Power Consumption Analysis For Households Using IBM Watson

Introduction

Overview

Electricity sector in India is the world's third largest producer and third largest consumer of electricity. The gross electricity consumption in 2018-19 was 1,181 kWh per capita. Energy use can be viewed as a function of total GDP, structure of the economy and technology. The household electricity consumption ranks in the second position after the industrial electricity consumption, across the globe.

The increase in household energy consumption is more significant than that in the industrial sector. To achieve reduction in electricity consumption, it is vital to have current information about household electricity use. This will help you track the power consumed on regular intervals for all kinds of appliances which use heavy loads such as Air Conditioners, Oven or a washing machine etc.

Purpose

This project mainly focuses on applying a machine-learning algorithm to calculate the power consumed by all appliances in a household.

Literature Survey

Existing Problem

- Existing system is highly manual and involves a lot of paper work and calculation and therefore may be erroneous. This lead to inconsistency and inaccuracy.
- It consumes a lot of time, thus causes inconvenience.
- The data may be lost or destroyed.

Proposed Solution

To achieve reduction in electricity consumption, it is vital to have current information about household electricity use. This analysis mainly focuses on applying a

machinelearning algorithm to calculate the power consumed by all appliances. This will help you track the power consumed on regular intervals for all kinds of appliances which use heavy loads such as Air Conditioners, Oven or a washing machine.

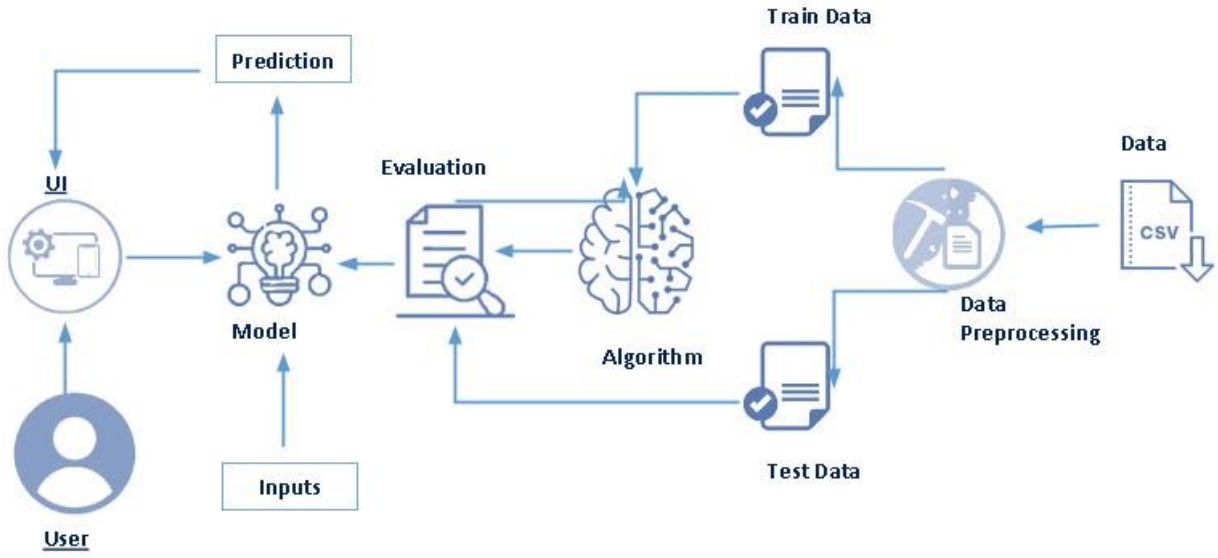
To eliminate all these difficulties and problems we proposed a system that predict the quality using Machine Learning services that we train the data from the past data that helps to predict the result very fast and accurate by using any regression algorithm

The algorithms can be chosen according to the objective. As per the dataset which we are using, it seems to be a regression problem, therefore we can use the following :

- Linear Regression
- Random Forest Regressor
- Decision Tree Regressor
- XGBoost Regressor and many more.

