

Deep Learning and Machine Learning Models to Predict Energy Consumption in Steel Industry

Kittisak Kerdprasop*, Nittaya Kerdprasop, and Paradee Chuaybamroong

Abstract—This paper present the study results of predicting energy consumption in the steel industry using modeling methods based on machine learning and deep learning techniques. Machine learning algorithms used in this work include artificial neural network (ANN), k-nearest neighbors (kNN), random forest (RF), and gradient boosting (GB). Deep learning technique is long short-term memory (LSTM). Linear regression, which is the statistical-based learning algorithm, is also applied to be the baseline of this comparative study. The modeling results reveal that among the statistical-based and machine learning-based techniques, GB and RF are the best two models to predict energy consumption, whereas ANN shows the predictive performance comparable to the linear regression model. Nevertheless, LSTM outperforms both statistical-based and machine learning-based algorithms in predicting industrial energy consumption.

Index Terms—Energy consumption prediction, deep learning, machine learning, long short-term memory, ensemble model

I. INTRODUCTION

Efficient energy usage planning and management is the primary concern in smart buildings and manufacturing. Researchers and practitioners in both sectors have long been search for accurate methods to estimate the amount of energy usage for proper design and development in smart building [1]. In the manufacturing process, efficient energy usage is also important to numerous industrial applications such as petrochemical industries, iron and steel mills, and many others.

In recent years, a data-driven approach using machine learning algorithms to build a model for predicting amount of energy use has gained much popularity among researchers. In the past decade, a non-linear algorithms such as support vector machine (SVM) and support vector regression (SVR) had been applied to the problem [2–5]. The SVM and SVR algorithms produce a single model to predict future value of energy use. A more sophisticate method deploys group of models to cooperatively predict energy usage. Such method is called ensemble machine learning. Many types of learning algorithms can be applied in the ensemble scheme, for example, swarm intelligence [6], a group of regression trees [7], and decision trees [8]. The two ensemble learning algorithms that had been proven accurate and effective for predicting energy consumption are random forest [9–11] and gradient boosting [12–14].

Recent advancement in the field of machine learning is the

introduction of deep learning technique. It has been widely used with promising results in many application areas including energy consumption estimation. The applied deep learning techniques are varied from auto-encoder [15, 16], Boltzmann machine [17], deep recurrent neural network [18–21], to long short-term memory [22, 23]. Most research work based on deep learning techniques report accurate results that outperform single modeling and ensemble learning methods.

This work thus propose a comparative study to assess experimentally the predictive performance of machine learning-based against deep learning-based modeling methods. The single-model machine learning algorithms include k-nearest neighbors and artificial neural network. The ensemble algorithms are random forest and gradient boosting. Deep learning algorithm is long short-term memory. The algorithm linear regression, which is statistical-based approach, is also applied to be a baseline of model performance comparison. Details of dataset and research methodology are explained in Section II. Experimental results are reported in Section III. Then, the conclusion are presented in Section IV.

II. METHODOLOGY

A. Data

Energy consumption data used in this research are obtained from the UCI repository [24]. This energy data are publicly available by the Daewoo Steel Company in South Korea. The original data are comprised of 11 attributes, but this research extracts only two of them (i.e., date-time and energy usage) to be used in the modeling process. Energy usage data had been recorded continuously in a 15-minute interval starting from the time 00:15 of January 1, 2018 and end recording at the time 00:00 of December 31, 2018 resulting in 35,041 data records. The unit of energy usage is kilowatt-hour (kWh). Within the year 2018, the maximum energy usage was 157.18 kWh, while the minimum was 0 and the standard deviation was 33.44. Therefore, the dataset used in this work is time series and characteristic of the series can be illustrated as in Fig. 1.

B. Predictive Modeling Method

The modeling process of this research is comprised of three main phases: data pre-processing, model creation, and model evaluation (as displayed in Fig. 2).

At the first phase of data pre-processing, there are two steps, that are, feature extraction and data normalization. Feature extraction is the step to select two variables (or attributes/features) from the original dataset. The selected attributes are date-time and energy consumption in the steel manufacturing. Energy consumption data were recorded continuously every 15 minutes for the whole year of 2018.

Manuscript received January 10, 2023; revised February 20, 2023; accepted March 28, 2023.

Kittisak Kerdprasop and Nittaya Kerdprasop are with the School of Computer Engineering, Suranaree University of Technology, Thailand.

Paradee Chuaybamroong is with the Department of Environmental Science, Thammasat University, Thailand.

