NAAN MUDHALVAN PROJECT 4

Topic : Measure Energy Consumption

Problem Definition:

Develop a software program to accurately measure and monitor energy consumption in residential and commercial buildings, providing real-time and historical data to enable energy efficiency improvementsand cost savings.

Design Thinking:

Building a script to automate data collection, analysis, and visualization processes involves several stepsand the use of appropriate programming languages and libraries. Here's a general guideline on how to create such a script:

Choose a Programming Language:

Select a programming language that is suitable for your task and the libraries available for data manipulation, analysis, and visualization. Python is a popular choice for these tasks due to its extensiveecosystem of libraries (e.g., pandas, NumPy, Matplotlib, Seaborn, Plotly) and ease of use. Other languages like R, Julia, or JavaScript may also be suitable depending on your specific needs.

Define the Workflow:

Clearly define the workflow of your automation script. Break it down into distinct stages: data collection,data preprocessing, analysis, and visualization.

Data Source:

Identify the data sources (e.g., APIs, databases, CSV files) from which you will collect data.Write code to

programmatically fetch or import data from these sources.

Consider scheduling data collection tasks using tools like cron jobs or Python schedulers to automateperiodic data updates.

The dataset for the project is given in the link below:

https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption

Data Preprocessing and Extraction:

Develop code to clean and preprocess the raw data.

Transform and reshape data as needed (e.g., feature engineering,

normalization, encoding).Store the cleaned data in a structured

format (e.g., CSV, Excel, database) for analysis.

Extracting relevant features and metrics from energy consumption data is essential for understandingpatterns, optimizing energy usage, and making informed decisions.

Maintain clear documentation of the extracted features, their calculations, their interpretations. Rememberthat the choices of features and metrics will depend on your specific energy consumption analysis goals, whether it's forecasting, anomaly detection, or optimization.

Model Development:

Implement the data analysis tasks that are relevant to project. This may include statistical analysis,machine learning, or time series analysis.

Utilize libraries such as pandas, NumPy, or scikit-learn for data analysis and modeling.

Develop functions or scripts that automate the analysis processes, including model training andevaluation.

Visualization:

Use libraries like Matplotlib, Seaborn, or Plotly to create visualizations.

Define functions or scripts that generate the necessary charts, graphs, and plots to convey insights.Customize

visualizations to make them informative and visually appealing.

Consider saving visualizations as image files or embedding them in reports or dashboards.

Automation:

Create a main automation script that orchestrates the entire workflow.

Sequence the steps for data collection, preprocessing, analysis, and visualization within this script. Implement error

handling and logging to handle exceptions or issues that may arise during automation.

Utilize programming constructs like loops and conditional statements to make the script flexible androbust.

Use functions or modular programming to keep the code organized and maintainable.