



Article Feature Extraction of Anomaly Electricity Usage Behavior in Residence Using Autoencoder

Chia-Wei Tsai ¹^[b], Kuei-Chun Chiang ^{1,2}, Hsin-Yuan Hsieh ¹^[b], Chun-Wei Yang ^{3,4}^[b], Jason Lin ⁵^[b] and Yao-Chung Chang ^{1,*}

- ¹ Department of Computer Science and Information Engineering, National Taitung University, No. 369, Section 2, University Road, Taitung 95092, Taiwan; cwtsai@nttu.edu.tw (C.-W.T.); chun@iii.org.tw (K.-C.C.); 0911137@gm.nttu.edu.tw (H.-Y.H.)
- ² Digital Transformation, Institute for Information Industry, 11F, No. 106, Section 2, Heping E. Road, Taipei 106, Taiwan
- ³ Center for General Education, China Medical University, No. 100, Section 1, Jingmao Road, Beitun District, Taichung 406040, Taiwan; cwyang@mail.cmu.edu.tw
- ⁴ Master Program for Digital Health Innovation, College of Humanities and Sciences, China Medical University, No. 100, Section 1, Jingmao Road, Beitun District, Taichung 406040, Taiwan
- ⁵ Department of Computer Science and Engineering, National Chung Hsing University, No. 145, Xingda Road, South District, Taichung 40227, Taiwan; jasonlin@nchu.edu.tw
- * Correspondence: ycc@nttu.edu.tw

Abstract: Due to the climate crisis, energy-saving issues and carbon reduction have become the top priority for all countries. Owing to the increasing popularity of advanced metering infrastructure and smart meters, the cost of acquiring data on residential electricity consumption has substantially dropped. This change promotes the analysis of residential electricity consumption, which features both small and complicated consumption behaviors, using machine learning to become an important research topic among various energy saving and carbon reduction measures. The main subtopic of this subject is the identification of abnormal electricity consumption behaviors. At present, anomaly detection is typically realized using models based on low-level features directly collected by sensors and electricity meters. However, due to the significant number of dimensions and a large amount of redundant information in these low-level features, the training efficiency of the model is often low. To overcome this, this study adopts an autoencoder, which is a deep learning technology, to extract the high-level electricity consumption information of residential users to improve the anomaly detection performance of the model. Subsequently, this study trains one-class SVM models for anomaly detection by using the high-level features of five actual residential users to verify the benefits of high-level features.

Keywords: energy saving; carbon reduction; advanced metering infrastructure; low-voltage users; anomaly detection; autoencoder

1. Introduction

Owing to the depletion of fossil energy and increasingly serious global warming problems, the effective decrease in fossil energy consumption and energy and carbon reduction has become a common concern for governments and enterprises around the world. According to the data from the Bureau of Energy, Ministry of Economic Affairs of Taiwan, sectors that consumed the most energy in Taiwan were the industrial (55.9%), service (17.7%), and residential sectors (17.6%). The data clearly indicated that the electricity consumption of the residential sector was the third highest, only slightly lower than that of the service sector. Therefore, if the electricity consumption of the residential sector can be effectively reduced, considerable energy-saving benefits will be achieved. However, unlike the industrial and service sectors that comprise medium and large users, the residential



Citation: Tsai, C.-W.; Chiang, K.-C.; Hsieh, H.-Y.; Yang, C.-W.; Lin, J.; Chang, Y.-C. Feature Extraction of Anomaly Electricity Usage Behavior in Residence Using Autoencoder. *Electronics* 2022, *11*, 1450. https:// doi.org/10.3390/electronics11091450

Academic Editor: Floriano De Rango

Received: 2 April 2022 Accepted: 28 April 2022 Published: 30 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). sector comprises a large number of small users (approximately 13.2 million non-business users and 1.03 million business users of low-voltage meters). In addition, the electricity consumption behaviors of different users vary significantly, complicating the development of a universal energy-saving strategy for residential users. Fortunately, in recent years, information and communication technology has developed rapidly; mobile devices, mobile networks, and Internet of Things (IoT) devices have gained a significant amount of popularity. Furthermore, power companies have actively promoted the advanced metering infrastructure (AMI) and smart meters to replace the existing mechanical meters for understanding the power-load behavior of low-voltage users quickly and efficiently.