*Correspondence: KittisakThailand@gmail.com (K.K.)

The pre-processed dataset is thus consisted of two features and 35,041 instances (or records). The values of energy consumption fluctuate in the range of 0 to almost 160 kWh. Data normalization to decrease the range is therefore applied in order to obtain an accurate model in the subsequent phase. We adopt z-score normalization as shown in (1) when E' is the normalized energy consumption value, E is the original value of energy consumption, μ is the mean value of energy consumption and σ is standard deviation.

$$E' = \frac{E - \mu}{\sigma} \quad (1)$$

The second phase of this research is model creation. We utilize three types of modeling algorithms: statistical-based, machine learning, and deep learning.

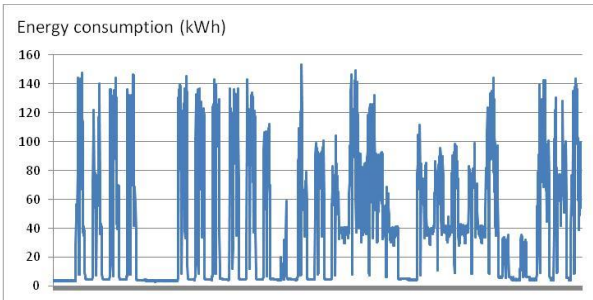


Fig. 1. Industry energy consumption (kWh) during January 1st to 31st, 2018 measured as a continuous value in every 15 minutes.

The statistical-based learning algorithm is linear regression. It is to be used as the baseline to compare performance of predictive models.

The machine learning-based algorithms are composed of two categories of learning algorithms: single model and a group of models (or ensemble). Learning algorithms that generate a single model for predicting future values are k-nearest neighbors (kNN) and artificial neural network (ANN). Ensemble algorithms that use a group of models to work in a cooperative manner in forecasting future values of energy consumption are random forest (RF) and gradient boosting (GB). In the first step of machine learning-based modeling process, there are ten more features (namely, lag1 up to lag10) to be added to the dataset. The augmented features are lagged data of energy consumption, which are data records in the previous ten periods of time. This feature augmentation step is for producing an accurate model.

The deep learning algorithm is long short-term memory (LSTM). The algorithm LSTM is a kind of regression network that utilizes historical data sequence to estimate future value. LSTM works in an iteration manner such that at each round (or epoch) a set of historical time series is used for training the network to predict future series one at a time and then shifts the series by one step and update the network state to predict the next time series.

