A Deep Learning-Based Strategy to the Energy Management-Advice for Time-of-Use Rate of Household Electricity Consumption

Lu-Xian Wu, Shin-Jye Lee

Institute of Management of Technology, National Chiao Tung University, Taiwan
wainshain@gmail.com, camhero@gmail.com

Abstract

With the high industrialization at a rapid pace, the demand for energy increases exponentially, but it is difficult to meet the balance of demand and supply. Therefore, how to effectively meet the balance of demand and supply intelligently has become a popular issue in this century. In recent years, Taiwan Power Company (Taipower), the biggest electric company in Taiwan, is committed to the construction of Advanced Metering Infrastructure (AMI), which provides communication channels and enables demand-side users to participate in load dispatch. On the other hand, the construction of AMI is expected to generate a tremendous number of valuable data on electricity consumption, but it is not easy to convert these data into effective information by the conventional quantitative methods. In as much as the rapid progression of AI technology in the industrial field, the application of AI technology in the technology management has become an increasing issue as an interdisciplinary study. To address this task, this work applies the recurrent neural network based on deep learning to predict low-voltage usage shortly by the electricity information of low-voltage user and meteorological data. After many vicissitudes, the electricity consumption per hour can be predicted and a sound energy arrangement can be therefore planned. Through introducing the proposed model, Taipower Company will have an effective capability that schedules power, reduces unnecessary backup power, and provides time-consuming electricity prices for industrial enterprises accurately among high usage of Taiwan industries.

Keywords: Machine learning, Deep learning, Energy management

1 Introduction

With global climate change, summer temperatures have increased year by year, resulting in the continuous growth of Taiwan's electricity demand. To ensure the stability of the power supply, the government has actively promoted renewable energy and energy conservation measures. The Advanced Metering Infrastructure (AMI) is one of the important tools. AMI, composed of smart meters, communication systems, and meter information management systems, is the most important construction for achieving smart grid goals. AMI has a communication function, including connecting electricity, power generation and storage systems, collecting data from multiple parties and achieving instant communication and communication of grid information. AMI provides a communication pipeline, so that supply at the demand side can actively participate in the dispatch of the power company so that the supply and demand sides can simultaneously profit, and the time price is one of the applications profited. Time-consuming electricity prices, by pricing the difference between peak and off-peak electricity consumption period, prompts users to change their original electricity use habits, reduces electricity use during the peak period. At present, high-voltage electricity consumers with high electricity consumption, such as commercial buildings and factories, have applied time-consuming electricity prices. As for low-voltage electricity users like residential and small stores, traditional electricity prices are relatively low, resulting in low prices. Generally, households have a lower willingness to choose time-consuming electricity prices. However, according to the Taipower Company analysis, the low-voltage power consumption during the peak period has accounted for more than 51%. Therefore, it is hoped that these users should apply time-consuming electricity price to reduce electricity usage during the peak period.

In addition, the Government expects to complete the establishment of 1 million smart meters in Taiwan by 2020. It is expected that a large amount of valuable electricity data will be generated. How to convert such data into effective information depends on the support of information science. To this end, this study will use the household smart meter data, taking advantage of deep learning to predict the electricity consumption of each user's future electricity usage, and hope to

*Corresponding Author: Shin-Jye Lee; E-mail: camhero@gmail.com
DOI: 10.3966/160792642020012101026

accurately predict the electricity usage. With data visualization, users can better understand the advantage of using time-consuming price. On the other hand, these predict results can also be provided to Taipower Company to plan more time-consuming electricity prices that are more in line with demand. Deep learning is a kind of machine learning that can analyze large amounts of data messages and handle more complex problems in a short period of time. When using deep learning for prediction, the results obtained by different algorithms are often different, and there is no so-called optimal solution. This study will establish models with multiple deep learning algorithms and experiment with a large number of models. Adjust with the parameters to find the model and parameters that are most suitable for this study.

Deep learning is a kind of machine learning that can analyze large amounts of data messages and handle more complex problems in a short period time. When using deep learning for prediction, the results obtained by different algorithms are often different, and there is no so-called optimal solution. This study will establish models with multiple deep learning algorithms and experiment with a large number of models. Adjust with the parameters to find the model and parameters that are most suitable for this study. Both Chang and Tseng [1], Liu et al. [2], used neural networks to achieve good result from different domain.

