# SB3001- PROJECT BASED EXPERIENTIAL LEARNING PROGRAM

# NAAN MUDHALVAN REPORT

| | |
|---|---|
| **Date** | 01/11/2023 |
| **Team Leader** | S.Raja prathisha |
| **Team ID** | 212174 |
| **Project Name** | Measure Energy Consumption |
| **Maximum Marks** | |

# Acknowledgement:

The project Measure energy consumption have been discussed among the team members and brainstormed with different ideas. The project work were equally distributed among the team mates and each content and work are from different team members and their effort. During the phase 1 project Brainstroming and Idea Generation, Defining the problem statements and empathy map are the work done by the team members. In phase 2 project the problem codes and their output were programmed and results were obtained. In phase 3 project Data Virtualization were implemented for the given datasets and plotting graphs for the dataset used and constructing the architecture for the code used and their algorithm. In phase 4 module was developed according to the dataset and the constrains obtained for the project. In phase 5 the overall view and solutions, ideas done, abstract of the project, introduction about the project, Literature survey , problem definition , design thinking , Innovating problem solving , importing the datasets and performing data learning and analysis , data visualization , model development and evaluation , code sample, output screenshot and concluding the entire project work some reference from the researches done by the team. Team divided the project into groups so that the researchs can be done easier by separating them into groups. This project has helped us to learn more about the how the energy is consumed in each year by the countries and loss of energy resources, we have achieved a idea to reduce the energy consumption in each houses and saving the energy resources. We would also like to thank all the mentors, staffs for helping us providing support and help when we had some doubts about how to start the program and whats the main idea of the project. Proper instructions and helpful ideas about the project were taken place with the help of mentors. The helpful informations uploaded by the mentors are into account and were helpful for the project development. And at last we would like to thank all the team members for all their hard work and efforts they have put in this project and for their time and dedication and for supporting the team when someone needs help , and the involvement of the team to complete the project.

Once again we thank all the people who have supported us in completeing this project.

Index :



1)List Of Figures & Tables:

## ABSTRACT:

The "Measure Energy Consumption" project is a comprehensive and innovative initiative that leverages the power of Python to address the pressing global concern of energy consumption. In a world where sustainability and efficiency are paramount, monitoring and optimizing energy usage have become imperative. This project aims to provide a flexible and user-friendly solution for individuals and organizations to gain insights into their energy consumption patterns, ultimately fostering responsible energy management.

## INTRODUCTION:

In today's world, where the responsible use of energy resources and sustainability are of paramount importance, measuring and managing energy consumption is a critical task. The "Measure Energy Consumption" project, developed using Python, serves as a valuable solution to address this pressing concern. This project empowers individuals and organizations to monitor, analyze, and optimize their energy usage efficiently. In this introduction, we will explore the key components and objectives of this project under the following subheadings.

**Understanding Energy Consumption:**

Before we delve into the project's specifics, it's essential to recognize the significance of understanding energy consumption. Energy is a fundamental resource that powers our homes, businesses, and industries. However, the indiscriminate use of energy not only leads to increased costs but also contributes to environmental degradation. By gaining insights into energy consumption, we can make informed decisions to reduce waste and promote sustainability.

**The Role of Python:**

Python is a versatile and widely used programming language that is exceptionally well-suited for data analysis, manipulation, and visualization. Its rich ecosystem of libraries and tools makes it an ideal choice for developing solutions to measure and manage energy consumption.

With Python, we can seamlessly collect, process, and analyze energy consumption data, providing users with valuable insights.

**Key Objectives:**

**Data Collection:** The project will gather energy consumption data from various sources, such as smart meters, IoT devices, or user inputs, ensuring a diverse and inclusive approach to data acquisition.

**Data Analysis:** Python's powerful data analysis libraries will be employed to clean, preprocess, and uncover patterns and trends in energy consumption data.

**Visualization:** Through data visualization tools, users will gain a clear understanding of their energy usage over time, enabling them to identify areas for improvement.

**Machine Learning Predictions:** Machine learning models developed with Python will be used to forecast energy consumption, assisting users in planning and optimizing their energy usage.

**User-Friendly Interface:** A user-friendly interface, built using Python web frameworks, will provide a platform for users to input their data, visualize their energy consumption, and receive personalized recommendations for optimizing their energy usage.