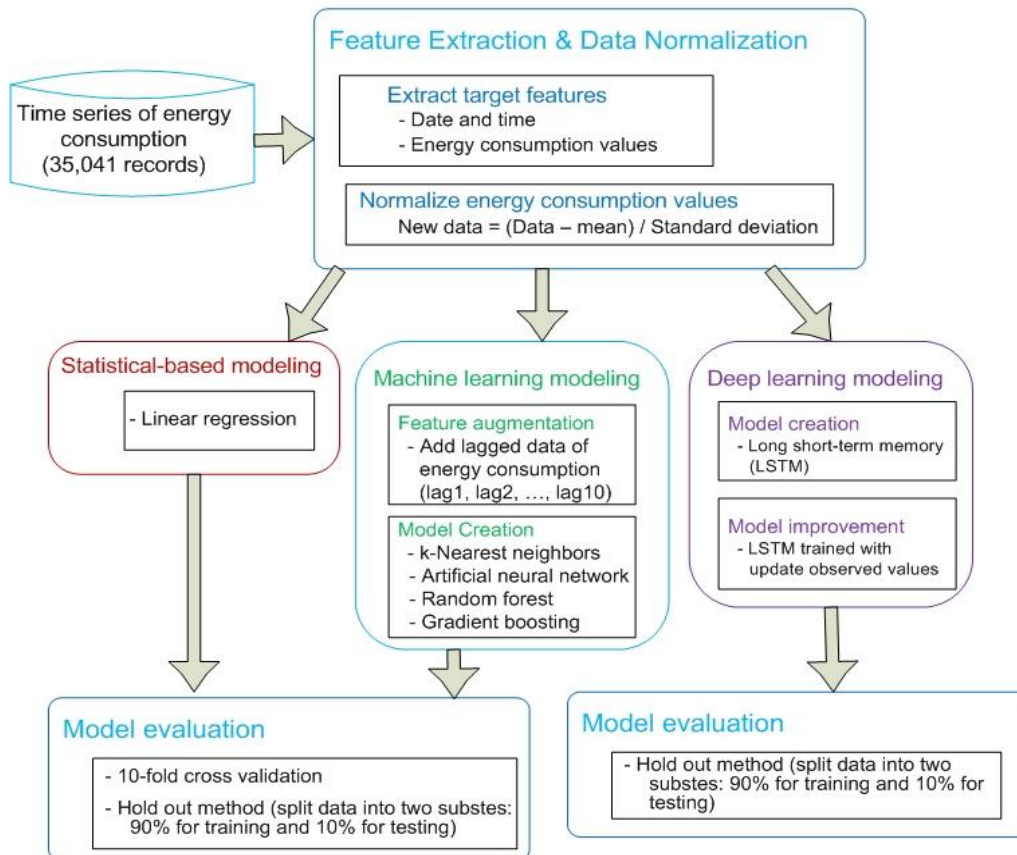


Fig. 2. The data-driven modeling process for predicting energy consumption in steel industry.

We specify the LSTM network to compose of 200 hidden units and train the network for 250 epochs. Initial learning rate of LSTM is 0.005 and then decrease the rate by a factor of 0.2 after 125 epochs. The improved LSTM model is also

generated. The performance of LSTM can be improved by using observed values instead of the predicted value to update the network.

The last phase of this research framework is model

evaluation. For statistical-based and machine learning-based modeling schemes, two types of model assessment have been deployed, that are, cross validation and hold out methods. Cross validation has been repeated ten times using ten data subsets. For each of the ten iteration, nine data subsets had been used as training set and the remaining one subset is for testing. Each round a different test set is applied. The hold out method separates data into two subsets: 90% of the data are for training and the remaining 10% are for testing. In deep learning, it is forecasting a time series such that the order of series is important, therefore cross validation does not make sense. We thus perform only the hold out evaluation method.

III. EXPERIMENTATION AND RESULTS

A. Performance of Machine Learning Models

To compare predictive performance of models built from different types of learning algorithms, we adopt three evaluation metrics: root mean squared error (RMSE), mean absolute error (MAE), and R-squared (R^2). RMSE and MAE measure the difference between actual energy consumption values and the values predicted by the model. Thus, the lower is the better. The metric R^2 (or coefficient of determination) represents the goodness of fit showing how well the data fit the regression model. The higher R^2 is the better because the model can capture more variability in the dataset.

Results of model performance evaluation using 10-fold CV method are illustrated in Table I. Gradient boosting is the best algorithm with lowest error and highest R^2 . Random forest is the second best one. Algorithm k-nearest neighbor comes in the third place showing performance better than artificial neural network. Linear regression is the worst model.

TABLE I: PERFORMANCE OF FIVE MACHINE LEARNING MODELS ON PREDICTING ENERGY CONSUMPTION (ASSESSED BY 10-FOLD CROSS VALIDATION METHOD)

Modeling Algorithm	RMSE	MAE	R^2
Linear Regression	13.521	7.189	0.837
k-Nearest Neighbors	12.850	6.173	0.852
Artificial Neural Network	13.520	7.173	0.837
Random Forest	12.318	5.812	0.864
Gradient Boosting	11.987	5.820	0.872

B. Performance of Deep Learning Model

Performance of the LSTM model to forecast future values of energy consumption is shown in Fig. 3. The network has been updated with the predicted values thus making RMSE quite high at 29.0043. When the model has been improved by using observed values to update the network, the RMSE is as low as 7.957 (in Fig. 4).

Predictive performances of the three groups of learning algorithms (i.e., statistical-based, machine learning, deep learning) are summarized and shown in Table II. The improved long short-term memory show the best performance when assessed the model with hold out method. Random forest and gradient boosting are quite comparable and come in the second and third place, respectively. Linear regression and artificial neural network are almost at the same performance level, whereas k-nearest neighbors shows

the worst performance. It can be noticed from the experimental results that k-nearest neighbors is subjective to the test set in the sense that using different test data, predictive performance can be affected significantly.

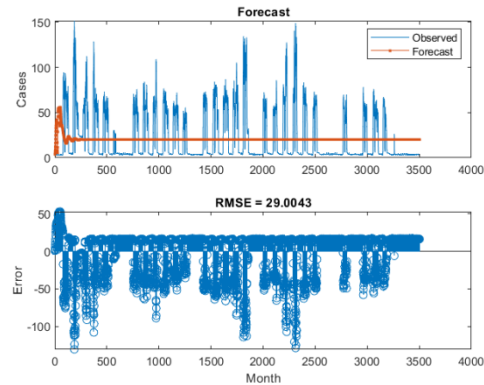


Fig. 3. Predictive performance of LSTM using predicted values to update network.

