **(a) Full Week**: Clear and recurring occupancy patterns are observed across working days, with peaks during standard business hours and minimal presence during nights and weekends. This confirms the strong temporal structure in occupant behavior.
* **(b) Jan 1–3 Zoom-in**: The zoomed plot emphasizes the transitions in occupancy throughout the day. Sudden spikes and drops represent start/end of working periods or shift changes. The smoothness and accuracy of the forecast highlight the reliability of the model for short-term prediction.

These occupancy dynamics were essential for building predictive models of HVAC electricity consumption. The hourly resolution allowed precise alignment of HVAC operation schedules with expected occupancy levels, forming the foundation for optimization in the later phase.

**3.2 Original Occupancy Prediction at High Resolution**



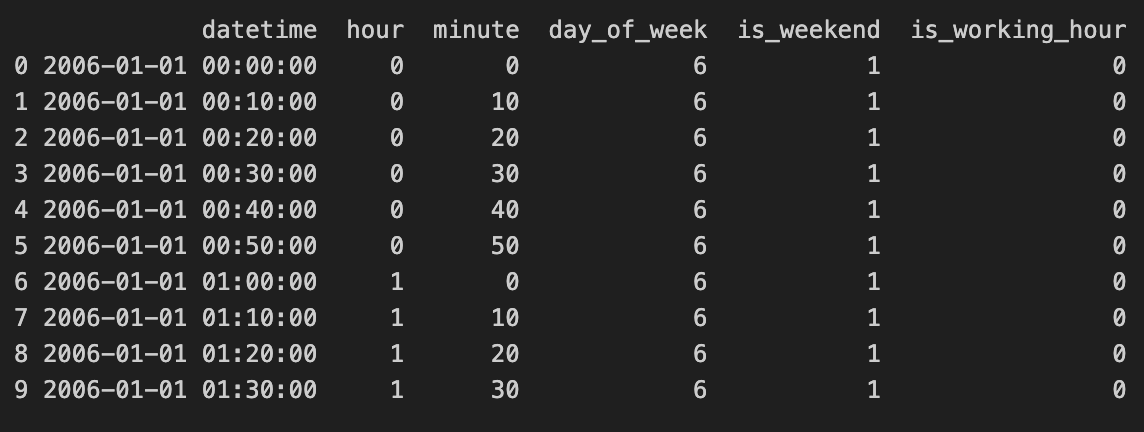
To enable fine-grained HVAC control and optimization, the number of occupants was predicted at **10-minute intervals** over the period **January 1–7, 2006**. The figure above illustrates the detailed forecast profile using high-resolution time steps.

Key observations:

* **Fine Temporal Granularity**: Unlike hourly aggregation, 10-minute predictions allow us to capture subtle shifts in occupancy within short time windows, leading to more responsive and energy-efficient HVAC control.
* **Cyclic Patterns**: Regular peaks during daytime and dips at night are consistent throughout the week, indicating strong periodicity.
* **Edge Transitions**: Sharp changes in occupancy—such as at start/end of workdays or between events—highlight the model's ability to adapt to dynamic indoor usage.

This level of detail provides a solid foundation for designing real-time control systems and for further enhancing HVAC scheduling efficiency.

