February 6, 2021

A Comparative Study of Machine Learning and Deep Learning Techniques for Short Term Load Forecasting

Electrical Engineering and Computer Science | MA19019 | Abraham KALIGAMBE

Power System Engineering | Supervisor | Prof. Dr. Goro FUJITA

1. Introduction

World over, commercial buildings consume a large proportion of the energy generated from fossil fuels hence are responsible for nearly 40% of both direct and indirect global CO_2 emissions. As a means to curb CO_2 emissions, lower energy costs, and minimize our dependence on fossil fuels, low carbon related technologies and methods have been developed to address these challenges. These include;

- Energy wastage reduction using optimization techniques
- Integration of renewable energy sources such as wind and solar PV
- Usage of highly energy efficient electrical appliances
- Innovative construction techniques

Most commercial buildings employ an Energy Management System (EMS) to monitor, control, and optimize both local generation and load consumption. Given the importance of an EMS in the viability of energy efficiency and optimization, plus the abundance of historical energy data from smart meters, the main purpose of this study was to develop a short term load forecasting (STLF) model for a Zero Energy Building (ZEB-Ready). The building used as a case study is Enefice Kyushu which is an office building for DAI-DAN CO., LTD. The major goal was to design an accurate 7 days ahead load forecasting model using state of the art Artificial Intelligence (AI) techniques i.e. machine and deep learning models. The accuracy of STLF models was based on RMSE and MAE metrics.

2. Load Forecasting Overview

In practice, load forecasting is performed for different purposes and time intervals. Researchers classified load forecasting into four categories concurring to the forecast duration. One of the categories is Short Term Load Forecasting whose duration ranges from 1 hour to one week. The Energy consumption of a commercial building is influenced by multiple factors. To model an accurate STLF model, all the factors shown in figure 1 were analyzed using data analysis techniques.

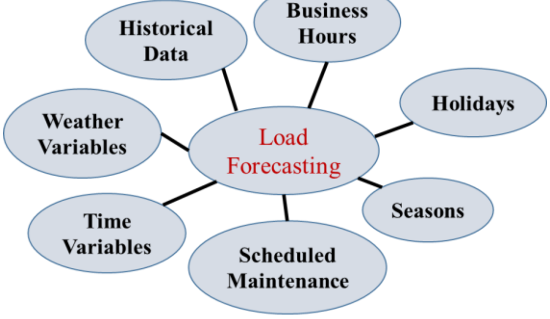

Figure 1. Factors affecting STLF in buildings

The importance of each factor is dependent on the magnitude of its influence on the energy consumption of a building. In addition to the factors illustrated in figure 1, time lags are other very important input features. Time lags are engineered from the historical energy consumption data. The selection of optimal time lags improves a model's performance[1]. Many studies on STLF were published using various approaches and methods with varying degrees of success. A survey on these approaches was carried out by Bouktif et al. in 2018. This survey, which covers the recent publications on the subject, classified STLF methods into four categories i.e. Statistical techniques, Artificial Intelligence (AI) techniques, Knowledge-Based Expert Systems, and Hybrid models. In our study, we only focused on AI techniques for STLF based on the recent publications and the success of AI in other fields.

3. Proposed AI methodology for Load Forecasting

In our study, a number of both popular Machine Learning (ML) and Deep Learning (DL) algorithms for regression were extensively reviewed[1]. From the ML side, we trained the dataset using linear regression, KNeighbors regressor, RandomForest regressor, GradientBoosting regressor, ExtraTrees regressor, and MLP regressor models using the Scikit Learn ML library in Python framework. We then compared the forecasting results with those of DL models. The DL models were DNN, LSTM, a hybrid of LSTM-CNN, and WaveNet-1D CNN model. The DL models were implemented using Keras and TensorFlow DL libraries.

3.1 Feature Engineering and analysis

The dataset for our case study was cleaned and organized. Input features were extracted from the data using various techniques such as a correlation matrix and visual analysis using box plots. Then the data were normalized using a min-max scaler and all categorical features were encoded using a one-hot encoder function. The final dataset contained 8229 data points with a mean of 9.5kWh. These were randomly split into training (92%), validation (6%), and testing (2%).

3.2 WaveNet-1D CNN model

WaveNet-1D CNN is a special architecture of Convolutional Neural Networks(CNN) that uses a stack of dilated causal convolutions[2]. A dilated convolution is more efficient than an ordinary convolution layer because a filter is applied over a longer input sequence than its length. This is done by skipping input values with a certain step known as a dilation rate. Lower layers learn short term patterns while higher layers learn long term patterns. Each added convolution layer doubles the receptive field [2] as shown in figure 2.

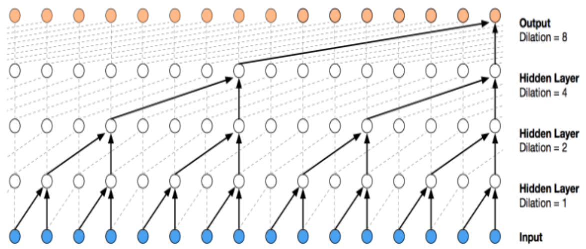

Figure 2. Illustration of WaveNet-1D CNN operation

The input feature for the WaveNet model is the optimal time-lagged energy consumption data. The filters in the input layer compute the weighted sums and biases and pass their output as an input to the dilated hidden layer and the process is repeated. Table I is a summary of the hyperparameters used to design the WaveNet model.

