ME 592: Data Analytics and Machine Learning for Cyber-Physical Systems

Homework 1

Homework Assigned on February 1, 2022 Homework Due on: February 10, 2022

Note that individual submissions are required for this homework, we will switch to group-based homeworks from the next one.

1 Git Repository & Code Management

All the homeworks would have to be submitted using git commits containing all the codes. So, you would have to commit `*.ipynb` file from your google colab to the git repository. Although these git repositories could be made in either BitBucket or GitHub, We suggest creating a repository in GitHub and sending us the link to the git repository. If you wish to form private repositories, then you could give access of the GitHub handle to *"xylee95"* for evaluation of the homeworks.

2 Simple Programming & Exploratory Analytics

The following problems must be coded using specified tool only in order to get everyone to use python.

2.1 Images

Task:

1. Load sudoku-original.png image in grayscale mode.
2. Plot the histogram of intensities.
3. Choose 4 appropriate points and perform a Perspective Transformation to obtain the sudoku as a proper square.

2.2 Time Series

Data: Experimental data used to analyze appliances energy use in an energy efficient building. The data set is collected at a frequency of 10 min for about 4.5 months. The house temperature and humidity conditions were monitored with a ZigBee wireless sensor network. Each wireless node transmitted the temperature and humidity conditions at a period of around 3.3 min. Then, the wireless data was averaged for 10 minute periods. The energy data was logged every 10 minutes with m-bus energy meters. Weather from the nearest airport weather station (Chievres Airport, Belgium) was downloaded from a public data set from Reliable Prognosis (rp5.ru), and merged together with the experimental data sets using the date and time column. Random variable is included in the data set for testing the regression models and to filter out non predictive attributes (parameters). This data is adopted from the UCI machine learning repositories [1] and several aspects of the data were analyzed [2]. Our motivation is to explore some aspects of this time series data.
Task: Load the data (energydata_complete.csv) and perform the following analysis.

1. Plot the appliances energy consumption for whole period and a closer look at any one week of consumption.
2. Plot heatmap of hourly consumption of appliances for a week. An example heatmap looks like Figure 1.
3. Plot the histogram of energy consumption of appliances.
4. Construct a feature variable NSM (no. of seconds from midnight) and plot energy consumption vs. NSM.
5. Plot appliances energy consumption vs. Press_mm_Hg.
6. It is observed that the major contributing factors for the energy consumption among all other features is NSM and Press_mm_Hg. Comment on it.

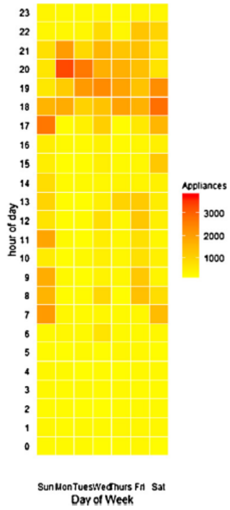

Figure 1: Example heatmap of hourly energy consumption of appliances over a week

2.3 Multi-variate

Data: The NASA data set comprises different sizes of NACA 0012 airfoils at various wind tunnel speeds and angles of attack. The span of the airfoil and the observer position were the same in all of the experiments. This problem has the following inputs:

1. Frequency, in Hz.
2. Angle of attack, in degrees.
3. Chord length, in meters.
4. Free-stream velocity, in meters per second.
5. Suction side displacement thickness, in meters.

The only output is Scaled sound pressure level, in decibels.

Task: Load the data and Compute the following descriptive statistics of the data:

1. Mean
2. Variance (or Standard Deviation)

3. Median
4. Kurtosis
5. Skewness
6. Range

Attachments

1. sudoku-orginal.png
2. energydata-complete.csv
3. airfoil_self_noise.dat

Expected Outcome

Final (committed) code with results for Section 2 should be pushed before the deadline.

References

1. https://github.com/LuisM78/Appliances-energy-prediction-data

2. Data driven prediction models of energy use of appliances in a low-energy house. Luis M. Candanedo, Véronique Feldheim, Dominique Deramaix. Energy and Buildings, Volume 140, 1 April 2017, Pages 81-97, ISSN 0378-7788

3. https://archive.ics.uci.edu/ml/datasets/Airfoil+Self-Noise

4. T.F. Brooks, D.S. Pope, and A.M. Marcolini. Airfoil self-noise and prediction. Technical report, NASA RP-1218, July 1989.