2 Related Works

Neural network is a mathematical model that mimics the structure and function of a biological neural network, using data for learning and summarizing. The existing large research literature applies it to energy consumption, power forecasting, electricity price analysis and power signal analysis. Wu [3] used integrated fuzzy systems, data mining methods and neural networks to explore single-family short-term load forecasting, and estimate and predict the 24-hour electricity usage change in the future. Based on the characteristics of short-term load, the influence of weather conditions and the user's lifestyle, this study summarized five kinds of simulation scenarios, with the load data during January-June, there are 100 single families in Yongkang City, Tainan, and the weather data from the Central Meteorological Bureau. As a predictive sample, the data was finally optimized with an integrated fuzzy neural network to obtain an accuracy of 3.68%. Su [4] studied the electrical load and daily meteorological data of residential areas and proposed a short-term electric usage forecasting model based on neural network-like residential areas. The historical temperature (T) and humidity (W) collected by the Keelung Meteorological Observatory were analyzed by the highest historical load (P) of a residential area in Keelung City from 2011-2012. The results show that the neural network trained by the Bayesian regularization algorithm has a mean error percentage of 4.924%. Li [5] used a neural-like algorithm to construct a training network through various reference factors that may affect the behavior of electricity consumption and then used the training network to predict the electricity consumption in Taiwan. According to the forecast results, provide the trend of electricity usage in each month of the coming year, and accurately predict the future consumption of electricity to judge the adjustment of load and energy consumption; in addition, the research results also show average temperature, average relative humidity, average The load, total domestic energy consumption, average hours and oil consumption combined to obtain an average absolute percentage error (MAPE) of 1.90% is the best prediction accuracy, and understand that climate change and the length of sunshine have a considerable degree of change in the behavior of Taiwan people.

Due to the prediction results, we can understand the power consumption behavior of the whole of Taiwan. The related strategy can be planned in advance to avoid future electricity shortage when Taipower Company predicts that there may have high energy consumption in the future.

3 Data Processing

The power consumption data of the low-voltage smart meter user in this paper was collected from ten different districts in Taiwan from 2016-2018 by Industrial Technology Research Institute. There are about 1000 household consumption data being collected. The data was collected and saves by each distinct household, that is, every household will have its prediction model. The information recorded in the 15-minute interval and also accumulated the previous period electricity consumption. Therefore, each piece of data needs to be differentiated to obtain the actual power consumption of each user for every 15 minutes. Also, to avoid the excessive amount of data and that causing training time be too long, we will integrate every four data, that is, the data will be shown by every hour. The next thing to deal with is the missing value. The following Figure 1 is one of the user's electricity consumption information from October 2017 to March 2018. The x-axis represents each different month, and the y-axis represents the electricity usage every day (kilowatt-hour). It can be seen that there is a missing value in the middle. In addition to this long-term missing value, there are also days and hours without data, so we take measures to resolve these errors.

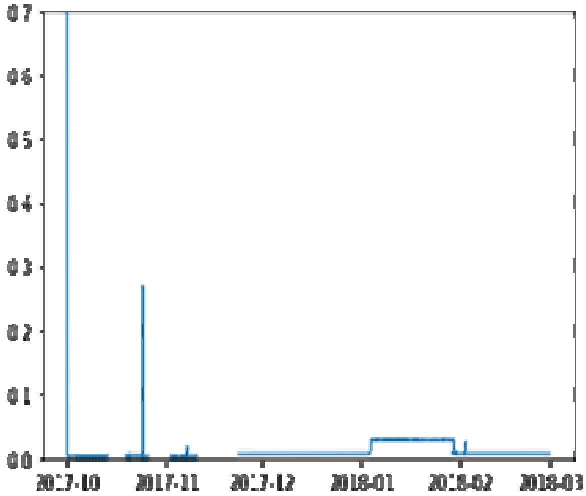

Figure 1. Example of original data

Moreover, we assume that electricity usage is similar at the same time every week. For example, the electricity usage on Monday afternoon may be similar to the electricity usage next Monday afternoon, so when there is a missing value in a period, it will search for a month before and after the time. The electricity consumption at the same time as each week (a total of eight periods will be obtained) and averaged. If eight of the same time is also missing value, it will be discarded. According to the literature review, weather conditions are important factors affecting electricity consumption and there are many open data about the weather that we can easily obtain. Through the web crawler program, we retrieve the meteorological data of the user's corresponding area. The meteorological data used include temperature, humidity and rainfall probability.