Theoretical Analysis

Block Diagram



Hardware requirements

processor : intel core i3
Ram : 4GB
Hard disk :30GB
input device : standard keyboard and mouse
output device : monitor
operating system: Windows
programming :python 3.6+

software requirements

Anaconda Navigator
Jupyter Notebook
Spyder

• IBM Watson Studio

Watson Studio provides you with the environment and tools to solve your business problems by collaboratively working with data. It provides a suite of tools for data scientists, application developers and subject matter experts, allowing them to collaboratively connect to data, wrangle that data and use it to build, train and deploy models at scale. Successful AI projects require a combination of algorithms + data + team, and a very powerful compute infrastructure.

• IBM Watson Machine Learning

IBM Watson Machine Learning is a full-service IBM Cloud offering that makes it easy for developers and data scientists to work together to integrate predictive capabilities with their applications. The Machine Learning service is a set of REST APIs that you can call from any programming language to develop applications that make smarter decisions, solve tough problems, and improve user outcomes.

- **IBM Cloud Object Storage**

IBM Cloud Object Storage is an application-data archive and backup platform that offers persistent cloud storage and data encryption by default. Made for electronic records retention, it helps with regulatory compliance.

EXPERIMENTAL INVESTIGATIONS

The household power consumption dataset is a multivariate time series dataset that describes the electricity consumption for a single household over four years.

The dataset was collected between December 2006 and November 2010 and observations of power consumptions within the household were collected every minute.

It is a multivariate series comprised of seven variables (besides the data and time), they are:

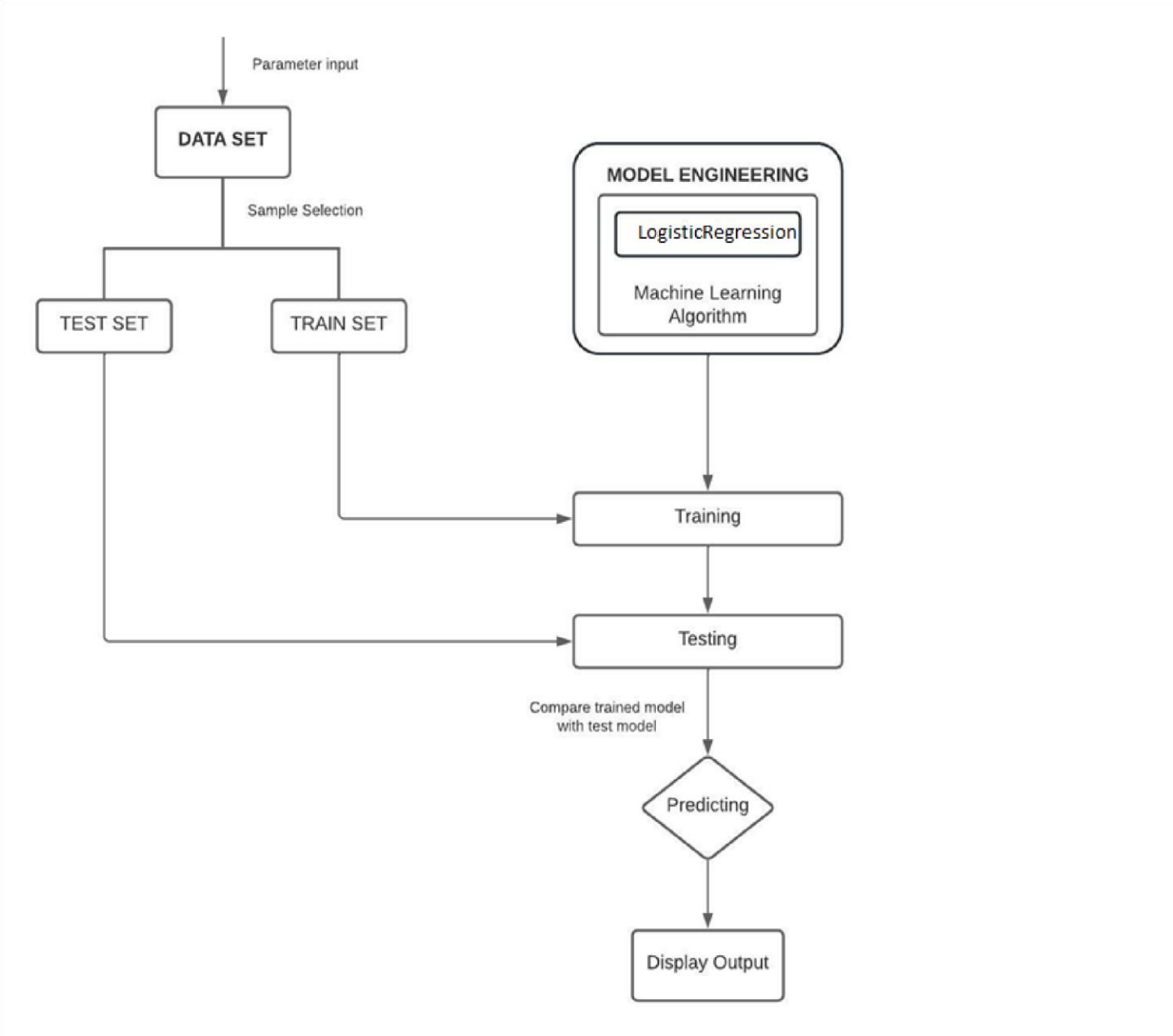
- **global active power**: The total active power consumed by the household (kilowatts).
- **global reactive power**: The total reactive power consumed by the household (kilowatts).
- **Voltage:** Average voltage(volts).
- **global intensity:** Average current intensity (amps).
- **sub_metering_1:** Active energy for kitchen(watt-hours of active energy).
- **sub_metering_2:** Active energy for laundry(watt-hours of active energy).
- **sub_metering_3:** Active energy for climate control systems(watt-hours of active energy).