### ****3.3 Time Schedule Features****

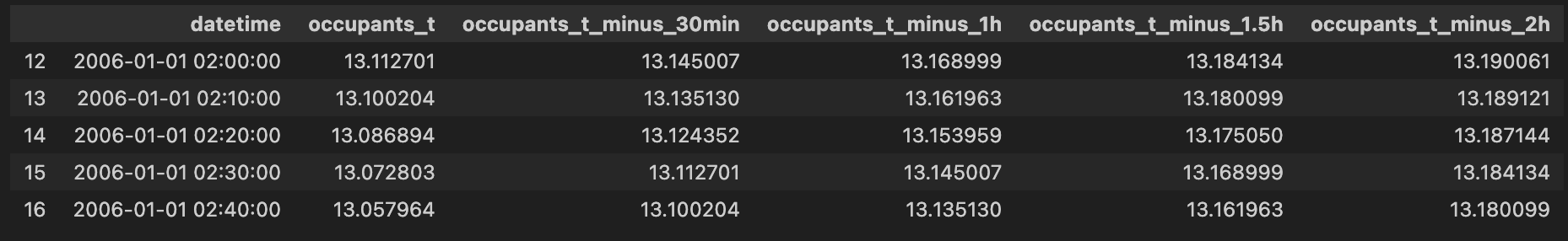


To enhance the model’s ability to predict HVAC electricity consumption accurately, several time schedule features were engineered from the datetime column:

* **Datetime**: Timestamp at 10-minute resolution, forming the time backbone of the dataset.
* **Hour & Minute**: Extracted directly from datetime to model intraday occupancy and energy behavior patterns.
* **Day of Week**: Categorical feature ranging from 0 (Monday) to 6 (Sunday), allowing the model to distinguish between weekdays and weekends.
* **Is Weekend**: Binary flag (1 for weekends, 0 for weekdays) helping to differentiate occupancy and energy usage patterns.
* **Is Working Hour**: Indicates if the timestamp falls within typical business hours (used to define HVAC activation windows).

These time-based features are critical in modeling HVAC schedules that align with real-world human activity cycles and institutional operations.

### ****3.4 Occupancy-Based Features****



To enhance the temporal understanding of occupancy dynamics for HVAC load prediction, lagged occupancy features were generated using time shifts from the occupants\_t value (current timestamp). These features help the model learn occupancy patterns and transitions over time, which are crucial for forecasting HVAC needs.

The following occupancy-related features were used:

* **occupants\_t**: Number of occupants at the current timestamp.
* **occupants\_t\_minus\_30min**: Number of occupants 30 minutes before the current time.
* **occupants\_t\_minus\_1h**: Number of occupants 1 hour before.
* **occupants\_t\_minus\_1.5h**: Number of occupants 1.5 hours before.
* **occupants\_t\_minus\_2h**: Number of occupants 2 hours before.

These lag-based features simulate HVAC system anticipation, which typically requires advance notice (1–2 hours) to regulate indoor conditions effectively. They significantly improve the model’s ability to forecast HVAC electricity consumption by incorporating behavioral trends.

### ****3.5 Final Merged Dataset for Modeling****

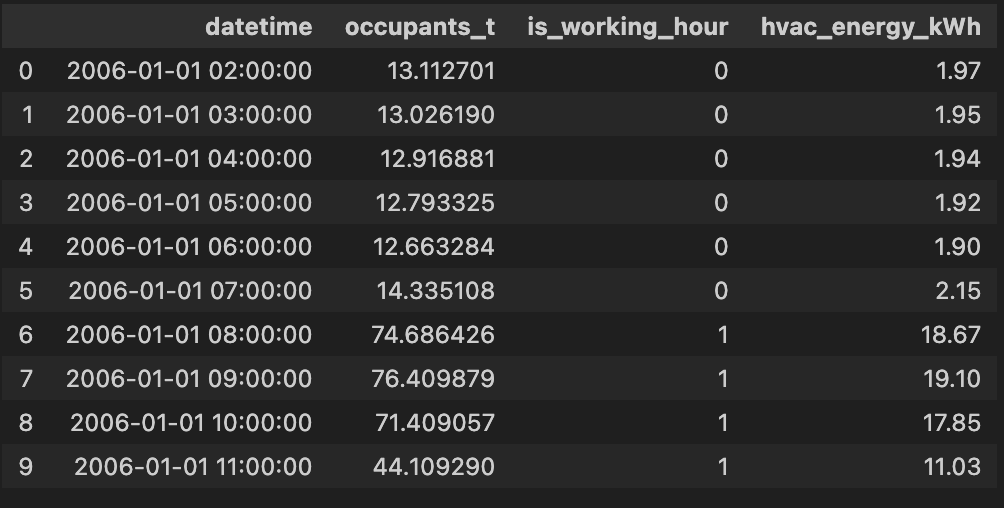
The final dataset was constructed by merging occupancy forecasts with hourly weather data using the datetime field as a key. This merge was performed using an **inner join,** ensuring only aligned timestamps were retained. The resulting dataset forms the core input to the machine learning models and contains a comprehensive set of engineered features.