4 The Proposed Deep Learning Model

The neural network algorithms used in this study are iterative, hoping to find the Global Minimum. The prediction based on future power usage is related to the application of time series. We use the recurrent neural network (RNN), suitable for time series and its variant, Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) as a predictive model. When training the model, we use 80% of the data to train a model, and the remaining 20% as a Testing Data set, which was used to verify whether overfitting occurred.

4.1 Inference Model-Recurrent Neural Network (RNN)

When using traditional neural network architecture, each piece of data is found to be independent of each other, without considering the implicit association between each piece of data. Therefore, when there is a relationship between the current input data and the next data (such as time series, natural language, etc.), it is more suitable to use the Recurrent Neural Network (RNN) [6-8]. RNN is suitable for dealing with such problems because RNN stores the state St-1 of the neuron at time t-1 and multiplies the weight W to obtain the current time state St. The overall structure is shown in Figure 2.

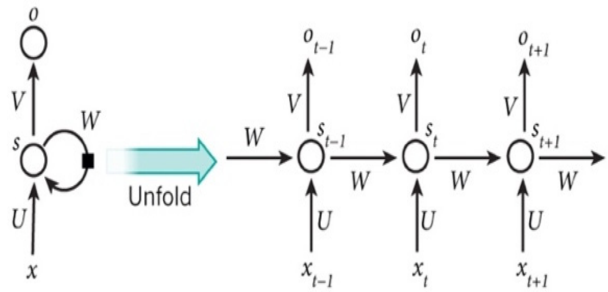

Figure 2. Structure of RNN

Source: http://www.wildml.com/2015/09/recurrent-neura-networks-tutorial-part-1-introduction-to-rnns/.

U and *V* are the Weight Matrixes of the Input layer and Output layer respectively, and every input shares the same weight among different input.

4.2 Long-Short Term Memory (LSTM)

Since the general recurrent neural network is prone to the problem of gradient disappearance or gradient explosion in the case of backpropagation. If dealing with a longer sequence of data, performance will not be as expected. Therefore, Sepp Hochreiter & Jürgen Schmidhuber [9] proposed a new improved RNN model in 1997: Long-Short Term Memory (LSTM) [10-11]. The LSTM sets three thresholds that control input, memory, and output. The principle is to add a logic function to the three gates so that the three threshold values will between 0 and 1. If the value of the Input Gate is approximately 0, the value here will be blocked and will not be delivered to the next layer. If the value of the Forget Gate is approximately 0, the value of the block memory is forgotten. If the value of the Output Gate is approximately 0, it means there won't have any output. The overall structure is shown in Figure 3. x_t, C_t, and h_t are the inputs, memory cell and hidden state in time t respectively, used to calculate the output of different steps. σ represents sigmoid function, which is between 0 and 1 and every parameter shares the same weight among different input.

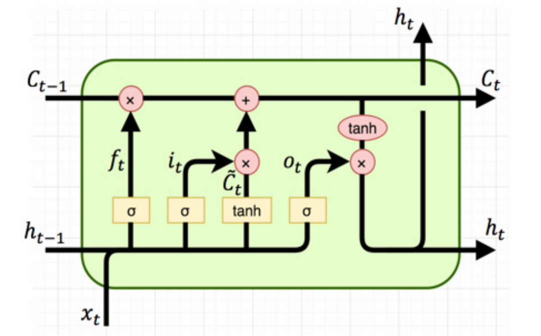

Figure 3. Structure of LSTM

Source: https://medium.com/@saurabh.rathor092/simple-rnn-vs-gru-vs-lstm-difference-lies-in-more-flexible-control-5f33e07b1e57.