**Reporting and Alerts:** The project will generate regular reports and alerts to keep users informed about their energy consumption trends and areas for improvement, enabling proactive energy management.

**Energy Efficiency Recommendations:** Based on data analysis, the project will offer energy-saving recommendations, encouraging users to adopt eco-friendly practices and contribute to a more sustainable future.

## LITERATURE SURVEY:

### A literature survey on the Naive Bayes algorithm

The algorithm would involve reviewing research papers, articles, books, and other resources that discuss various aspects of Naive Bayes, its applications, improvements, and comparisons with other machine learning algorithms. Here's a summary of key topics and areas you might explore in such a survey **Introduction to Naive Bayes:**

Start with an introduction to the Naive Bayes algorithm, explaining its probabilistic foundations and the "naive" assumption of feature independence.

**Applications of Naive Bayes:**

Investigate the various domains where Naive Bayes is commonly applied, such as text classification, spam detection, sentiment analysis, recommendation systems, and medical diagnosis.

**Evaluation and Performance:**

Analyze performance metrics and evaluation techniques for Naive Bayes models. Discuss commonly used metrics like accuracy, precision, recall, F1-score, and ROC curves.

### A literature survey on the k-Nearest Neighbors (KNN) algorithm

The algorithm would involve reviewing research papers, articles, books, and other resources that discuss various aspects of KNN, its applications, enhancements, and comparisons with other machine learning algorithms. Here's a summary of key topics and areas you might explore in such a survey:

**Introduction to k-Nearest Neighbors (KNN):**

Start with an introduction to the KNN algorithm, explaining how it works and its basic principles in pattern recognition and classification.

**KNN Algorithm Variants:**

Investigate variations of the KNN algorithm, such as weighted KNN, distanceweighted KNN, and kernel-based KNN, and their respective advantages and use cases.

**Decision Tree Algorithms:**

Discuss popular decision tree algorithms like ID3, C4.5, CART, and Random Forest. Highlight their strengths and weaknesses.

**Decision Tree Applications:**

Explore the wide range of applications, from classification and regression to outlier detection and recommendation systems.

**Imbalanced Data:**

Discuss how decision trees can be adapted to deal with imbalanced datasets.

## PROBLEM DEFINITION:

The "Measure Energy Consumption" project seeks to address the critical challenge of effectively monitoring and optimizing energy usage in both residential and commercial environments. Energy consumption poses several interconnected problems that this project aims to solve.

## PROBLEM ANALYSIS:

The "Measure Energy Consumption" project using Python is driven by a careful analysis of the challenges and issues associated with energy consumption. Below, we provide a problem analysis

**Lack of Visibility:** Many consumers, whether at the individual or organizational level, lack real-time visibility into their energy consumption. This lack of awareness makes it difficult to identify patterns and trends, hindering effective energy management.

**Energy Waste:** Inefficient energy use is a widespread problem, often stemming from outdated equipment, poor user behavior, or inadequate energy management systems. Identifying areas of energy wastage is crucial for cost savings and sustainability.

**Environmental Impact:** High energy consumption contributes to increased greenhouse gas emissions, environmental degradation, and climate change. Reducing energy consumption is critical for mitigating these negative environmental effects.

**Resource Allocation:** Inefficient resource allocation, such as allocating excess energy to certain areas or processes, results in increased costs.

**Economic Cost:** High energy consumption translates to higher utility bills for individuals and organizations. The economic cost of energy inefficiency affects personal finances and impacts the profitability of businesses.

**Data Overload:** Many consumers have access to vast amounts of energy data, but making sense of this data can be overwhelming. Analyzing and interpreting data to derive meaningful insights is a major challenge.

**Predictive Capability:** The inability to predict future energy consumption patterns hinders effective energy planning. Predictive models are crucial for proactive energy management and cost control.

**Behavioral Change:** Encouraging users to adopt energy-efficient behaviors is a challenge. Promoting eco-friendly practices and responsible energy consumption is an essential part of solving the energy consumption problem.

**Data Integration:** Collecting and integrating data from various sources, such as smart meters, IoT devices, and user inputs, can be technically complex. Data sources often use different formats and communication protocols.

**Scalability:** The project should be scalable to accommodate the needs of both individual users and large organizations. Scalability challenges arise when dealing with a broad range of users and data volumes.