Active and reactive energy refer to the technical details of alternative current .In general terms, the reactive energy is the real power consumed by the household, whereas the reactive energy is the unused power in lines.

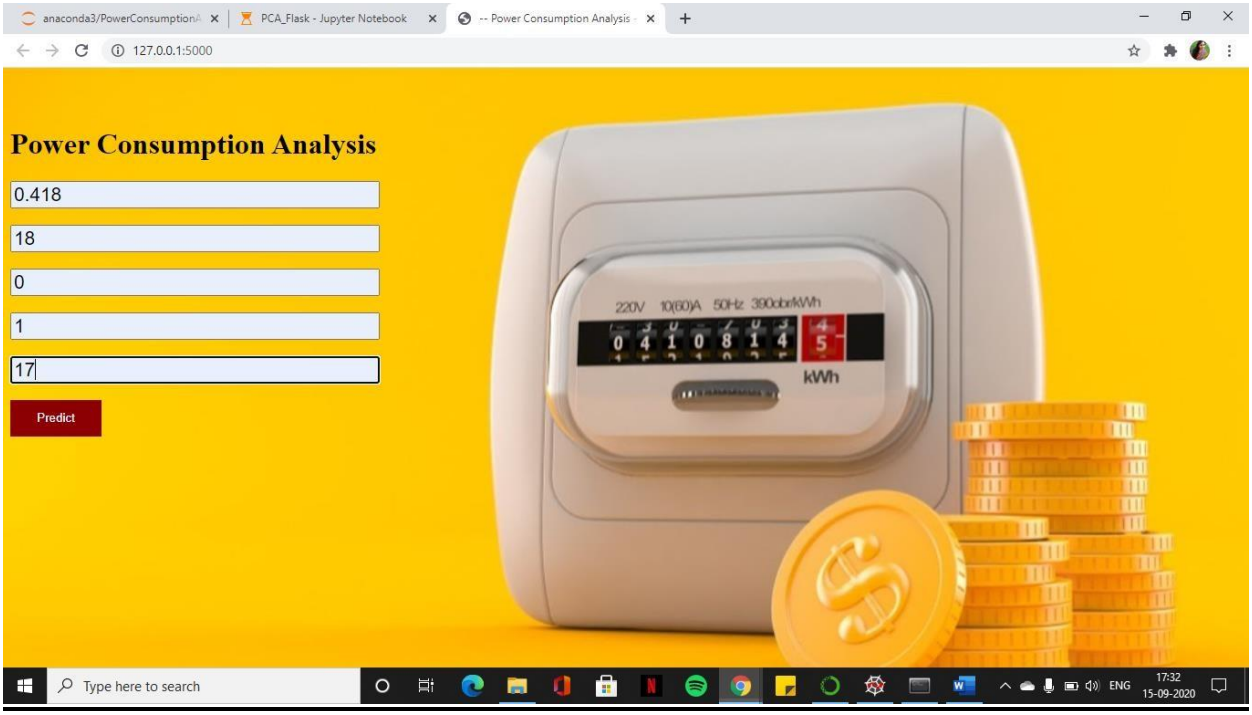
We can see that the dataset provides the active power as well as some division of the active power by main circuit in the house, specifically the kitchen, laundry and climate control. These are not all the circuits in the household.

