AI MEASURE ENERGY CONSUMPTION

NAME : KARUPPAYI.C
REGISTER NUMBER : 720421104026

INTRODUCTION:

The concept of energy consumption is directly related to energy efficiency since higher consumption results in lower energy efficiency.

It's estimated that during an hour about 1,000 watts are consumed, so this measure is used to calculate the consumption of homes, businesses, or any other type of building in order to issue the corresponding bills.

There are various factors that directly influence energy consumption such as:

- The activity that takes place in the home or business.
- The number of people in a household or workers.
- The consumption habits of each person.
- The energy performance of household appliances.

With the right information and technology, it's possible to use energy more responsibly and efficiently. This results in a reduction in energy consumption and, therefore, in significant savings on utility bills.
Energy consumption is the total amount of energy required for a given process and is measured in kilowatt hours (kWh). This includes the use of electricity, gas, diesel, oil, and biomass.

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

DATASET

Decision making process on knowledge of organisation policies in Data Analytics	120	1.991667	0.47892	0.69204	The result concludes that the decision making is showing the greatest impact in the Data Analytics to make quick decisions using Business Intelligence and Decision Making Systems.
Data Analytics is the concrete base for organisational success	120	2.1	0.79282	0.62857	Organisational success is also contributing to the Data Analytics which helps in the organizational growth.
One of the fascinating feature you consider in Data Analytics is Predictive Analytics	120	1.9	0.68087	0.46358	Predictive analytics helps in the organization to forecast the future through data.
Employee Performance and Compliance are the Prominent features in the Data Analytics	120	1.8	0.68087	0.46358	Performance of employees and Compliance helps in the effective implemention of Data Analytics.
Exclusive Induction Programs are one of the Significant feature in the Data Analytics	120	1.9166	0.68087	0.46358	Induction programs about Data Analytics given to the employees helps in the ease of use .

Here's a list of tools and software commonly used in the process:

1. Programming Language:

- Python is the most popular language for machine learning due to its extensive libraries and frameworks. You can use libraries like NumPy,pandas, scikit-learn, and more.

2. Integrated Development Environment (IDE):

- Choose an IDE for coding and running machine learning experiments. Some popular options include Jupyter Notebook, Google Colab, or traditional IDEs like PyCharm.

3. Machine Learning Libraries:

- You'll need various machine learning libraries, including:

- scikit-learn for building and evaluating machine learning models.

- TensorFlow or PyTorch for deep learning, if needed.

- XGBoost, LightGBM, or CatBoost for gradient boosting models.

4. Data Visualization Tools:

- Tools like Matplotlib, Seaborn, or Plotly are essential for data exploration and visualization.

5. Data Preprocessing Tools:
- Libraries like pandas help with data cleaning, manipulation, and preprocessing.

6. Data Collection and Storage:
- Depending on your data source, you might need web scraping tools (e.g., BeautifulSoup or Scrapy) or databases (e.g., SQLite,PostgreSQL) for data storage.

7. Version Control:

- Version control systems like Git are valuable for trackingchanges in your code and collaborating with others.

8. Notebooks and Documentation:

- Tools for documenting your work, such as Jupyter Notebooks or Markdown for creating README files and documentation

PROGRAM:

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns
```

```python
# customize the style pd.options.display.float_format

= '{:.5f}'.formatpd.options.display.max_rows = 12

# load the data

filepath = '../input/hourly-energy-consumption/PJME_hourly.csv'df =

pd.read_csv(filepath)

print("Now, you're ready for step one")

# turn data to datetime

df = df.set_index('Datetime') df.index =

pd.to_datetime(df.index)

# create the plot

df.plot(style='.',

        figsize=(15, 5),

        title='PJM Energy (in MW) over time')

plt.show()
```

MODEL CHOSEN FOR THE PREDICTION

SVR ALGORITHM

```python
import numpy as nm
import matplotlib.pyplot as mtp
import pandas as pd
#importing datasets

   data_set= pd.read_csv('user_data.csv')7.
#Extracting Independent and dependent Variable
```

```
 x= data_set.iloc[:, [2,3]].values
 y= data_set.iloc[:, 4].values
 # Splitting the dataset into training and test set.
 from sklearn.model_selection import train_test_split
 x_train, x_test, y_train, y_test= train_test_split(x, y, test_size= 0.25, random_state=0)
#feature Scaling
from sklearn.preprocessing import StandardScaler
st_x= StandardScaler()
x_train= st_x.fit_transform(x_train)
x_test= st_x.transform(x_test)
```

OUTPUT:

Index	User ID	Gender	Age	EstimatedSalary	Purchased
0	15624510	Male	19	19000	0
1	15810944	Male	35	20000	0
2	15668575	Female	26	43000	0
3	15603246	Female	27	57000	0
4	15804002	Male	19	76000	0
5	15728773	Male	27	58000	0
6	15598044	Female	27	84000	0
7	15694829	Female	32	150000	1
8	15600575	Male	25	33000	0
9	15727311	Female	35	65000	0
10	15570769	Female	26	80000	0
11	15606274	Female	26	52000	0
12	15746139	Male	20	86000	0
13	15704987	Male	32	18000	0
14	15628972	Male	18	82000	0

	0	1
0	-0.804802	0.504964
1	-0.0125441	-0.567782
2	-0.309641	0.157046
3	-0.804802	0.273019
4	-0.309641	-0.567782
5	-1.1019	-1.43758
6	-0.70577	-1.58254
7	-0.210609	2.15757
8	-1.99319	-0.0459058
9	0.878746	-0.770734
10	-0.804802	-0.596776
11	-1.00287	-0.422817

	0
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	1
8	0
9	0
10	0
11	0

Output:

```
Out[8]:
SVC(C=1.0,
coef0=0.0,    cache_size=200,   class_weight=None,


   decision_function_shape='ovr',     degree=3,
gamma='auto_deprecated',
   kernel='linear',                 max_iter=-1,
probability=False, random_state=0,
   shrinking=True, tol=0.001, verbose=False)
```

ADAVANTAGES:

1. Reduce Living Expenses. Saving energy usuallyreduces living expenses.
2. Benefits The Environment And Protects Wildlife. Lessconsumption benefits the environment and protects wildlife.
3. Less Power Plants.
4. Promote Health.
5. Reduce Dependence.
6. Finite Resources.
7. Make A Positive Impact.

DISADAVANTAGES:

The environmental problems directly related toenergy production and consumption include
1.air pollution,
2.climate change,
3.water pollution,
 4.thermal pollution,
5.solid waste disposal.

CONCLUSION:

In conclusion, the development of an automated energy consumption monitoring system with data analysis and visualization is a complex and multifaceted task. It encompasses various stages, including data collection, storage, analysis, visualization, and user interface development. Ensuring data security, scalability, and reliable automation further add to the complexity. Regular maintenance and updates are necessary to keep the system accurate and secure. This endeavor demands a multidisciplinary team with expertise in hardware, software, data science, and user interface design.