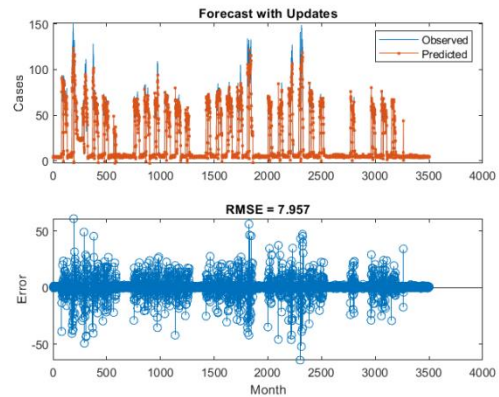


Fig. 4. Performance of LSTM that has been improved by using the observed values to update the network.

TABLE II: PREDICTIVE PERFORMANCE OF DEEP LEARNING MODEL COMPARATIVE TO MACHINE LEARNING MODELS (ASSESSED BY HOLD-OUT METHOD USING 10% OF DATA AS A TEST SET)

Type of Algorithm	Modeling Algorithm	RMSE
Statistical-based	Linear Regression	13.590
	k-Nearest Neighbors	16.430
Machine learning	Artificial Neural Network	13.593
	Random Forest	12.059
	Gradient Boosting	12.117
Deep learning	Long Short-Term Memory	7.957

IV. CONCLUSION

This research presents the experimental studied results of applying data-drive methods to predict energy consumption, which is the area of interest in many domains including manufacturing and smart buildings. Modeling methods to forecast energy usage can utilize different schemes such as physical modeling based on accurate equation formulation, intelligent modeling based on historical data, and a hybrid method. This research focuses on the intelligent modeling method because of its less time consuming compared to the physical method. The intelligent modeling method is based on the application of various machine learning algorithms.

Performance of intelligent modeling using three different

groups of machine learning schemes has been explored. The machine learning schemes include statistical-based using linear regression algorithm, the machine learning-based using k-nearest neighbors, artificial neural network, random forest, and gradient boosting algorithms, and deep learning using long short-term memory algorithm. The experimental results reveal that deep learning is the best method to be applied for energy consumption prediction, whereas the ensemble scheme using random forest and gradient boosting algorithms are among the second best one.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

The first author is the main contributor in designing research process and experimentation. The second author is responsible for manuscript preparation and submission. The third author helps revising the manuscript.