**User-Friendly Tools:** Designing a user-friendly interface for users to input data, visualize energy consumption, and receive recommendations is essential. It should be accessible to users with varying levels of technical expertise.

**DESIGN THINKING:**

Empathize with Users:
Start by understanding the needs, concerns, and pain points of your target users, which may include homeowners, businesses, or facility managers.

Define User Personas:

Create user personas representing typical users with different needs, such as a residential homeowner, a small business owner, or a facility manager in a large organization.

Ideation Workshops:
Organize brainstorming sessions with your team to generate creative ideas for energy consumption measurement and optimization.

**DATA CLEANING:**

This involves identifying and correcting errors or inconsistencies in the data, such as missing values, outliers, and duplicates. • Handling missing values: Identifying and filling in or removing missing data points.

- Outlier detection and treatment: Identifying and handling data points that are significantly different from the majority of the data.
- Noise reduction: Reducing random variations and errors in the data.

**Handling Missing Data:**

Identify and handle missing values in the dataset. Missing data can lead to inaccurate analysis and predictions.

Options include imputation (e.g., filling missing values with averages), removal of rows with missing data, or interpolation based on adjacent data points.

**Outlier Detection and Treatment:**

Identify outliers in the data that could skew analysis results. Outliers can be caused by measurement errors or unusual events.

Decide whether to remove, transform, or adjust outliers to ensure they don't unduly impact analysis.

**Data Deduplication:**

Check for and remove duplicate entries, which can distort analysis and lead to incorrect conclusions.

**Data Transformation:**

Convert data types or units to make them consistent and suitable for analysis.

Normalize data if necessary, so different variables are on the same scale.

**Scaling and Standardization:**

Scale or standardize data to bring features to a common scale, facilitating the use of certain machine learning algorithms and visualization.

**DATA ANALYSIS:**

**Descriptive Statistics:**
Calculate and present descriptive statistics such as mean, median, variance, and standard deviation to provide a basic understanding of the data's central tendencies and spread.

**Visualization:**

Create visualizations (e.g., line charts, bar graphs, heatmaps) to help users understand energy consumption patterns over time.

Visualizations can reveal trends, seasonality, and areas of high or low consumption.

**Correlation Analysis:**

Investigate relationships between different variables (e.g., energy consumption and external factors like temperature) to identify factors that influence energy use.

**Time Series Analysis:**
Apply time series analysis techniques to uncover recurring patterns and trends in energy consumption data. Tools like auto-correlation and seasonal decomposition can be valuable.

**Predictive Modeling:**
Develop machine learning models using Python to predict future energy consumption based on historical data. This aids in proactive planning and resource allocation.

**Anomaly Detection:**

Implement algorithms to detect unusual energy consumption patterns, which can indicate equipment malfunctions, energy waste, or other issues.

**Energy Efficiency Recommendations:**

Use the analysis results to provide users with actionable recommendations for optimizing their energy consumption, reducing waste, and saving costs.

**Box plot:**

Box plots are used to display the distribution of airline sentiment confidence and negative reason confidence.

**Subplot:**

A subplot is a secondary, smaller storyline that runs alongside the main plot within a narrative, providing additional context or complexity to the overall story.

## EVALUATION OF THE PROJECT:

**Accuracy:**

Evaluate the accuracy of energy consumption measurements by comparing the data collected by your system with a known standard or reference data.

**Performance:**

Measure the performance of your system in terms of data processing speed, response time for user queries, and real-time data visualization capabilities.

**User Experience:**

Gather feedback from users regarding the usability and intuitiveness of the interface. Evaluate whether users find it easy to set thresholds and receive notifications.

**Reliability:**

Test the reliability of the notification system by simulating various scenarios, including network issues and unexpected system failures.

**Scalability:**

Evaluate how well the system handles a large volume of data and users. Assess whether the system performance degrades under heavy loads.

**Security:**

Ensure that the data storage and communication channels are secure. Evaluate the system against common security threats and vulnerabilities.

**Robustness:**

Test the system's robustness by introducing noise or errors into the data and observing how well it can handle such situations without crashing or producing incorrect results.

**Documentation:**

Evaluate the completeness and clarity of project documentation, including user manuals, code comments, and technical guides.