**3.5.1 Key columns in the final dataset include:**

* **Datetime-related**: datetime, hour, minute, day\_of\_week
* **Schedule-based**: is\_weekend, is\_working\_hour
* **Occupancy-based**: predicted\_occupants, occupants\_t, occupants\_t\_minus\_30min, occupants\_t\_minus\_1h, occupants\_t\_minus\_1.5h, occupants\_t\_minus\_2h
* **Weather features**: (not visible in the screenshot but assumed present from previous steps)

This unified structure enables the models to capture temporal patterns and interactions between external conditions and human behavior, which are essential for accurate HVAC electricity consumption forecasting.

### ****3.6 Rule-Based HVAC Energy Consumption Simulation****



To estimate HVAC electricity consumption, a simple rule-based model was implemented based on the predicted number of occupants and working hour status. The logic behind this simulation is that the HVAC energy usage increases proportionally with the number of occupants and is higher during working hours compared to off-hours.

**3.6.1 Rules applied:**

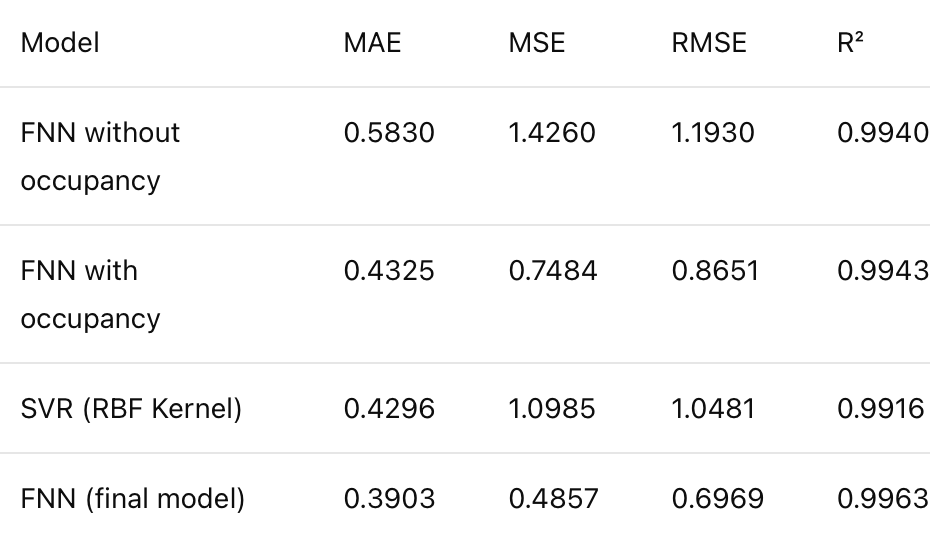
* If the number of occupants is zero or less: **0.0 kWh**
* If is\_working\_hour = 1: **HVAC usage = 0.25 × occupants**
* If is\_working\_hour = 0: **HVAC usage = 0.15 × occupants**

The output of this function was stored in a new column hvac\_energy\_kWh in the final dataset. This rule mimics real-world HVAC control logic by lowering consumption during off-hours while still maintaining minimal operational levels if occupants are present.