REFERENCES

- [1] Y. Sun, F. Haghghat, and B. C. M. Fung, "A review of the-state-of-the-art in data-driven approaches for building energy prediction," *Energy and Buildings*, vol. 221, 110022, Aug. 2020. doi: 10.1016/j.enbuild.2020.110022
- [2] F. Zhang, C. Deb, S. E. Lee, J. Yang, and K. W. Shah, "Time series forecasting for building energy consumption using weighted support vector regression with differential evolution optimization technique," *Energy and Buildings*, vol. 126, pp. 94–103, Aug. 2016. doi: 10.1016/j.enbuild.2016.05.028
- [3] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, "Vector field-based support vector regression for building energy consumption prediction," *Applied Energy*, vol. 242, pp. 403–414, May 2019. doi: 10.1016/j.apenergy.2019.03.078
- [4] S. Paudel *et al.*, "A relevant data selection method for energy consumption prediction of low energy building based on support vector machine," *Energy and Buildings*, vol. 138, pp. 240–256, Mar. 2017. doi: 10.1016/j.enbuild.2016.11.009
- [5] M. K. M. Shapi, N. A. Ramli, and L. J. Awalin, "Energy consumption prediction by using machine learning for smart building: Case study in Malaysia," *Developments in the Built Environment*, vol. 5, 100037, Mar. 2021. doi: 10.1016/j.dibe.2020.100037
- [6] K. Li, J. Tian, W. Xue, and G. Tan, "Short-term electricity consumption prediction for buildings using data-driven swarm intelligence based ensemble model," *Energy and Buildings*, vol. 231, 110558, Jan. 2021. doi: 10.1016/j.enbuild.2020.110558
- [7] Z. Wang, Y. Wang, and R. S. Srinivasan, "A novel ensemble learning approach to support building energy use prediction," *Energy and Buildings*, vol. 159, pp. 109–122, Jan. 2018. doi: 10.1016/j.enbuild.2017.10.085
- [8] Z. Dong, J. Liu, B. Liu, K. Li, and X. Li, "Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification," *Energy and Buildings*, vol. 241, 110929, Jun. 2021. doi: 10.1016/j.enbuild.2021.110929
- [9] Z. Wang, Y. Wang, R. Zeng, R. S. Srinivasan, and S. Ahrentzen, "Random forest based hourly building energy prediction," *Energy and Buildings*, vol. 171, pp. 11–25, Jul. 2018. doi: 10.1016/j.enbuild.2018.04.008
- [10] A. González-Vidal, V. Moreno-Cano, F. Terroso-Sáenz, and A. F. Skarmeta, "Towards energy efficiency smart buildings models based on intelligent data analytics," *Procedia Computer Science*, vol. 83, pp. 994–999, 2016. doi: 10.1016/j.procs.2016.04.213
- [11] V. E. Sathishkumar *et al.*, "Industry energy consumption prediction using data mining techniques," *International Journal of Energy, Information and Communications*, vol. 11, no. 1, pp. 7–14, Feb. 2020. doi: 10.21742/IJEIC.2020.11.1.02
- [12] S. Touzani, J. Granderson, and S. Fernandes, "Gradient boosting machine for modeling the energy consumption of commercial buildings," *Energy and Buildings*, vol. 158, pp. 1533–1543, Jan. 2018. doi: 10.1016/j.enbuild.2017.11.039
- [13] Y. Ding, L. Fan, and X. Liu, "Analysis of feature matrix in machine learning algorithms to predict energy consumption of public buildings," *Energy and Buildings*, vol. 249, p. 111208, Oct. 2021. doi: 10.1016/j.enbuild.2021.111208
- [14] B. Carrera, S. Peyrard, and K. Kim, "Meta-regression framework for energy consumption prediction in a smart city: A case study of Songdo in South Korea," *Sustainable Cities and Society*, vol. 72, 103025, Sep. 2021. doi: 10.1016/j.scs.2021.103025
- [15] C. Fan, F. Xiao, and Y. Zhao, "A short-term building cooling load prediction method using deep learning algorithms," *Applied Energy*, vol. 195, pp. 222–233, Jun. 2017. doi: 10.1016/j.apenergy.2017.03.064
- [16] C. Fan, Y. Sun, Y. Zhao, M. Song, and J. Wang, "Deep learning-based feature engineering methods for improved building energy prediction," *Applied Energy*, vol. 240, pp. 35–45, Apr. 2019. doi: 10.1016/j.apenergy.2019.02.052
- [17] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, "Deep learning for estimating building energy consumption," *Sustainable Energy, Grids and Networks*, vol. 6, pp. 91–99, Jun. 2016. doi: 10.1016/j.segan.2016.02.005
- [18] R. Olu-Ajayi, H. Alaka, I. Sulaimon, F. Sunmola, and S. Ajayi, "Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques," *Journal of Building Engineering*, vol. 45, 103406, Jan. 2022. doi: 10.1016/j.jobe.2021.103406
- [19] C. H. Kim, M. Kim, and Y. J. Song, "Sequence-to-sequence deep learning model for building energy consumption prediction with dynamic simulation modeling," *Journal of Building Engineering*, vol. 43, 102577, Nov. 2021. doi: 10.1016/j.jobe.2021.102577
- [20] L. Lei, W. Chen, B. Wu, C. Chen, and W. Liu, "A building energy consumption prediction model based on rough set theory and deep learning algorithms," *Energy and Buildings*, vol. 240, 110886, Jun. 2021. doi: 10.1016/j.enbuild.2021.110886
- [21] C. Fan, J. Wang, W. Gang, and S. Li, "Assessment of deep recurrent neural network-based strategies for short-term building energy predictions," *Applied Energy*, vol. 236, pp. 700–710, Feb. 2019. doi: 10.1016/j.apenergy.2018.12.004
- [22] K. Gaur and S. K. Singh, "CNN-Bi-LSTM based household energy consumption prediction," in *Proc. 2021 3rd International Conference on Signal Processing and Communication (ICSPSC)*, Coimbatore, India, May 2021, pp. 233–237. doi: 10.1109/ICSPSC51351.2021.9451797
- [23] Y. Gao and Y. Ruan, "Interpretable deep learning model for building energy consumption prediction based on attention mechanism," *Energy and Buildings*, vol. 252, p. 111379, Dec. 2021. doi: 10.1016/j.enbuild.2021.111379
- [24] D. Dua and C. Graff (2019). UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Steel+Industry+Energy+Consumption+Dataset>

Copyright © 2023 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).