4.3 Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) [12] has two gates: a reset gate r and an update valve *z*. Intuitively, the reset gate determines the combination of the new input and the memory of the previous moment, and the update gate determines the degree of retention of the previously remembered information. If all reset valves are set to 1 and all update valves are set to 0, the traditional RNN model can be obtained again. The overall structure is shown in Figure 4. In the figure, *z* represents Update Gate and *r* represents Reset Gate respectively, and every parameter shares the same weight among different input.

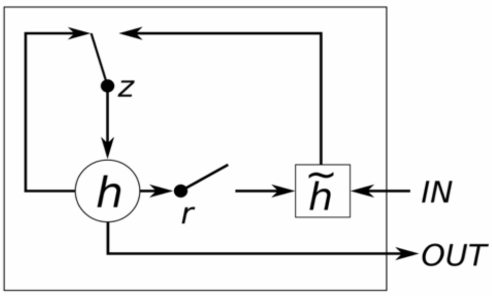

Figure 4. Structure of GRU

Source: http://www.gabormelli.com/RKB/Gated_Recurrent_Unit.

The basic principle of using the gate mechanism to learn long-term memory in GRU is the same as that of LSTM, but there are some differences: GRU has two gates, while LSTM has three gates. There is no difference between the GRU and the internal memory unit (*c_t*), and there is no output gate in the LSTM. The output gate and forget gate of LSTM are integrated into an update gate z in the GRU; the reset gate r is used directly in the previous hidden state. Therefore, the function of the reset gate in the LSTM is replaced by the reset gate r and the update gate z in the GRU. When calculating the output, the nonlinear function is no longer used. GRU has fewer parameters, so it can be trained faster and requires less data to be summarized. Correspondingly, if there is enough training data, LSTM may be a better choice.

5 Simulation Experiments

We use three models for prediction, LSTM, five-layer LSTM, and GRU, which predict the amount of electricity usage per hour for each of the two kinds of time scales, next day and next week. According to different time scale as the forecast target, different historical usage data are used for training. The standard used to measure the prediction accuracy is the Mean Absolute Percentage Error (MAPE). Compared with other error definitions, MAPE is a much more objective measurement, as represented in Equation 1.

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \quad \textbf{(1)}$$

where *At* is the actual value and *Ft* is the predicted value.

In 1982, scholar Lewis [13] divided the model prediction ability into four grades according to the size of MAPE, as shown in Table 1.

Table 1. MAPE criterion

MAPE	**< 10%**	**10%~20%**	**20%~30%**	**>50%**
Represent	Excellent	Good	Reasonable	Poor

Also, the input of our model are the electricity usage and meteorological data (temperature. humidity. rainfall probability) on that hour, which is shown in Table 2, and the output is the future electricity usage we are going to predict. Table 3 shows the parameter setting of each model.

Table 2. Input data

Time	MeterID	Value	Temp	Humidiy	Rain
10/1 00	AD1729	0.9935	27.5	83	0.0
10/1 01	AD1729	0.0117	27.9	74	0.0
10/1 02	AD1729	0.4161	27.4	75	0.0
10/1 03	AD1729	0.4155	27.6	74	0.0
10/1 04	AD1729	0.4149	25.6	77	0.0
10/1 05	AD1729	0.4141	26.7	79	0.0
10/1 06	AD1729	0.4149	26.6	79	0.0

Table 3. Model parameters

Model	Cell Size	Learning Rate	Batch Size	Training Step
5-layer LSTM	64	0.05	256 (prediction of next day) 128 (prediction of next week)	3001 (prediction of next day) 5001 (prediction of next week)
LSTM	128	0.05	256 (prediction of next day) 128 (prediction of next week)	3001 (prediction of next day) 5001 (prediction of next week)
GRU	64	0.05	256 (prediction of next day) 128 (prediction of next week)	3001 (prediction of next day) 5001 (prediction of next week)

As the prediction of the next day (24 hours), we used the data of the previous 48 hours to predict the next 24 hours, getting the best MAPE 6.38%, Figure 5 shows the prediction results of some of Testing Data. The blue line is the result of the model prediction and the orange line is the actual electricity usage. As we can see from the figure, the accuracy of the electricity consumption forecast for the next day is very high, and it can accurately capture the electricity usage of the user every hours. The x-axis represents each hour, and the y-axis represents the electricity usage (kilowatt-hour).