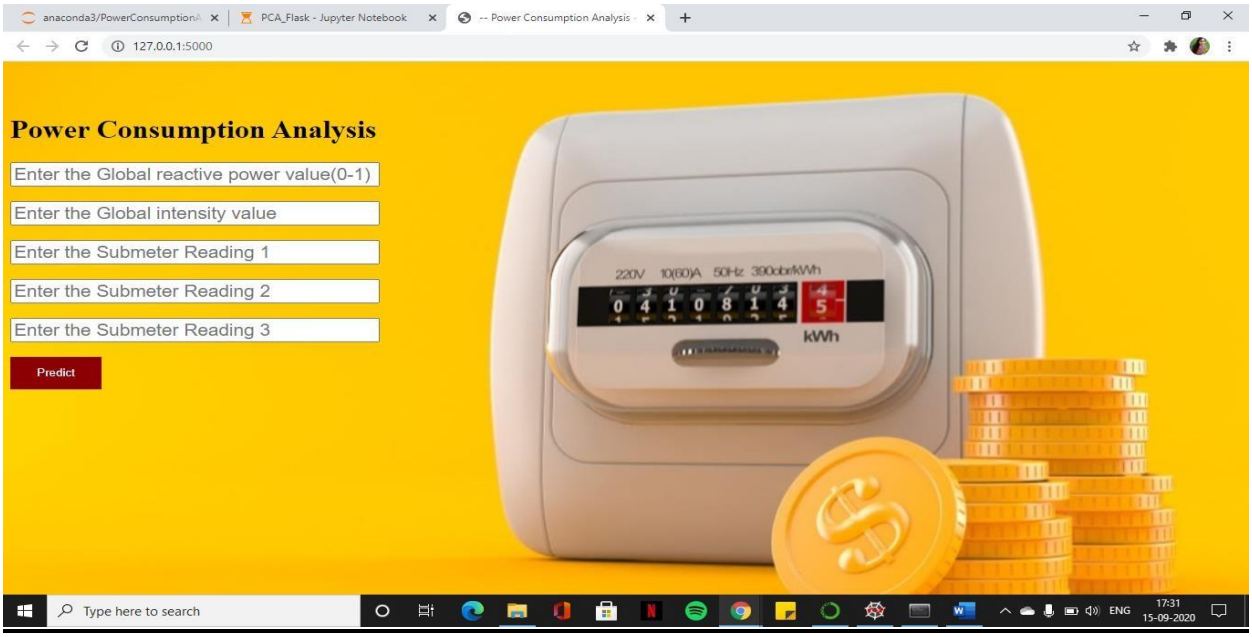
The remaining watt-hours can be calculated from the active energy by first converting the active energy to watt-hours then subtracting the other sub-metered active energy in watt-hour.