The cost of collecting the electricity consumption data of low-voltage users decreases every year, and related energy management systems and devices are gradually implemented. Despite this, residential users lack the motivation to apply energy saving and carbon reduction measures and introduce home energy management systems (HEMS) due to the low electric valance in Taiwan. Therefore, some research [1–13] has begun using machine learning techniques to collect and analyze the electricity consumption data of residential users and establish artificial intelligence (AI) models to provide appropriate and tailored energy-saving suggestions. Among them, if a mechanism can identify the abnormal electricity consumption behavior of residential users and propose appropriate energy management or saving suggestions, it will be particularly effective in improving user motivation in terms of energy-saving measures. Therefore, various anomaly detection techniques for the energy consumption of residential users and buildings have been proposed and discussed.

At present, the studies predominantly use the low-level feature data (i.e., the row data without being extracted), including electricity consumption data and associated features (temperature and humidity, summer/non-summer months, working/non-working days, etc.) to train machine learning models. However, the significant number of dimensions and a large amount of redundant information of these low-level features can compromise the training performance of subsequent anomaly detection algorithms. Although techniques such as principal component analysis (PCA) and feature selection can be adopted to improve this issue, how to provide a more efficient solution is still an important research topic. Therefore, this study further discusses this topic. In summary, this study wants to improve a method to extract the essences of the low-level features by using a deep neural network-autoencoder. That is, this study wants to use the autoencoder to extract the high-level features (i.e., code in the autoencoder) from the low-level features of the electricity consumption data of residential users. As the high-level features can decode to the original low-level features and the dimensions of the high-level features are less than those of the low-level features, the high-level features are more representative of the power consumption behaviors of users. That is, the high-level features are the essence of the user's power consumption data. This study uses the actual power consumption data of the five resident users to execute the experiments to verify whether using the high-level features to train the anomaly detection algorithm can benefit performance over using the low-level features. We use an anomaly detection algorithm, one-class SVM, to train the two anomaly detection models with the high-level and low-level features, respectively, and then analyze the performances of the two models to verify the feasibility of the proposed high-level feature extraction method.

The remainder of this paper is organized as follows. Section 2 reviews the related literature and technologies, Section 3 describes the research methods and processes, and Section 4 reports the implementation of the function and result comparison. Section 5 provides a conclusion and recommendations for future research topics and directions.

2. Background

This chapter first reviews the relevant literature on anomaly detection and then briefly explains the machine learning techniques used in this study, particularly the autoencoder and one-class SVM.

2.1. Anomaly Detection

As early as the 19th century, the statistical community had already started detecting anomalies in data. Anomalies are also referred to as outliers, biases, inconsistencies, and exceptions [14]. Anomalies typically include (1) point, (2) contextual, and (3) collective anomalies. The detection of point anomalies is the most simple and common anomaly detection method and strategy. However, rather than a pattern, point anomalies often represent a noise and consequently possess a low practical value. Alternatively, contextual anomalies are typically analyzed in a specific time sequence and spatial data to determine abnormal behaviors in the specific context, whereas collective anomalies often analyze group data comprising multiple pieces and evaluate whether the resultant model is anomalous. The occurrence of contextual anomalies depends on the availability of contextual attributes in the data. Therefore, when point anomaly detection is supplemented with contextual anomaly detection or part of the group data are categorized as contextual attributes, both point and collective anomalies are considered equivalent to contextual anomalies. Consequently, during anomaly detection, most studies convert anomaly events to contextual anomalies for analysis and processing. Anomaly detection strategies can be divided, according to the inclusion of labels in the analysis datasets, into three types: (1) supervised, (2) unsupervised, and (3) semi-supervised.

The primary techniques adopted by the existing literature to detect abnormal power consumption behaviors include [15] (1) anomaly detection models based on regression models, (2) anomaly detection models based on classifiers, and (3) others. This section divides anomaly detection techniques for electricity consumption according to their type and provides a brief review and description.