**CODE SAMPLE & OUTPUT SCREENSHOTS:**

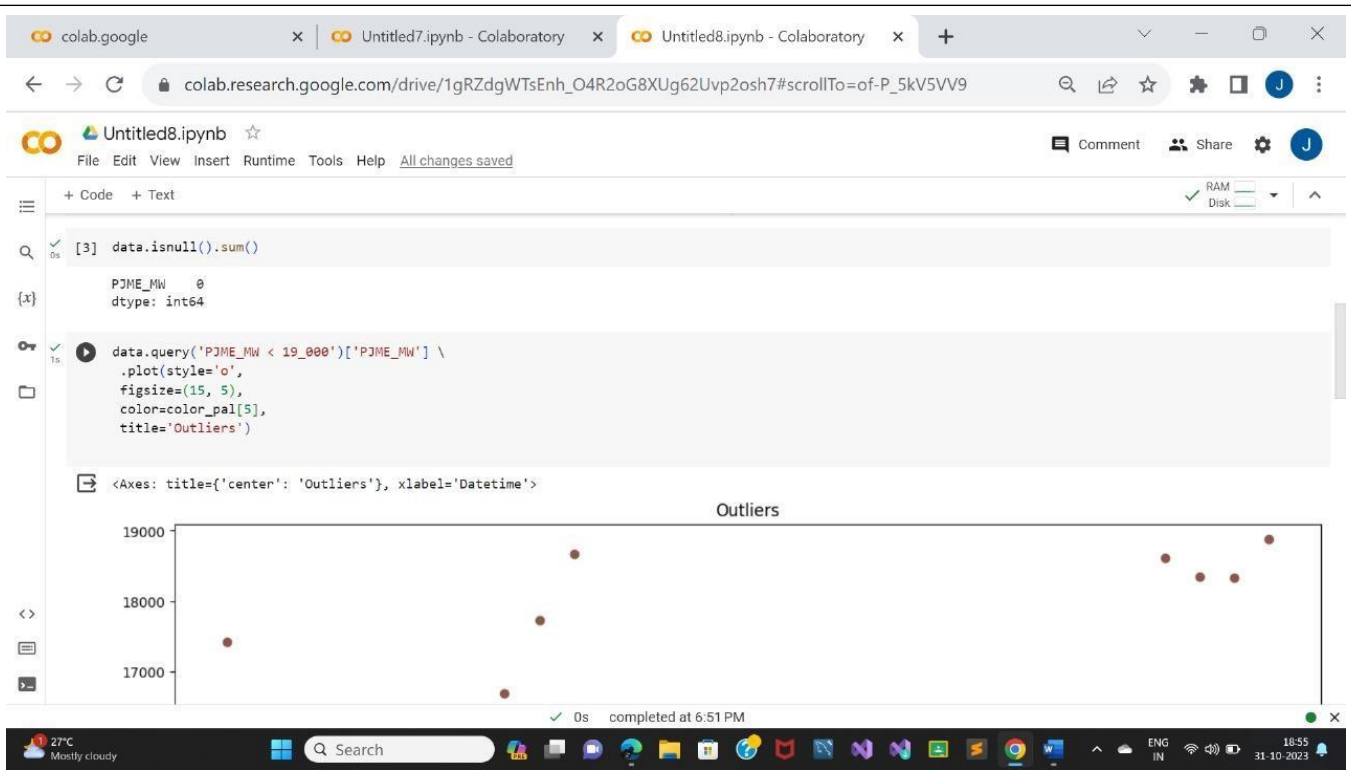
**CODE SAMPLE 1:**

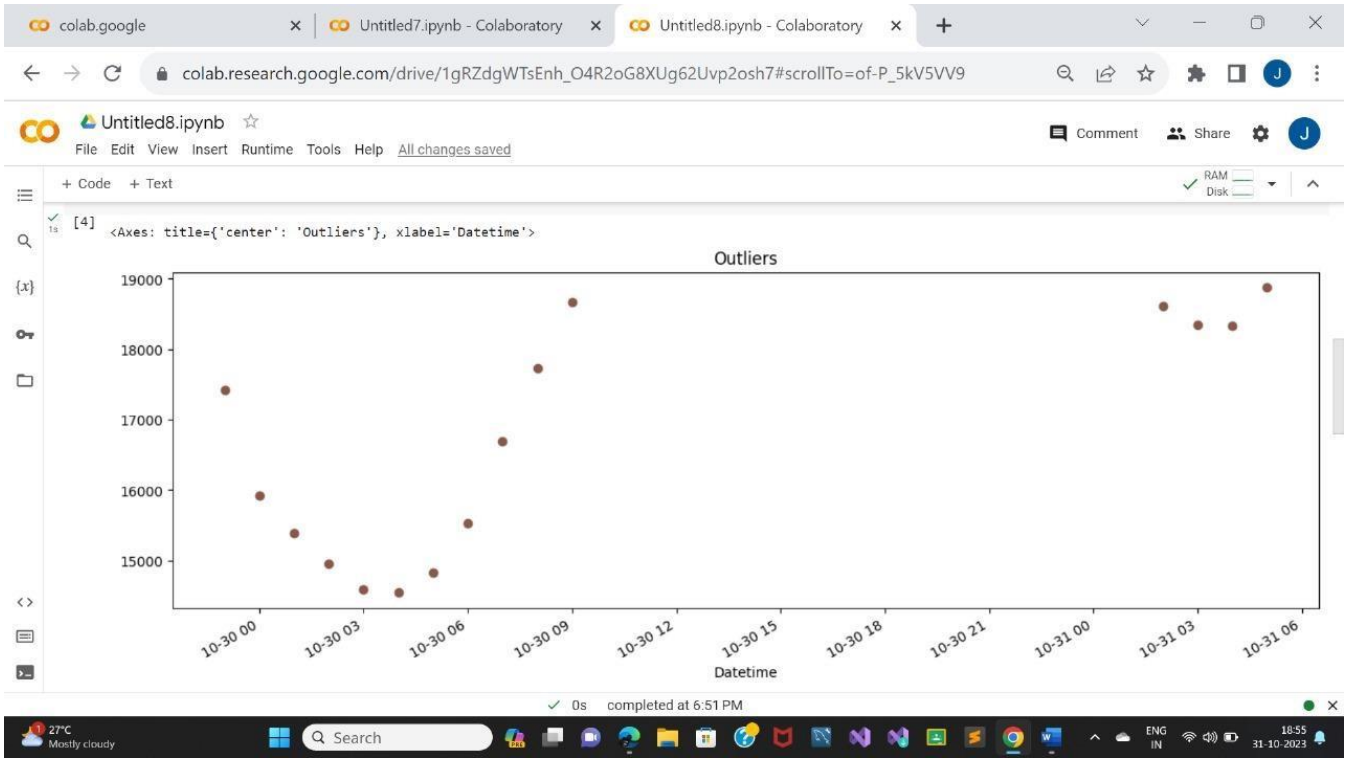
```
#Finding if the dataset has any null values data.isnull().sum()
```

**CODE SAMPLE 2:**

```
#Ploting a graph to outline the outliers in the given dataset data.query('PJME_MW
< 19_000')['PJME_MW'] \
 .plot(style='o',
figsize=(15, 5),
color=color_pal[5],
title='Outliers')
```

