YCNG 233, Assignment 5 (Final Presentation)

Presentation Due March 21st

Instruction

Machine learning-based methods have shown be effective for short-term electric load forecasting [1, 2, 3]. However, in real world, we may not have enough data to harness the full power of machine learning models. In this homework, we aim to investigate short-term load forecasting with low data regime [4, 5]. Specifically, we aim to forecast the one-hour ahead load forecasting with the last four observations as shown in Figure 1.

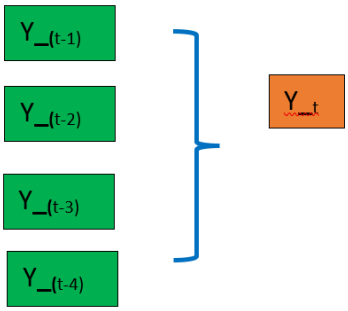


Figure 1. One-hour ahead load forecasting

Data set. You are given hourly data sets for four residential houses (House A, B, C, D) in New York City. For house A, as shown in Figure 2, we have the data for the last 2 days in November (48 data points) as the training set and the data for December as the testing set (744 data points). For house B, C, D, we have the data sets for the first 11 months (January-November) for these three houses (8016 data points), as shown in Figure 3.

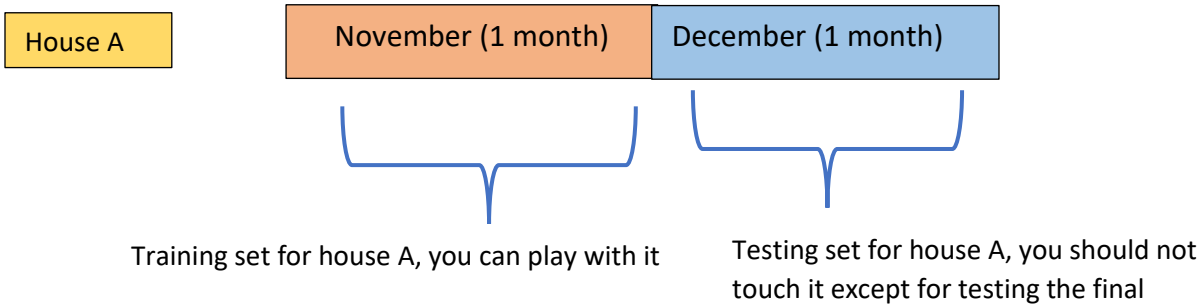


Figure 2. Overview for the house 1 hourly data set

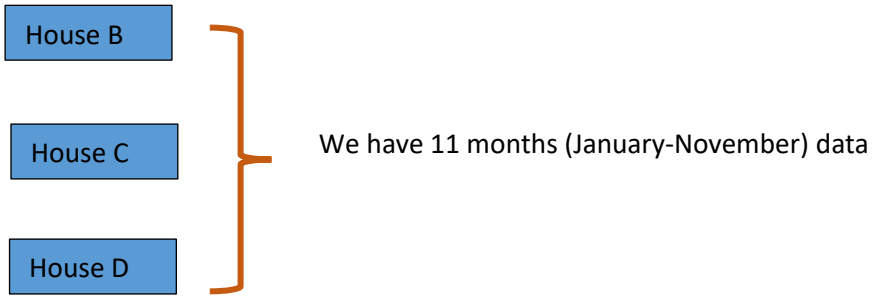
There are **two subtasks** for this assignment.

Task 1: Only use the two days data of house A as the training data and report the performance on the testing set of house A. Please note that you should report the performance of different models including deep feedforward neural networks, LSTM-based model, linear regression, repeat the last one, and at least

one statistical method. Besides the aforementioned methods, you can also try to test more methods for this problem.

Task 2: Now assuming that besides house A, you also have data sets for house B, C, D, and you are free to use the data sets of house B, C, D as well the training set of house A. **Can you try to come out a new solution and report its performance on the testing set on house A.** Please note that the testing set of house A should never be touched in the training phase.

Metrics we should use: **MAPE** (Mean Absolute Percentage Error), **MAE** (Mean Absolute Error).



Requirements for Presentation

You are required to prepare a 12-min (followed with 3 mins QA) presentation on March 21st.

In the presentation, you are required to report following items.

- One page to introduce the problem setup and data set.
- Any data preprocessing.
- Detailed results for task 1. You should report the MAPE, MAE for different models
- Detailed results for task 2. You should report the MAPE, MAE for your new solution.
- Other findings if you have.

Reference

[1] Machado, Eduardo, et al. "Electrical Load Demand Forecasting Using Feed-Forward Neural Networks." *Energies* 14.22 (2021): 7644.

[2] Zheng, Jian, et al. "Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network." *2017 51st Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2017.

[3] https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/

[4] Wu, Di, et al. "Multiple kernel learning-based transfer regression for electric load forecasting." *IEEE Transactions on Smart Grid* 11.2 (2019): 1183-1192.

[5] Weixuan Lin, Di Wu: Residential Electric Load Forecasting via Attentive Transfer of Graph Neural Networks. IJCAI 2021: 2716-2722