Anomaly detection models based on regression models first train the regression model using historical power-related data and then use the model to predict future consumption. An anomaly is detected upon a large deviation between the predicted and actual values (for example, the actual value is greater than the predicted threshold). Zhang et al. [16] developed an abnormal electricity load detection model based on a linear regression model and used its predications as the baseline. Power consumption data were considered abnormal when either significantly lower or higher than the threshold. Although the study provided a load anomaly detection solution that incorporated environmental factors, it could not accurately identify anomalies for residential users owing to their sensitivity to temperature. In addition, as the model was only trained with environmental factors, it might be inapplicable in an environment with a constant annual temperature. Alternatively, Zhou et al. [17] proposed an anomaly detection model based on a hybrid prediction model. The hybrid model integrated the ARIMA model with the ANN model, compensating the prediction error of the former in nonlinear regression and providing the advantages of both linear and nonlinear models. Although this approach improved the prediction accuracy, the anomaly detection strategy used was excessively simple and required further improvement. To eliminate detection errors caused by simple detection methods, Luo et al. [18] proposed an anomaly detection model based on dynamic regression. Instead of a fixed threshold, the model could calculate a dynamic, adaptive threshold for the difference between the predicted and actual loads during anomaly detection. The proposed dynamic-detection rule could improve the accuracy of anomaly detection. However, because the study used the results of the prediction model as the only reference for anomaly detection, an independent detection mechanism was lacking for anomaly detection, risking a decrease in anomaly detection accuracy when the prediction value was inaccurate. Fenza et al. developed a driftaware methodology for detecting anomalies in smart grids [19]. Historical data were used to train the long short-term memory (LSTM) and then to determine the anomaly detection thresholds from the prediction error trends obtained by the LSTM over time. As the study aimed to explore the abnormal load profile of users, the basis of anomaly detection was the error trend rather than the error between the predication and actual result for a specific time. Inayah et al. [20] used SARIMA and ANN models to predicate power consumption of the college buildings, and they adopted the difference between the actual and prediction

values to identify the anomaly events. Then, the results of the experiment proved that the ANN model has a better performance than the SARIMA model. Additionally, it is noteworthy that this kind of anomaly detection technology can also be used to protect the cybersecurity issue. For example, Zhang et al. [21] proposed a robustness assessment framework for wind power, and they evaluated the performances of the six forecasting models in terms of protection against the false data injection attack.

Anomaly detection models based on classifiers can be further divided into supervised and unsupervised/semi-supervised models according to the type of classifier. Jokar et al. developed an anomaly detection model for power theft based on supervised learning [22]. During the training process, the k-means cluster analysis algorithm and silhouette coefficient determined the number of patterns in the dataset, and an SVM-based classifier learned the normal and abnormal patterns. Pinceti et al. [23] conducted a model comparison study, during which different supervised learning models detected abnormal load redistribution events. After comparing kNN, SVM, and RNN models, the study suggested that the performance of the kNN model was superior. Fang et al. [24] adopted the extreme learning machines and the ensemble learning strategy to design a supervised learning anomaly detection system for various users (i.e., the low-voltage non-resident, the lowvoltage resident, the high-voltage resident, and the photovoltaic user). Wang et al. [25] proposed a semi-supervised learning anomaly detection model, sample efficient home power anomaly detection (SEPAD), in which the k-means and z-score function [26] were used to point out the suspicious data, and a semi-SVM based pattern matching algorithm was proposed to identify anomaly power consumption events. Hosseini et al. [27] focused on the appliance-level anomaly detection and trained the classification modes by using the operation patterns for the refrigerators depending on the semi-supervised learning strategy. Fan et al. [28] proposed a building electricity anomaly detection model based on unsupervised classification to reduce the training cost lower than that of supervised learning-based models. The study first determined the primary load frequency of users using spectral density analysis and features affecting the electricity consumption behavior using a decision tree, and then calculated the anomaly score of each event using the autoencoder, which is an unsupervised learning model, and ensemble learning. An event was defined as an anomaly if its anomaly score was higher than the preset threshold. Pereora et al. [29] developed an autoencoder-based unsupervised anomaly detection model for detecting anomalies in solar power generation. They also applied a variational self-attention mechanism to improve the performance of the autoencoder. Although anomaly detection techniques based on unsupervised learning do not require additional training to identify abnormal data, and therefore have a low training cost, evaluating their detection results is difficult due to the lack of reference labels [30]. Additional analysis (such as normal distribution analysis, data visualization analysis, and consulting domain experts) is often required to verify that the specific event is an anomaly.

Others include Janetzko et al. [31], who used the visual analysis to identify the anomaly power consumption events. The study [32] adopted the Hilbert-Huang transform and instantaneous frequency analysis to analyze the hidden anomaly events in commercial buildings. Cabrera et al. [33] adopted an anomaly detection method based on rule-based learning to analyze the waste of electricity in school buildings. They reduced the number of features using data mining methods and introduced various rules to identify wasteful behaviors. Li [34] uses statistical methods and clustering algorithms to identify the anomaly power consumption events in the short-term and long-term time scale data, respectively.

