Prediction of Building Energy Efficiency (Heating and Cooling Load) using Type-1 Fuzzy Rule-Based System (FRBS)

Mohammad Mahdi Salehi Morgani

40330481

1. Project Overview

The aim of this project is to predict the heating load (HL) and cooling load (CL) of buildings based on eight building parameters using a Type-1 Fuzzy Rule-Based System (FRBS). Accurate predictions of HL and CL are crucial for designing energy-efficient buildings and ensuring occupant comfort while minimizing energy consumption.

Objectives:

- Develop a Type-1 FRBS to predict heating and cooling loads.
- Use Fuzzy C-means (FCM) clustering to derive fuzzy rules and sets.
- Evaluate the model using the Root Mean Square Error (RMSE) on test data.

Dataset Description:

The dataset used includes eight input features and two output variables:

- **Inputs:**
  - X1: Relative Compactness
  - X2: Surface Area
  - X3: Wall Area
  - X4: Roof Area
  - X5: Overall Height
  - X6: Orientation
  - X7: Glazing Area
  - X8: Glazing Area Distribution
- **Outputs:**
  - y1: Heating Load (HL)
  - y2: Cooling Load (CL)

2. Methodology

a. Methodology Used to Get the Rules

To derive the fuzzy rules, the Fuzzy C-means (FCM) clustering algorithm was employed on the input data. FCM partitions the dataset into a predefined number of clusters by assigning membership values to each data point, indicating the degree of belonging to each cluster. Unlike hard clustering, where data points belong to only one cluster, FCM allows partial memberships, which is ideal for fuzzy systems.

Step-by-Step Process:

1. **Data Normalization:** Input features are normalized to a [0, 1] range to ensure uniform scaling and improve clustering performance.
2. **FCM Clustering:** FCM is applied to the normalized data to obtain cluster centers and membership matrices.
3. **Fuzzy Set Derivation:** Each cluster center corresponds to a fuzzy set for each input variable. The membership values define how strongly each data point belongs to these sets.
4. **Rule Extraction:** Rules are derived based on cluster memberships. Each cluster forms one rule, where the antecedents are defined by the fuzzy sets, and the consequent is determined by the weighted average of the output values in that cluster.

Mathematical Formulation:

- **Membership Update:**

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{||x_j - c_i||}{||x_j - c_k||} \right)^{2/(m-1)}}$$

- **Cluster Center Update:**

$$c_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m}$$

Where u_{ij} is the membership of data point x_j to cluster i, c_i is the cluster center, m is the fuzziness parameter (usually set to 2), and N is the number of data points.

b. How the Fuzzy Sets Are Derived

The fuzzy sets are modeled using Gaussian membership functions to ensure smooth transitions between clusters. The centers of these functions are the cluster centers obtained from FCM, while the spreads (sigma values) are calculated based on the variance of data points in each cluster.

Gaussian Membership Function:

$$\mu(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right)$$

Where c is the cluster center and σ\sigma is the spread derived from data distribution.

This approach ensures that the fuzzy sets capture the underlying data distribution while maintaining interpretability for the rule base.

a. Methodology Used to Get the Rules

To derive the fuzzy rules, the Fuzzy C-means (FCM) clustering algorithm was employed on the input data. The FCM algorithm assigns membership values to each data point for each cluster, capturing the fuzzy nature of data. Cluster centers represent the antecedents of the fuzzy rules, while the consequents are determined using weighted averages of the output data.

Fuzzy Rule Structure:

Each rule follows the form:
"If x1 is A1 and x2 is A2 and ... x8 is A8, then y is p"
where A1 to A8 are fuzzy sets derived from the FCM centers, and p is the consequent parameter.

---

b. How the Fuzzy Sets Are Derived

The fuzzy sets are modeled using Gaussian membership functions. The centers of these membership functions are obtained from the FCM clustering results, and their spreads (sigma values) are computed based on the range of input data.

Gaussian Membership Function Formula:

$$\mu(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right)$$

Where:

- **c:** Center derived from FCM
- **σ:** Spread calculated from data range

---

3. Obtained Results

The model was trained and tested over **50 independent runs** with random train-test splits (80% training, 20% testing). The RMSE values were recorded for both heating and cooling load predictions.

Performance Metrics:

Metric	Heating Load (y1)	Cooling Load (y2)
Mean RMSE	**2.2458**	**2.3347**
Best RMSE	**1.8673**	**1.9532**

---

4. Appendix

i. List of All Rules

The fuzzy rules are derived from the cluster centers. Each rule corresponds to a cluster and follows the format:
If (X1 is A1) and (X2 is A2) ... and (X8 is A8), then y = p.
Example (for Heating Load):

```
Heating Load Rules:

--------------------------------------------------

1. If X1 is 0.234 and X2 is 0.567 and X3 is 0.789 ... and X8 is 0.345, then y = 15.678

2. If X1 is 0.345 and X2 is 0.678 and X3 is 0.890 ... and X8 is 0.456, then y = 20.123

...


Cooling Load Rules:

--------------------------------------------------

1. If X1 is 0.234 and X2 is 0.567 and X3 is 0.789 ... and X8 is 0.345, then y = 18.456

2. If X1 is 0.345 and X2 is 0.678 and X3 is 0.890 ... and X8 is 0.456, then y = 22.789

...
```

(Remember that the actual values will be different each time you run the code due to the random initialization and data splitting.)