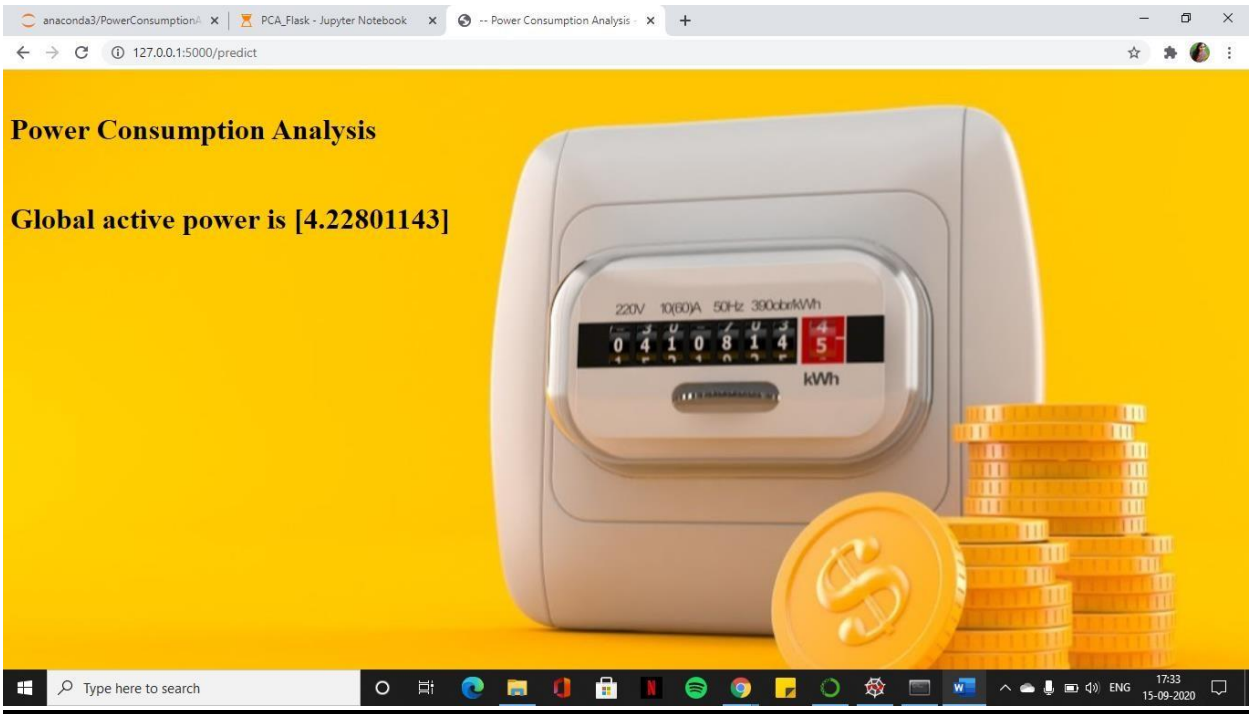
Flowchart


Parameter input
DATA SET
Sample Selection
TEST SET
TRAIN SET
MODEL ENGINEERING
LogisticRegression
Machine Learning Algorithm
Training
Testing
Compare trained model with test model
Predicting
Display Output