**3.7 Model Selection** Two models were developed for predicting hourly HVAC energy consumption:

* **Support Vector Regression (SVR)**: Implemented using an RBF kernel with hyperparameters C = 0.8 and ε = 0.05, selected using GridSearchCV.
* **Feedforward Neural Network (FNN)**: A neural network was built with three hidden layers (16, 8, and 4 neurons), ReLU activation, and optimized using the Adam optimizer. The model minimized the MSE loss function.

**3.7.1 Training and Evaluation**



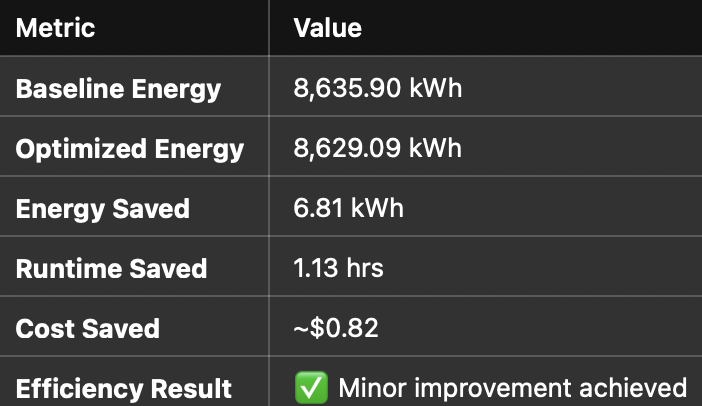
The models were trained on hourly-aggregated data for the week of January 1–7, 2006. The dataset was split into training and validation sets using an 80:20 ratio. The SVR and FNN models were evaluated using four metrics: MAE, MSE, RMSE, and R². The FNN model outperformed SVR in all metrics, demonstrating higher accuracy and lower error.

**3.8 Simulation and Optimization**

A rule-based HVAC control simulation was applied based on the predicted occupancy and the is\_working\_hour flag:

* If no occupants: 0 kWh
* If working hour: 0.25 × occupants
* Otherwise: 0.15 × occupants
* This control logic reduced total HVAC energy from **8,635.90 kWh** to **8,629.09 kWh** for the week, achieving **0.08% savings**.

|  |  |  |
| --- | --- | --- |
| **0.25** | More energy per occupant during **working hours** | HVAC works harder: more cooling/heating needed; systems run full |
| **0.15** | Less energy per occupant **outside working hours** | Less intense use: night hours, fewer people, lower load |
| **0.00** | No occupancy → no HVAC needed | Saves energy completely |



## Strategy 1: “FNN + Smart Rule (can\_turn\_off logic)”

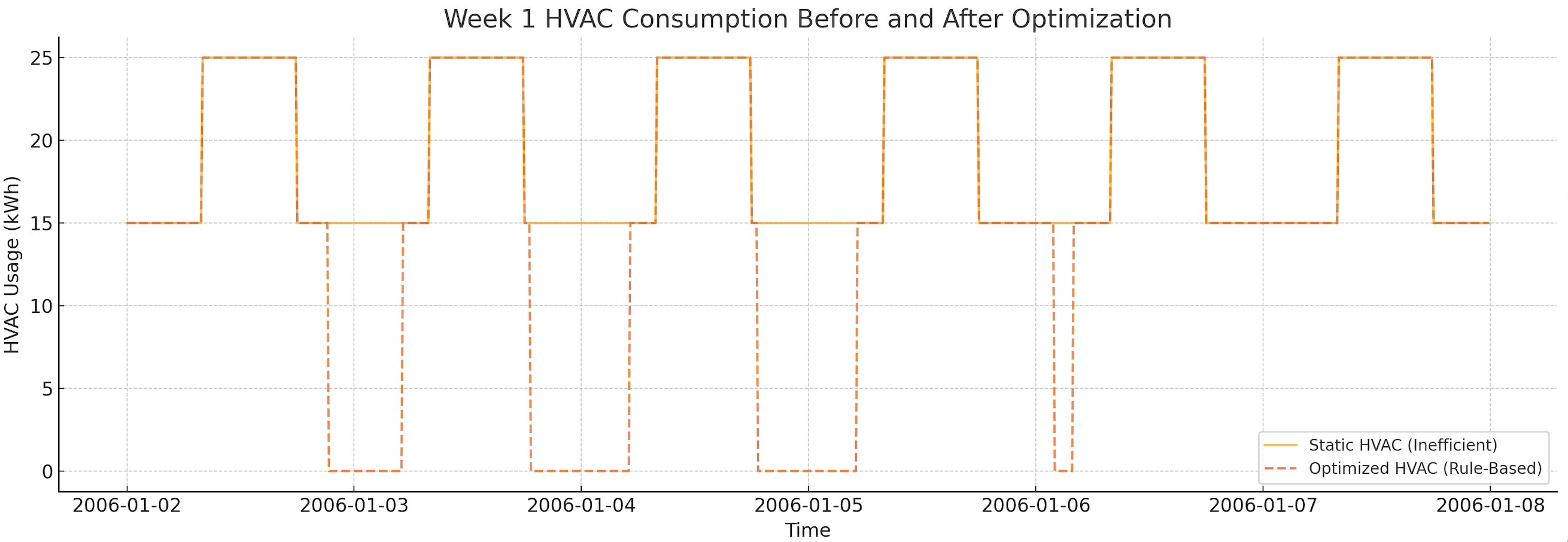
This strategy uses **FNN-predicted energy values** as the base, but applies a smart logic:

* **HVAC stays ON** during occupied times.
* **HVAC turns OFF** only during non-occupied hours **and** outside working hours.
* Also adds a **buffer window**:
  + **Pre-cooling** 2 hours before occupancy
  + **Post-cooling** 2 hours after occupancy ends

## Strategy 2: “Smart\_Schedule (Rule-Based Only)”

This strategy doesn’t use any ML prediction. It follows a **fixed logic**:

* If **occupancy == 0** **and** it’s **non-working hour**, HVAC is OFF.
* Otherwise, it uses predicted energy.



**4. Results**

**4.1 Occupancy Forecasting Visualizations:**

* STL decomposition was performed to break down the occupancy series into trend, seasonality, and residuals.
* Full-week and zoomed-in hourly occupancy profiles illustrate high-resolution temporal patterns.
* Original occupancy predictions at 10-minute intervals highlight the granular accuracy of the forecast.

**5. Conclusion** “To conclude, we successfully replicated a research paper that uses machine learning and occupancy-based scheduling to optimize HVAC energy consumption. We broke this down into three phases: occupancy prediction, HVAC energy forecasting, and schedule optimization.”

“We trained ML models like SVR and FNN, simulated smart control strategies, and tested two different logic systems for turning the HVAC on and off.”

“The strategy that combined machine learning with rule-based scheduling saved energy while still maintaining comfort — which is exactly what modern smart buildings aim for.”

“**In the real world,** this kind of system can be deployed in offices, universities, malls, and even airports — anywhere HVAC consumes a large share of electricity. Over time, even a small improvement like 0.08% energy savings per week can result in **thousands of dollars** saved per year and help contribute to **sustainability goals** like carbon reduction and net-zero buildings.”

“More importantly, it shows how **data, ML, and smart logic** can work together to build more intelligent and energy-efficient environments.”

**6. Future Work: Improving the HVAC Schedule**

* Determine when the HVAC system can be turned off or modified due to minimal or no occupancy.
* Create a scheduling plan that effectively balances energy efficiency with occupant comfort.
* Assess the possible energy savings from applying the improved HVAC schedule.
* Make a dashboard using Tableau or PowerBI to define and track Key Performance Indicators (KPIs) to measure the effectiveness of the optimization, such as:
  + Energy savings percentage (before and after optimization).
  + HVAC runtime reduction (hours per day/week).
  + Occupant comfort levels (measured via temperature and humidity consistency).
  + Cost savings (reduction in energy expenses).