

Django/Flask Programming Question:

1. Write a program in Django/Flask which shows real-time data on a browser, i.e. whenever there is any change/addition reflected on the database, it should reflect on the browser. This program should not use any update/refresh call from browser side. You can use packages like Celery or Websockets to implement it. You can assume the data in the database is a temperature data of a sensor which receives value every 10 seconds from a source. The data could be displayed over a table or graph (whichever is easier for you)

Data Science Task Overview

Non intrusive load monitoring or NILM is a process of analysis of disaggregation of electric appliances connected to a power within a certain area.

This understanding would allow occupants to make intelligent, informed decisions on how they conserve the use of energy.

Current task challenges prospective employee to build a machine learning based algorithm that automatically and accurately identifies what appliances are connected to the power supply.

Data Description

Data is collected using current(mA) sensor attached to the wire of main power supply. In current simulation environment total 3 devices are connected: 2 normal PC display and 1 temperature sensor. PC display can be in 3 states: OFF, IDLE and ON. Temperature sensor can be only in 2 states: ON or OFF. **test_data1536128609.98_screen1.csv** contains the data when only 1st PC display is connected to the power supply and other 2 devices are disconnected. **test_data1536128871.84_screen2.csv** contains the data when only 2nd PC display is connected to the power supply and other 2 devices are disconnected. **test_data1536129261.77_device253.csv** contains the data when only temperature sensor is connected to the power supply and other 2 devices are disconnected. **test_data1536129670.55_main.csv** - contains the data when all 3 devices are connected to the power supply. All csv files follows the format defined below:

- Timestamp
- Maximum current(mA) value detected within one second
- Effective current(mA) value detected within one second

Data:

- **test_data1536128609.98_screen1.csv** - current data from display 1
- **test_data1536128871.84_screen2.csv** - current data from display 2
- **test_data1536129261.77_device253.csv** - current data from temperature sensor

- **test_data1536129670.55_main.csv** - current data consumed by the devices which are connected to the central power.

Problem Statement

Provide machine learning based program which receives as an input file **test_data1536129670.55_main.csv** and as an output returns csv file with labeled state for each device.

For example if first row of **test_data1536129670.55_main.csv** is

| Timestamp | Maximum current | Effective current |
|------------------|-----------------|-------------------|
| 1536129671.00052 | 24.41 | 17.3 |

Output csv file's first row should be:

| timestamp | Maximum current | Effective current | PC screen1 state | PC screen2 state | Temperature sensor state |
|------------------|-----------------|-------------------|------------------|------------------|--------------------------|
| 1536129671.00052 | 24.41 | 17.3 | OFF | IDLE | OFF |

Also please submit all the auxiliary files(source codes, EDA, testing charts, etc.) that you have developed during the process of building the machine learning model. It will be great to understand your approach in solving the problem.