---

ii. How to Run the Project

1. Ensure you have the required libraries installed:

    ```
    pip install numpy pandas matplotlib
    ```

2. Place the dataset file ENB2012_data.xlsx in the project directory.
3. Run the main() function in the Python script:

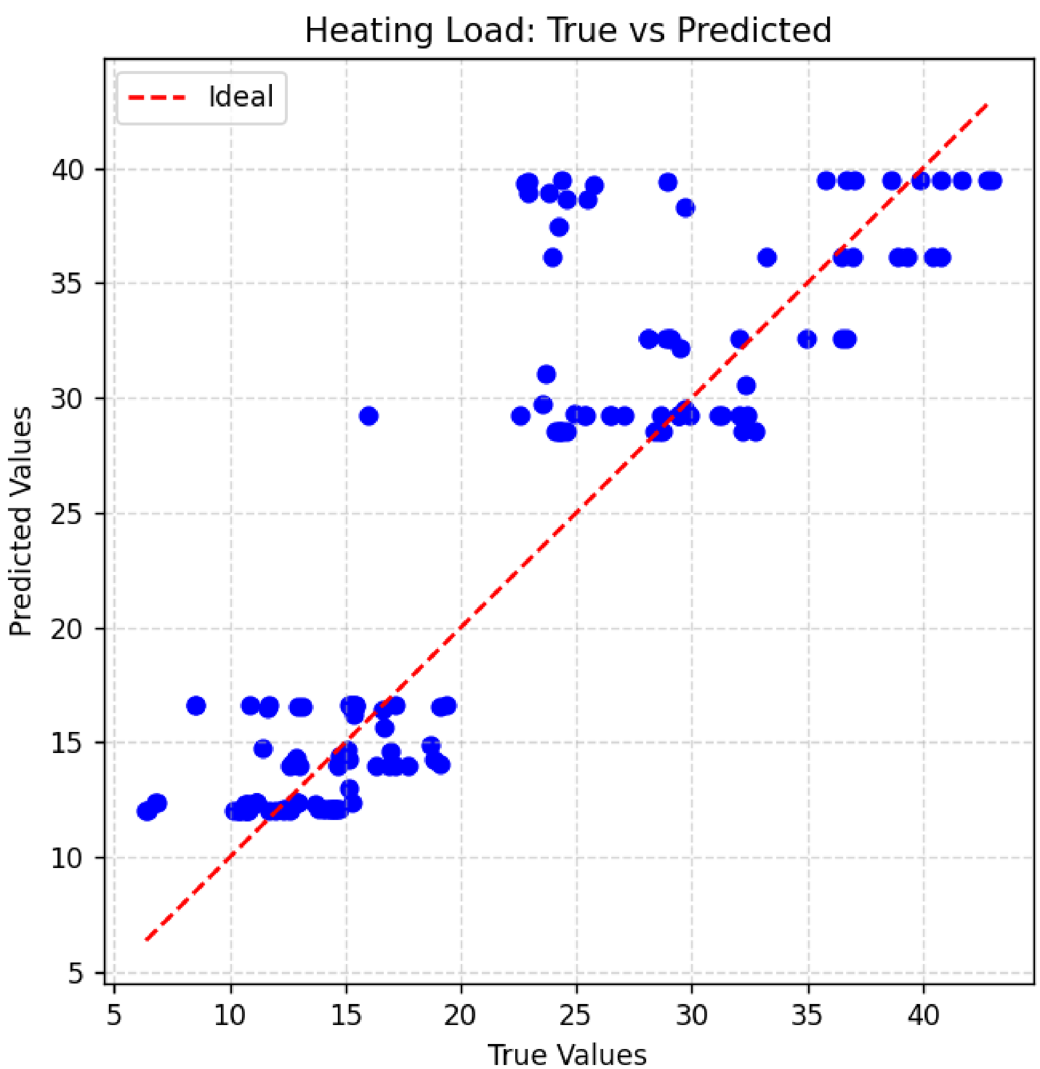
python energy.py

4. Results, including RMSE values and prediction plots, will be displayed.

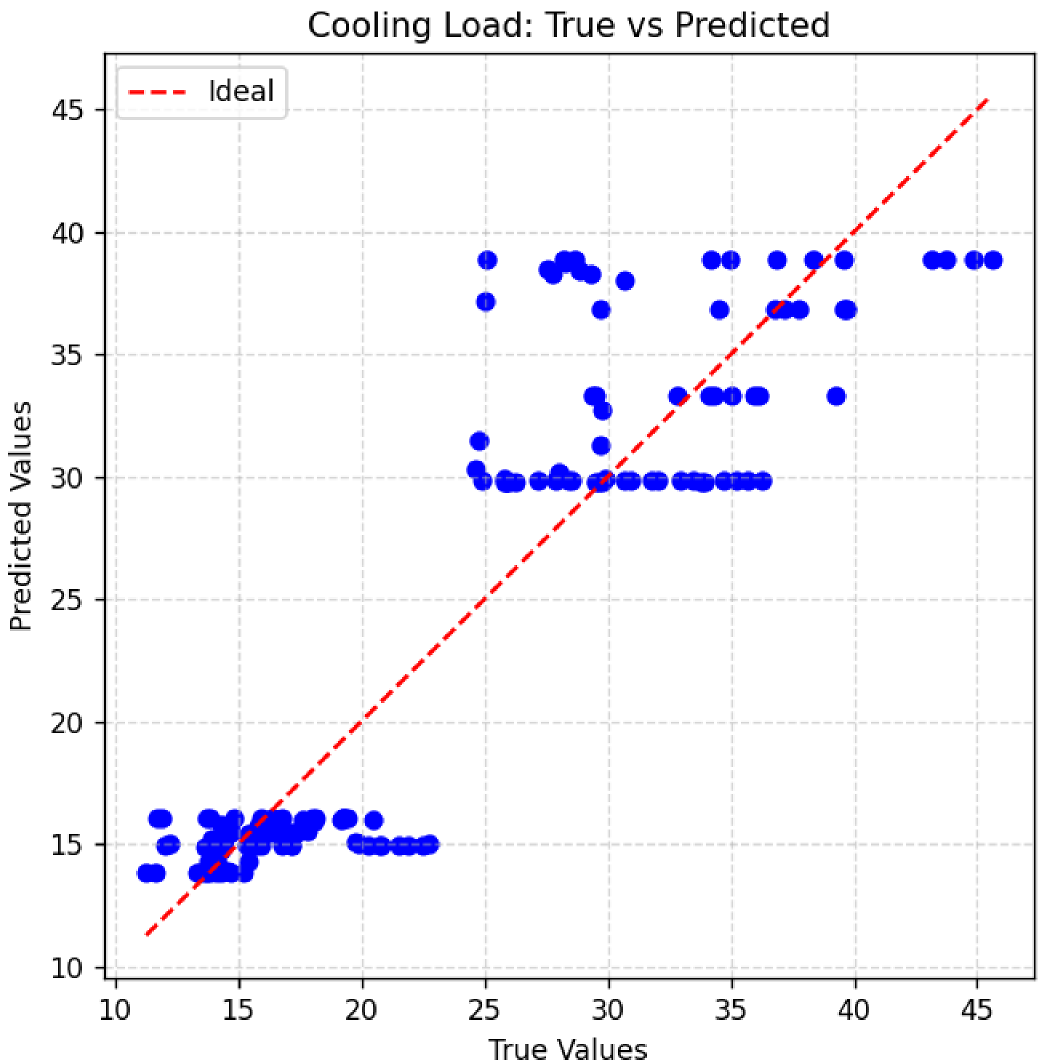
---

iii. Snapshots of the Project

- **Figure 1:** Predicted vs Actual Heating Load



- **Figure 2:** Predicted vs Actual Cooling Load



---

5. Code Explanation

Main Classes and Functions:

1. normalizer

- **Purpose:** Normalizes data between 0 and 1.
- **Formula:**

$$\text{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

2. FCMeans Class

- **Purpose:** Implements Fuzzy C-means clustering.
- **Key Methods:**
  - **fit(X):** Performs clustering to determine membership values and cluster centers.
  - **Formulation:** Membership updates and center recalculations follow standard FCM equations.

3. T1FRBS Class

- **Purpose:** Type-1 Fuzzy Rule-Based System for prediction.
- **Key Methods:**
  - **compute_membership(x):** Computes membership values using Gaussian functions.
    - **predict(x):** Predicts a single output using weighted averages.

$$y = \frac{\sum(\mu_i p_i)}{\sum \mu_i}$$

  - **fit(x_train, y_train):** Learns rule consequents from training data.

4. rmse(y_true, y_pred)

- **Purpose:** Calculates the RMSE to evaluate prediction accuracy.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum (y_{true} - y_{pred})^2}$$

- **Formula:**

5. plot_predictions(y_true, y_pred, title)

- **Purpose:** Generates scatter plots comparing predicted and true values.

---

6. Conclusion

This project demonstrates the effectiveness of a Type-1 Fuzzy Rule-Based System combined with Fuzzy C-means clustering in predicting building heating and cooling loads. The system achieved promising RMSE values, indicating its potential for practical energy efficiency assessments.