**Read Me**

Andrew Bailey (April, 2024)

For all Python notebooks, comments are in the code with explanations. Below is a short description of the purpose of each file, with those bolded being the most important.

* **Sched.py**
  + This is the main file that was used for the final project demonstration. It runs on a schedule that performs all of the necessary tasks such as the following
    - Get influx connection details from the user
    - Read the past 15 minutes of total power and timestamp from influx
    - Initialize prediction data frame
    - Load the trained ML model
    - Make predictions
    - Compute the cost calculation
    - Send all this data back to the influx database
    - Wait 20 minutes then run again
* **sched.exe** 
  + This is the application that runs the script talked about above. Pyinstaller was used to make the script. If you want to change the sched.py file or make a new python script to run as an executable look up pyinstaller or py2exe to turn it into an application
* **CNN LSTM Categorized**
  + This is the main file used for training the machine learning model. It uses a convolutional neural network and long short-term memory to make predictions. Further comments are located in the code.
* **Write Data to Influx (Read Instructions)**
  + This is a template that uses the influx library to send data to a database. The best way to use this is to make a copy of the notebook and fill in the information specific to your database
* **Combined\_training\_data.csv**
  + This is a good example of the training data used for training the machine learning model. More training data files can be found in the datasets folder or other/JE Data in the 2022-2023 project
* Pecan Street CNN LSTM
  + This is for creating a model using the open-source Pecan Street training data
* Simplifying Consolidated Training Data
  + This simplifies the appliances for training. For instance, all the different parts of the HVAC system are added into one row called “HVAC”
* Import and Test Model
  + This is used for importing a trained machine-learning model and making predictions with it
* Consolidate Training Data
  + This script is used for training data consolidation. It can take many spreadsheets and form them into usable training data
* Read from InfluxDB
  + This is a notebook only containing code that uses the influx library to read data from an influx bucket
* Pecan Street File Extraction
  + The pecan street data is downloaded from their website as a tar.gz file. This is used for unzipping it and converting to CSVs
* Archive - Folder
  + This contains old files which are not relevant anymore
* Datasets
  + This contains many CSV files that are datasets. The blind tests just contain the timestamp and total power measurement. The open/training data contain the timestamp, total, and individual power measurements