2.2. Autoencoder

An autoencoder [35,36] is an unsupervised learning algorithm in deep learning. The model is trained by defining the data (X) and output data (Y). According to the neural network architecture, an autoencoder comprises an encoder and decoder, which have neural networks with the same number of neurons. The encoder converts the input data into high-level features (Z) through the hidden layer, and the decoder reconstructs these

high-level features into input data through the hidden layer. The autoencoder aims to restore the high-level features of the input data as much as possible using the decoder. Its loss function often uses mean squared error (MSE) or cross-entropy losses. Two common autoencoder structures exist: undercomplete autoencoders whose number of neurons in the hidden layer is smaller than or equal to that in the decoder, and overcomplete autoencoders whose number of neurons in the hidden layer is larger than or equal to that in the decoder. Basic autoencoder structures comprise three fully connected layers: an input, hidden, and output layer. Both the number of hidden layers and its number of neurons can be adjusted to improve the model performance.

2.3. One-Class SVM

One-class SVM is an unsupervised algorithm [37–39]. As the name suggests, it classifies incoming training data into one category. A decision boundary is first learned using the characteristics of these normal samples, which are then used to determine the similarity between the new and training data. Abnormal data are identified when they exceed the boundary. If the kernel function adopts the Gaussian Redial Basis (RBF), features of the training data are first projected to high dimensions and then projected back to the original data dimension once the largest segmentation platform, the hyperplane, is determined in the high dimension. The one-class SVM algorithm is similar to that of two-class SVM. The only difference is that the former searches for the hyperplane that contains all the normal training data instead of the hyperplane that splits training samples into two categories.

3. Research Methods

This chapter describes the research methods and processes (Figure 1). Data from a full year of electricity consumption of lower-voltage residential users were analyzed to assess whether the proposed method could effectively detect abnormal electricity consumption behaviors. The details of the research methods and process are listed below.



Figure 1. Research structure and process.

3.1. Data Preprocessing

Although the sources of the electricity consumption data of low-voltage users were predominantly smart meters and home energy management systems, the collected data could still have noises, outliers, or missing values owing to the noise and short-term failure of sensors during communication and data transmission. These noisy and abnormal data could affect the subsequent model training and performance and must therefore be filtered and processed prior to model training and analysis. As assigning missing value and restoring abnormal data (e.g., using the regression model) could compromise the accuracy of model establishment, abnormal and missing data were directly deleted in this study, that is, the electricity load data of a day was deleted upon the presence of any missing or abnormal value. Furthermore, because the international electric power industry routinely used 15 min as the sampling frequency of user electricity load data, this study resampled the original load data to 1 record/15 min using data averaging. During data preprocessing and after the removal of missing and abnormal values, to accelerate model training and increase model accuracy, a min–max normalization was performed to convert the data to a range of [0:1], the formula of which is:

$$s_{norm}^{i} = rac{s_{original}^{i} - s_{min}}{s_{max} - s_{min}}$$

where $s_{original}^{i}$ is the *i*-th original sample data, s_{min} and s_{max} the minimum and maximum values of the original sample, respectively, and s_{norm}^{i} the normalized value of the *i*-th sample. Finally, the normalized data were then divided into a training and test dataset at a ratio of 80:20.

3.2. High-Level Feature Extraction

To facilitate the identification of abnormal electricity consumption behaviors, the study used the autoencoder for feature extraction. The autoencoder then encoded and compress the features of the 96 daily electricity load records to obtain low-dimensional electricity load features. This study adopted the multilayer undercomplete autoencoder as the primary high-level feature extraction model, which compressed and extracted the original low-level features (i.e., 96-dimensional features) into two-dimensional high-level features. Figures 2 and 3 present the model architecture, in which the intention of the first and the second dimensions of input and out shapes are the batch size and input size, and the term "None" in the first dimension means the batch size depends on how many samples we give for training.

input_3: InputLayer		input:	[(None, 96)]
		output:	[(None, 96)]
	model: Functional	¥	A
		mput:	(None, 96)
		output:	(None, 2)
	model 1: Functional	input:	(None, 2)
		output:	(None, 96)

Figure 2. Structure of the autoencoder model.



Figure 3. Structures of the encoder (left) and decoder (right).