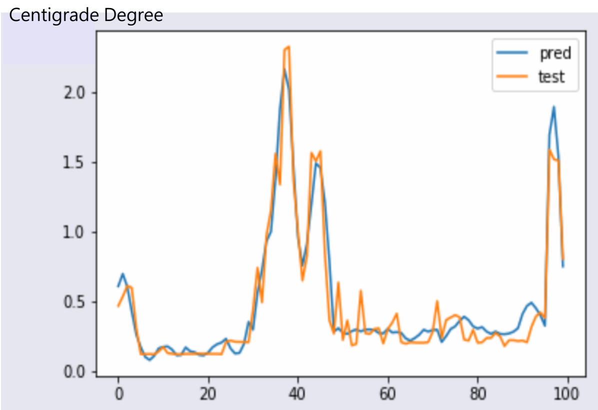

Figure 5. Result of prediction of next day

As the prediction of next week (168 hours), we used the data of previous 168 hours to predict the next 168 hours, getting the best MAPE 3.59%, Figure 6 shows the prediction results of some of Testing Data. The blue line is the result of the model prediction and the orange line is the actual electricity usage. As we can see from the figure, is the forecast of electricity usage of the next week, the forecast result per hour is not very well. Although it can be known whether the trend of electricity consumption in the hour is upward or downward, it has gradually failed to grasp the peak. However, in the prediction of next week, MAPE is lower than the prediction of the next day, which means that although the accurate electricity consumption per hour is difficult to predict in the next week, for the entire month, each user's electricity usage amount is not much changed. The x-axis represents each different hour, the y-axis represents the electricity usage (kilowatt-hour).

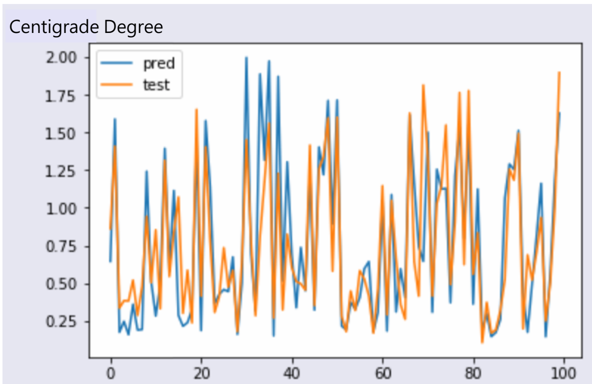

Figure 6. Result of prediction of next week

Table 4 shows the different prediction results obtained by different algorithms. The implementations of these experiments are in Python with the help of TensorFlow [14]. We run all the experiments on a computer with a single Intel i5-7500 CPU, without GPU. It can be seen that the five-layer LSTM consumed the most time in all two kinds of time scales, had the best results, but it takes a lot of computing resource. If using the forecast results and the resources consumption as the judging standard, using GRU may be a good choice.

Table 4. Prediction result

		Next Day	Next Week
LSTM	MAPE	19.18%	15.16%
	Time Consumption	8 mins	53 mins
five-layer LSTM	MAPE	6.38%	3.59%
	Time Consumption	19 mins	192 mins
GRU	MAPE	11.23%	6.92%
	Time Consumption	8 mins	51 mins

Table 5. Model setting

Model	Predict Scale	Cell Size	Learning Rate	Batch Size	Training Step
LSTM	Next day	128	0.05	256	5001
	Next week	128	0.05	256	5001
Five-layer LSTM	Next day	64	0.05	256	3001
	Next week	64	0.05	256	3001
GRU	Next day	128	0.05	256	5001
	Next week	128	0.05	256	5001

Table 5 shows the model setting in different tests, Five-layer LSTM can use lower cell size and training steps to outperform other models.

