POWER CONSUMPTION ANALYSIS FOR HOUSEHOLDS

Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. This crucial phase establishes project parameters, identifies key team members, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a wellorganized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

Activity 1: Define Problem Statement

Problem Statement: The problem statement for power consumption analysis in households involves examining patterns and drivers of electricity usage to optimize efficiency and reduce costs. This analysis aims to leverage data to develop predictive models that accurately forecast energy demand based on factors such as household size, appliance usage, time of day, and external influences like weather

Ref. template:

Power Consumption Problem Statement Report: click here

Activity 2: Project Proposal (Proposed Solution)

The proposed solution for power consumption analysis in households involves deploying advanced data analytics and machine learning techniques to develop accurate predictive models. By leveraging historical electricity usage data and integrating variables such as household size, appliance usage patterns, weather conditions, and time-of-day factors, these models will forecast future energy demands effectively.

Ref. template:

Power Consumption Project Proposal Report: click here

Activity 3: Initial Project Planning

Initial project planning for power consumption analysis in households involves defining objectives, assembling relevant data sources (such as electricity usage records, weather data, and household demographics), selecting appropriate analytical methods (including statistical analysis and machine learning algorithms), and establishing milestones for model development and validation. Key tasks include data preprocessing to ensure quality and compatibility, feature selection to identify influential factors on energy consumption, and initial model training and evaluation.

Ref. template:

Power Consumption Initial Project Planning Report : click here

Milestone 2: Data Collection and Preprocessing Phase

The Data Collection and Preprocessing Phase involves executing a plan to gather relevant loan application data from Kaggle, ensuring data quality through verification and addressing missing values. Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysis and machine learning model development.

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

The dataset for "Power Consumption Analysis For Households" is sourced from Kaggle. The download data set is not suitable for training the machine learning model as it might have so much randomness so we need to clean the dataset properly in order to fetch good results. This activity includes the following steps

• Handling missing values

Ref. template:

Power Consumption Data Collection Report: click here

Activity 2: Data Quality Report

The Data Quality Report for power consumption analysis of households indicates that the dataset is comprehensive, containing records from a diverse range of households over a specified period. The data appears to be accurate, with minimal missing values or outliers identified through initial exploratory analysis..

Ref. template:

Power Consumption Data Quality Report: click here

Activity 3: Data Exploration and Preprocessing

 data exploration and preprocessing in power consumption analysis for households, the initial step involves loading and inspecting the dataset to understand its structure, variables, and any missing values. Exploratory data analysis (EDA) techniques such as summary statistics, distribution plots, and correlation matrices are then employed to uncover patterns and relationships among variables like electricity usage, time of day, and household characteristics.

Ref. template:

Power Consumption Data Exploration and Preprocessing Report: click here

Milestone 3: Model Development Phase

In the model development phase for power consumption analysis of households, the focus is on selecting appropriate machine learning algorithms such as regression models, time series forecasting methods, or clustering techniques, depending on the specific analysis goals. Models are trained using the training set and tuned using the validation set to optimize parameters and improve accuracy. Finally, the model's performance is evaluated on the test set, assessing metrics like RMSE (Root Mean Squared Error) or MAE (Mean Absolute Error) to validate its predictive capability and suitability for understanding household power consumption patterns.

Activity 1:Feature Selection Report

Creating a feature selection report for power consumption analysis in households involves several steps. The goal is to identify the most relevant features that influence power consumption. Here's a structured approach to produce this report

Ref. template:
Power Consumption Model Selection Report click here

Activity 2: Model Selection Report
The Model Selection Report for power consumption analysis of households involves evaluating various machine learning models such as linear regression, decision trees, random forests, and possibly more advanced techniques like gradient boosting or neural networks. Each model is assessed based on performance metrics such as RMSE, MAE, and R-squared to determine its accuracy in predicting household power consumption patterns.
Ref. template:
Power Consumption Model Selection Report : click here

Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

The initial model training for power consumption analysis of households involves splitting the dataset into training and testing sets, typically 80% for training and 20% for testing, ensuring randomization for unbiased evaluation. A machine learning algorithm, such as a Random Forest regressor, is chosen for its ability to capture complex relationships in data. The model is trained on the training set, optimizing parameters like the number of trees and maximum depth to minimize mean squared error (MSE) and mean absolute error (MAE).

Ref. template:

Power Consumption Model Development Phase Template: click here

Milestone 4: Model Optimization and Tuning Phase

In the Model Optimization and Tuning phase for power consumption analysis of households, the focus is on refining the selected machine learning model to improve its predictive performance. Techniques such as hyperparameter tuning using methods like grid search or randomized search are employed to find the optimal values for parameters such as the number of trees (for ensemble methods like Random Forests), learning rate (for Gradient Boosting), or regularization strength (for linear models). Cross-validation is utilized to assess model performance across different subsets of data, ensuring generalizability.

Ref. template:

Power Consumption Model Development Phase Template: click here

Milestone 5: Project Files Submission and Documentation

For project file submission in Github, Kindly click the link and refer to the flow
For the documentation, Kindly refer to the link.: click here

Milestone 6: Project Demonstration

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.