ENERGY CONSUMPTION OPTIMIZATION IN INDUSTRIAL PLANTS

USING UCI MACHINE LEARNING REPOSITORY - ENERGY EFFICIENCY DATA SET

PROBLEM STATEMENT

Develop an AI model to optimize energy consumption in an industrial plant by anticipating and regulating energy demand in real time, lowering costs and increasing efficiency.

DATASET OVERVIEW

Source: UCI Machine Learning Repository

Dataset Details:

- **Features:** 8 input variables representing building characteristics (e.g., relative compactness, surface area, wall area).
- **Target Variables:** Heating Load (HL) and Cooling Load (CL).
- **Dataset Size:** 768 samples.
- **File Format:** CSV (ENB2012_data.xlsx)

PREPROCESSING STEPS

- **Data Cleaning:** Handling missing values and outliers.
- **Normalization:** Applied MinMax scaling to features.
- **Train-Test Split:** Divided dataset into 80% training and 20% testing.

MODEL SELECTION AND TRAINING

- **Models Used:**
  - Ridge Regression
  - Lasso Regression
  - Random Forest
  - Gradient Boosting
  - Support Vector Regression (SVR)
- **Training Process:** Applied cross-validation and randomized search for hyperparameter tuning.

HYPERPARAMETER OPTIMIZATION

- **Method:** Randomized Search with Cross-Validation (3-fold)
- **Objective:** Find the best model configuration to minimize Mean Squared Error (MSE).
- **Key Hyperparameters Tuned:**
  - Regularization strengths (Ridge, Lasso)
  - Number of estimators, max depth (Random Forest, Gradient Boosting)
  - C, gamma (SVR)

RESULTS

Model	Mean Squared Error (MSE)	R^2 Score
Heating Load		
Ridge Regression	9.21	0.91
Lasso Regression	12.42	0.88
Random Forest	0.24	1.00
Gradient Boosting	0.26	1.00
SVR	7.97	0.92

refer to project2.ipynb

RESULTS

Model	Mean Squared Error (MSE)	R^2 Score
Cooling Load		
Ridge Regression	9.94	0.89
Lasso Regression	13.75	0.85
Random Forest	3.15	0.97
Gradient Boosting	2.29	0.98
SVR	10.62	0.89

refer to project2.ipynb

RESULTS (HYPERPARAMETER OPTIMIZATION)

Model	Best Parameters	Mean Squared Error (MSE)	R^2 Score
Ridge Regression	alpha: 0.1	9.16	0.91
Lasso Regression	alpha: 0.001	9.16	0.91
Random Forest	n_estimators: 100, max_features: 'log2', max_depth: 20	0.31	0.997
Gradient Boosting	n_estimators: 200, max_depth: 3, learning_rate: 0.3	0.21	0.998
SVR	kernel: 'rbf', gamma: 1.0, C: 100.0	1.71	0.98

refer to project2.ipynb

THANK YOU VERY MUCH!