Result


anaconda3/PowerConsumptionA
PCA_Flask - Jupyter Notebook
-- Power Consumption Analysis
127.0.0.1:5000
Power Consumption Analysis
0.418
18
0
1
17
Predict
220V 10(60)A 50Hz 3200imp/kWh
kWh
Type here to search
ENG
17:32
15-09-2020


anaconda3/PowerConsumptionA
PCA_Flask - Jupyter Notebook
-- Power Consumption Analysis
127.0.0.1:5000
Power Consumption Analysis
Enter the Global reactive power value(0-1)
Enter the Global intensity value
Enter the Submeter Reading 1
Enter the Submeter Reading 2
Enter the Submeter Reading 3
Predict
220V 10(60)A 50Hz 3200imp/kWh
kWh
Type here to search
ENG
17:31
15-09-2020



ADVANTAGES & DISADVANTAGES

Advantages

Faster Claim Settlements & Save time :

We can predict the output very fast compared to other system and it save time due to machine learning so we can predict the power consumption using machine learning techniques

Cost Efficiency :

Due to automation of everything therefore the cost is reduced compared to the manual process

- Automation of Everything

Machine Learning is responsible for cutting the workload and time. By automating things we let the algorithm do the hard work for us. Automation is now being done almost everywhere. The reason is that it is very reliable. Also, it helps us to think more creatively. Due to ML, we are now designing more advanced computers. These computers can handle various

Machine Learning models and algorithms efficiently. Even though automation is spreading fast, we still don't completely rely on it. ML is slowly transforming the industry with its automation.

- Wide Range of Applications

Machine Learning has a wide variety of applications. This means that we can apply ML on any of the major fields. ML has its role everywhere from medical, business, banking to science and tech. This helps to create more opportunities.

- Scope of Improvement

Machine Learning is the type of technology that keeps on evolving. There is a lot of scope in ML to become the top technology in the future. The reason is, it has a lot of research areas in it. This helps us to improve both hardware and software.

In hardware, we have various laptops and GPUs. These have various ML and Deep Learning networks in them. These help in the faster processing power of the system. When it comes to software we have various UIs and libraries in use. These help in designing more efficient algorithms.

- Efficient Handling of Data

Machine Learning has many factors that make it reliable. One of them is data handling. ML plays the biggest role when it comes to data at this time. It can handle any type of data. Machine Learning can be multidimensional or different types of data. It can process and analyze these data that normal systems can't. Data is the most important part of any Machine Learning model. Also, studying and handling of data is a field in itself.

Disadvantages

- Possibility of High Error

In Machine Learning, we can choose the algorithms based on accurate results. For that, we have to run the results on every algorithm. The main problem occurs in the training and testing of data. The data is huge, so sometimes removing errors becomes nearly impossible. These errors can cause a headache to users. Since the data is huge, the errors take a lot of time to resolve.

- Algorithm Selection

The selection of an algorithm in Machine Learning is still a manual job. We have to run and test our data in all the algorithms. After that only we can decide what algorithm we want. We choose them on the basis of result accuracy. The process is very much time-consuming.

- Data Acquisition

In Machine Learning, we constantly work on data. We take a huge amount of data for training and testing. This process can sometimes cause data inconsistency. The reason is some data constantly keep on updating. So, we have to wait for the new data to arrive. If not, the old and new data might give different results. That is not a good sign for an algorithm.

- Time and Space

Many Machine Learning algorithms might take more time than you think. Even if it's the best algorithm it might sometimes surprise you. If your data is large and advanced, the system will take time. This may sometimes cause the consumption of more CPU power. Even with GPUs alongside, it sometimes becomes hectic. Also, the data might use more than the allotted space.

Applications

1. To assess the energy efficiency and energy usage of their homes.
2. To use the new understanding of the problem to consider different framings of the prediction problem, ways the data may be prepared and modeling methods that may be used.
3. The household power consumption dataset is a multivariate time series dataset that describes the electricity consumption for a single household over years.

Conclusion

Finally, total power consumption by all the appliances is calculated and displayed.

BIBLOGRAPHY:

https://machinelearningmastery.com/how-to-load-and-explore-household-electricity-usage-data/