4222-SURYA GROUP OF INSTITUTION
VIKRAVANDI-605 652

NAAN MUDHALVAN PROJECT

MEASURE ENERGY CONSUMPTION

PREPARED BY:

K.ARUN

REG NO :422221106003

ECE DEPARTMENT

3RD YEAR 5TH SEM

OBJECTIVE:

The purpose of this research was to predict energy consumption using the data of Finland's transmission system operator. The objective of this project was to test if a machine learning model can yield good enough results in a complex forecasting problem, exploring machine learning techniques and developing a data-driven model for forecasting energy.

DATA SOURCE:

To predict the total energy consumption of a data center, we calculated the PUE according to different internal and external parameters, and estimated the IT equipment energy consumption. The overall energy consumption, carbon emissions, and electricity cost of the data center were obtained directly.

DATA PREPROCESSING:

Data preprocessing is an important step before applying machine learning methods for energy or load prediction. The common steps include data imputation, data resolution processing, data normalization, outlier detection and data smoothing. Data smoothing could reduce the influence of noisy data and improve data quality, which contributes to the improvement of model performance. There are several approaches of data smoothing widely used in the field of energy

consumption prediction, such as moving average approach seasonal exponential smoothing model Gaussian kernel density estimation exponential moving average filter neighbor-averaging interpolation and Savitzky-Golay filter. Most studies either have no data smoothing or only choose one method for data smoothing according to their experience.

FEATURE EXTRACTION:

Features should be selected and determined before model construction and they are important to the model performance. Engineering knowledge or statistical methods are two main methods to determine model inputs . According to engineering knowledge, HVAC energy consumption relates to outdoor weather and occupancy so the dry-bulb temperature, dew point temperature and occupant schedules are widely used as model inputs. With the availability of system operation data from building management systems, many studies consider historical energy consumption data as model inputs because they can reflect the delay of weather influence and similar operation patterns on the dynamic HVAC energy consumption. When the historical data are used as inputs, the length of historical data can affect the model performance. Existing studies adopt historical data to predict future energy consumption according to their experience and these models can achieve accurate prediction results.

#features about the hour of the day, day of the week type things

def create_features(df):

```
    df=df.copy()     df['hour'] = df.index.hour
df['dayofweek'] = df.index.dayofweek     df['quarter']
= df.index.quarter     df['month'] = df.index.month
df['year'] = df.index.year     df['dayofyear'] =
df.index.dayofyear     df['dayofmonth'] = df.index.day
df['weekoftheyear'] = df.index.isocalendar().week
return df
```

MODEL DEVELOPMENT:

The model performance of load prediction models based on de learning methods is closely related to training sets used to map t relationship betwee inputs and outputs. Large training set volum contribute to more accurat results but the computation time is longe Sometimes it is hard to colle sufficient data in practical projects. Small training set volumes use fewe calculation resources but the model m produce worse prediction resul Therefore, a suitable training set volum should be selected to ensure mode performance with less computing co The training sets are usually selected i a certain proportion from the who dataset in most research, ignoring th balance of accuracy and computation time in practice . As a result, guidanc lacks in the applications of predicti models. Therefore, it is necessary t address how to select suitable trainin set volumes in practical applications t ensure reasonable computing tim and accuracy.

VISUALIZATION:

With data visualization, consumers can acquire a bett understanding of their appliances' consumption, showing how th contribute to the whole ene consumption. Thus, with the right incentiv consumers can become more aw and more energy-efficient. For instanc some studies have already shown in in energy consumptio visualization inside buildings, where the former is interested in lig commercial buildings' energy consumption visualization, a the latter focused on domestic consumption visualization and human behav simulation. In addition, augmented reality is also utilized to enhance t consumers' awareness regarding electronic devices/appliances consumptio

```python
import pandas as pd  import numpy as np

import matplotlib.pyplot as plt  import seaborn

as sns from sklearn.metrics import

mean_squared_error from

sklearn.model_selection import train_test_split

import pickle


import xgboost as xgb  from xgboost

import plot_importance
```
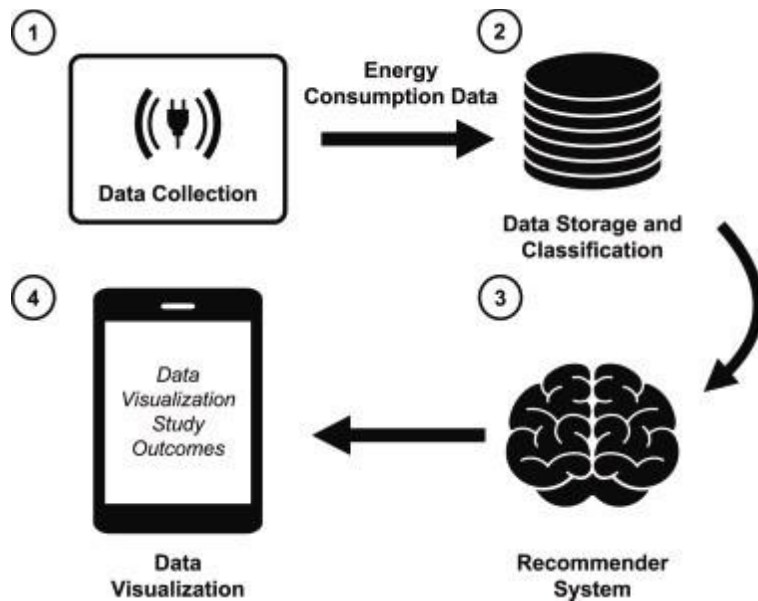
1. To identify consumption trends which will enable them to reduce the consumption. For example, the trends might show that the consumption is at its highest in the evening between when the user's family is back from work/school.

2. To determine the energy consumed by individual appliances so that the users can identify which appliances are responsible for high energy consumption

AUTOMATION:

The dataset is obtained by combining the two data sources. It is classified by probe measurement channel, where each channel corresponds to a server connection. The analysis focuses on power consumption and the types of tests performed. The seven tests carried out are described in Section This categorization enables a detailed impact analysis of the system

charging tests according to energy consumption and system process usage.

The no charge test was developed as a baseline for future analysis and comparison. Using this approach, it is possible to isolate the system processes from the benchmark techniques applied to stimulate the machines. Therefore, the no charge test served as a benchmark for comparison of the energy consumption of the other tests.

Graphical analysis helps to understand the relationship between the different identified characteristics. The conclusions drawn thanks to the evaluation enable the creation of a correct and reliable dataset. Data are analyzed with a special focus on finding possible information patterns of interest. Additionally, the evolution of measurements throughout the benchmarks is considered, taking into account the CPU progression for each test.