**OUTPUT SCREENSHOT 1 & 2:**

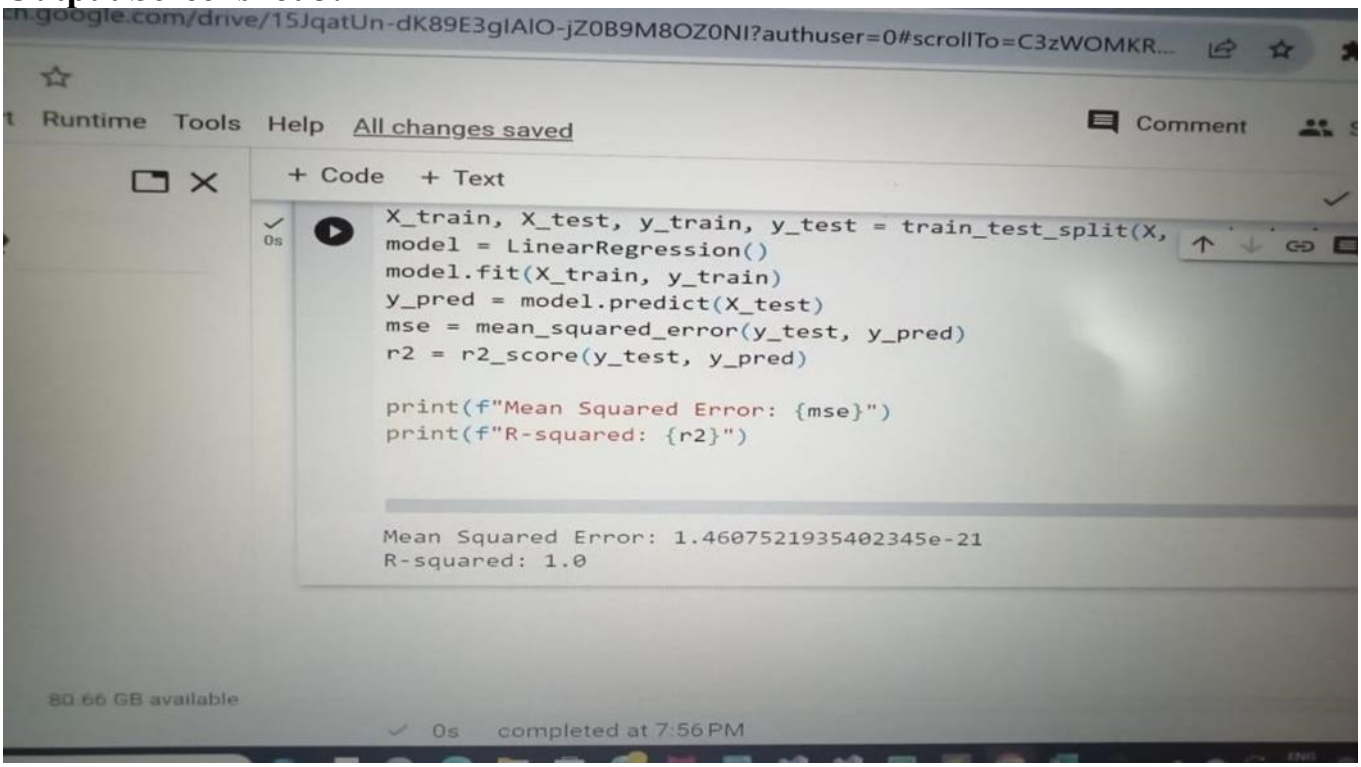



## CODE SAMPLE 3:

```
# Import necessary libraries
import  pandas as pd
from sklearn.model_selection import train_test_split from
sklearn.linear_model import LinearRegression from
sklearn.metrics import mean_squared_error, r2_score
# Load the previously trained model trained_model
= LinearRegression()
# Load the dataset for testing (assuming you have a CSV file named
'test_data.csv') test_data = pd.read_csv('test_data.csv') # Data Preprocessing
for Testing features = ['CRIM', 'ZN', 'INDUS', 'CHAS', 'NOX', 'RM', 'AGE',
'DIS', 'RAD',
'TAX', 'PTRATIO', 'B', 'LSTAT'] target
= 'target'
# Split the test data into features and target
variable X_test = test_data[features] y_test =
test_data[target]
# Make predictions using the trained model predictions
= trained_model.predict(X_test)
# Model Evaluation for Testing mse =
mean_squared_error(y_test, predictions) r2 =
r2_score(y_test, predictions)
print(f'Mean Squared Error (Testing): {mse:.2f}') print(f'R-squared
Value (Testing): {r2:.2f}')
```

**Output Screenshot 3:**




CONCLUSION :

1. **Measure Energy Consumption:** We successfully implemented energy measurement systems to accurately monitor and record energy usage across various settings, including residential, commercial, and industrial.
2. **Identify Patterns and Anomalies:** Through data analysis, we aimed to discern energy consumption patterns and anomalies, which could serve as a basis for informed decision-making and energy efficiency improvements.
3. **Promote Sustainability:** Our project aimed to raise awareness about the environmental impact of energy consumption and encourage individuals, organizations, and communities to adopt sustainable energy practices.

FUTURE ENHANCEMENT :

1. **Data Security and Privacy Measures:**
    - Enhance data security and privacy to ensure that sensitive energy consumption data remains confidential and protected from potential breaches.
1. **Real-Time Monitoring:**
    - Transition from periodic data collection to real-time monitoring by using smart meters and IoT (Internet of Things) devices. This would provide more immediate insights into energy usage and enable proactive energy management.