6 Conclusion

In this paper, we have successfully applied the deep learning method to the electricity usage prediction and got excellent results that haven't be done. We have got good results in predicting the electricity usage in the next day. As the prediction in next week, because of its long-timescale, although it is impossible to accurately predict each hour, but still can catch the trend of the overall power consumption, and the prediction results can be visualized so that users can realize whether they are suitable for time-consuming electricity price. The forecast for usage in the next day can also be used as the basis for Taipower Company's power dispatching because as for load dispatch, it needs the nearer upcoming prediction, rather than farther one. Till this moment, we need to train different model for different household because of their different electricity consumption behavior, we hope that we can create a more general model to fit more households at one time in the future. For the next step, we hope to classify these users according to different electricity use habits, so that Taipower Company can provide a more flexible and suitable time price for different types of users.

Acknowledgements

This research is partially supported by the Ministry of Science and Technology research grant in Taiwan (MOST 108-2218-E-009-041-).

References

[1] J.-H. Chang, C. Y. Tseng, Analyzing Google Trends with Travel Keyword Rankings to Predict Tourists into a Group, *Journal of Internet Technology*, Vol. 20, No. 1, pp. 247-256, January, 2019.

[2] L. Liu, P. Wu, X. Wang, Short-time Traffic Flow Prediction Based on Wavelet Neural Network, *Journal of Internet Technology*, Vol. 20, No. 4, pp. 1237-1246, July, 2019.

[3] J.-C. Wu, *An Integrated Research on Short-Term Load Forecasting of a Single-Person Family in Taiwan: Application of Fuzzy Logic Controllers and Artificial Neural Networks*, Master Thesis, National Chung Hsing University, Taichung, Taiwan, 2012.

[4] C.-P. Su, *The Study of Short-term Load Forecast Based on the Neural Network*, Master Thesis, National Taiwan Ocean University, Keelung, Taiwan, 2014.

[5] T.-C. Li, *Prediction of Electric Power Consumption in Taiwan by Back Propagation Neural Network*, Master Thesis, National Chung Cheng University, Chiayi, Taiwan, 2016.

[6] D. Mandic, J. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, Wiley, 2001.

[7] P. Campolucci, A. Uncini, F. Piazza, B. D. Rao, On-Line Learning Algorithms for Locally Recurrent Neural Networks, *IEEE Transactions on Neural Networks*, Vol. 10, No. 2, pp. 253-271, March, 1999.

[8] C. L. Giles, C. B. Miller, D. Chen, H. H. Chen, G. Z. Sun, Y. C. Lee, Learning and Extracting Finite State Automata with Second-Order Recurrent Neural Networks, *Neural Computation*, Vol. 4, No. 3, pp. 393-405, May, 1992.

[9] S. Hochreiter, J. Schmidhuber, LSTM Can Solve Hard Long Time Lag Problems, *Advances in Neural Information Processing Systems 9*, Denver, CO, USA, 1996, pp. 473-479.

[10] J. Schmidhuber, F. Gers, D. Eck, Learning Nonregular Languages: A Comparison of Simple Recurrent Networks and LSTM, *Neural Computation*, Vol. 14, No. 9, pp. 2039-2041, September, 2002.

[11] J. A. Perez-Ortiz, F. A. Gers, D. Eck, J. Schmidhuber, Kalman Filters Improve LSTM Network Performance in Problems Unsolvable by Traditional Recurrent Nets, *Neural Networks*, Vol. 16, No. 2, pp. 241-250, March, 2003.

[12] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*, https://arxiv.org/abs/1412.3555.

[13] C. D. Lewis, *Industrial and Business Forecasting Methods*, Butterworths, 1982.

[14] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, *Tensorflow: Large-scale Machine Learning on Heterogeneous Distributed Systems,* https://arxiv.org/abs/1603.04467.

Biographies

Lu-Xian Wu is a postgraduate at the Institute of Technology Management, National Chiao Tung University, Taiwan. His research interests mainly focus on data science, especially in financial field.

Shin-Jye Lee is currently with the Institute of Technology Management, National Chiao Tung University, Taiwan. Before that, he was the Professor of National Pilot School of Software, Yunnan University, China, and he also made his academic career in Poland. Further, his research interests primarily comprise machine learning and technology policy.