TABLE I. WaveNet-1D CNN MODEL DESIGN

Model Design	Quantity/Type
Number of filters	32
Kernel Size	2
Strides	1
Padding	Causal
Dilation_rate	{1,2,4,8,16,32}

The output from WaveNet was then fed into a fully connected (FC) layer. At this stage, all the other input features were also fed into the FC layer and finally to the output layer with only one neuron. The latter is the layer responsible for the forecasting. The training was done using a gradient descent algorithm. Adam optimizer was used to achieve this process. To increase non-linearity, we applied an activation function called ReLU on all layers. The model was compiled using the Mean Squared Error (MSE) as the loss function.

3.3 Evaluation metrics

The performance evaluation of our models was computed using RMSE and MAE[1] metrics shown in the equations (1) and (2) respectively;

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}} \tag{1}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{2}$$

Where $\hat{y}_i$ is the forecasted load value, y_i is the actual load value and n is the number of testing samples.
The comparison of the ML and DL models was done based on their RMSE and MAE scores.

4. Experimental Results

Several case studies were performed to select optimal input features, hyperparameters for model training, and determination of a baseline model for comparison of ML and DL models. To begin with, the dataset was trained and tested on our selected ML models. MLP Regressor achieved the best score of both RMSE and MAE and was hence selected as the baseline model. Then, artificial neural network models were trained and tuned on the same dataset. Table II is a summary of the hourly performance results after model tuning.

TABLE III. PERFORMANCE COMPARISONS

Algorithm	Test MAE	Test RMSE
MLPRegressor (Baseline)	2.1885	3.5161
DNN	2.6631	3.8791
LSTM	1.1120	2.5143
1D CNN & LSTM	0.8244	1.9821
WaveNet-1D CNN	**0.6377**	**1.3863**

Figure 3 depicts the comparison of actual load consumption and WaveNet-1D CNN model's hourly predictions for one week on the test dataset.

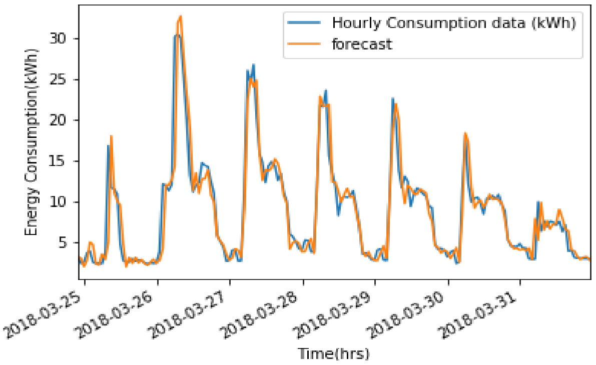

Figure 3. One week forecast Vs Actual Consumption

It shows a good fit and manifests the power of WaveNet architecture to extract useful features from a raw dataset. MAE of 0.6377 implies that the model's prediction is off by an average of 0.6377 kWh per hour.

5. Conclusion

Our study presented a comparison of various artificial intelligence approaches to short term load forecasting for commercial buildings. Notably, it proposed the use of a Deep Learning technique that utilized 1D Convolutional Neural Networks with a WaveNet architecture that was able to make a week ahead forecast and showed the lowest value of test MAE and RMSE compared to other AI models. The accuracy of STLF models is dependent on the dataset in question. It's, therefore, our recommendation to always train the dataset on several AI models and then select the most accurate model for the Energy Management System.

References

[1] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches," *Energies*, vol. 11, no. 7, 2018, doi: 10.3390/en11071636.

[2] A. van den Oord *et al.*, "WaveNet: A Generative Model for Raw Audio," pp. 1–15, 2016, [Online]. Available: http://arxiv.org/abs/1609.03499.

Research Achievements

1) A. Kaligambe and G. Fujita, "Short-Term Load Forecasting for Commercial Buildings Using 1D Convolutional Neural Networks," *2020 IEEE PES/IAS PowerAfrica, PowerAfrica 2020*, 2020, doi: 10.1109/PowerAfrica49420.2020.9219934.
2) A. Kaligambe and G. Fujita, "Short Term Load Forecasting for Zero Energy Buildings Using A Deep Learning LSTM Model for Multivariate Time Series", proceedings in *the14th South East Asian Technical University Consortium Symposium (SEATUC2020)*, KMTT, Thailand February 27-28, 2020
3) A. Kaligambe and G. Fujita, "A Comparative Study of Machine Learning and Deep Neural Network Models for Load Forecasting", *The15th South East Asian Technical University Consortium Symposium (SEATUC2021)*, non-reviewed (In progress), ITB, Indonesia February 25-26, 2021