Here, this study analyzes the time complexity of this network model to briefly obtain a time complexity formula. We know that the time complexity of matrix multiplication for $M_{ij} \times M_{jk}$ is $O(i \times j \times k)$. In the forward propagation, from the *i*-th layer to the (i + 1)-th layer, a matrix multiplication and an activation function must be computed. The time complexity of matrix multiplication is $O(n_{neuron}^i \times n_{neuron}^{i+1} \times n_{sample})$, where n_{neuron}^i (n_{neuron}^{i+1}) denotes the number of neurons in the *i*-th ((*i*+1)-th) layer, and n_{sample} is the number of samples used to train the network. Due to the element-wise operation, the time complexity of the activation is $O(n_{neuron}^{i+1} \times n_{sample})$. Therefore, the total time complexity is $O(n_{neuron}^{i} \times n_{neuron}^{i+1} \times n_{sample} + n_{neuron}^{i+1} \times n_{sample}) \approx O(n_{neuron}^{i} \times n_{neuron}^{i+1} \times n_{sample})$. For all networks, the time complexity is $O\left(\sum_{i=1}^{n_{layer}} (n_{neuron}^{i} \times n_{neuron}^{i+1}) \times n_{sample}\right)$, where n_{layer} denotes the number of layers in the network model. In the backward propagation, from (i + 1)-th layer to the *i*-th layer, we must compute the error signal matrix by an elementwise multiplication operation, use a matrix multiplication to compute the delta weights, and then adjust the weights by using element-wise operation. Therefore, the total time complexity is $O\left(2 \times n_{neuron}^{i+1} \times n_{sample} + n_{neuron}^{i} \times n_{neuron}^{i+1} \times n_{sample} + n_{neuron}^{i} \times n_{neuron}^{i+1}\right) \approx$ $O(n_{neuron}^{i} \times n_{neuron}^{i+1} \times n_{sample})$ which is the same as the time complexity in the forward propagation. Thus, the total time complexity of the network model in both propagations is $O\left(\sum_{i=1}^{n_{layer}} \left(n_{neuron}^{i} \times n_{neuron}^{i+1}\right) \times n_{sample} \times n_{epoch}\right)$, where n_{epoch} denotes the number of training iteration.

3.3. One-Class Classifier Training

Once high-level feature extraction is completed, a one-class classifier trained the model to detect abnormal electricity consumption behaviors. As a type of classifier, the one-class

classifier primarily uses single-class samples for model training, allowing the model to identify a new event and determine whether it belongs to the specific class of events. A positive result indicates that the new event belongs to the class whereas a negative result indicates that the new event does not belong to the class. The one-class classifier cannot provide further information on which class it belongs to. This study adopted the one-class SVM algorithm as the one-class classifier. As the proportion of the abnormal power consumption behaviors of general users was normally low, the study assumed that abnormal behaviors accounted for 2% of the overall load and used this assumption to train the model. In addition, to verify whether the proposed high-level feature extraction method could effectively escalate the anomaly detection efficiency of the one-class classifier, the study also used low-level features to train the anomaly detection model and compared its performance with that of the model trained using high-level features.

3.4. Performance Comparison

Owing to the lack of labels, anomalies detected by unsupervised strategies often had a low explanatory power. As insufficient evidence was available for proving whether the identified anomalies were true anomalies, domain experts should assist in the detection. Therefore, this study used data visualization to analyze the performances of models trained using high-level and low-level features. During data visualization, cluster analysis was adopted to determine the main load pattern of users among the electricity loads of normal consumption behaviors identified by the model. In addition, abnormal electricity consumption behaviors and characteristics identified by the two models were compared and their differences were analyzed to assess the pros and cons of the model trained using high-level features. The k-means++ algorithm was used during cluster analysis to determine the main electricity consumption characteristics of users, and a silhouette coefficient determined the appropriate number of clusters. Finally, the center point of each cluster was defined as the load characteristics of users, plotting its range using the Q1 and Q4 of the quartile to evaluate the usefulness of the abnormal electricity consumption detection model.

According to the steps and procedures, the study collected a full year of electricity consumption data of five congregate residences and performed data preprocess, high-level feature extraction, abnormal electricity consumption detection model training, main electricity consumption feature extraction, and performance analysis and evaluation, the results of which are described in the next chapter.

4. Results and Discussion

The data sources of this empirical research were five residences randomly selected from 200 residential users who had installed energy management systems. Data were collected at a frequency of one electricity consumption record per minute (for a total of 1440 records per day) between 1 January 2020 and 31 December 2020. The study then resampled the data to one record per 15 min (for a total of 96 records per day) using data averaging to match the main measurement unit adopted by power companies in Taiwan. Personal information was removed prior to data acquisition.

Subsequently, during model training, the epochs of the autoencoder was set to 10,000. MSE was chosen as the loss function owing to its sensitivity toward extreme values, and adaptive moment estimation (ADAS) was selected as the optimizer. Figure 4 presents the evaluation results of the trained model. According to the time complexity formula in Section 3.2, we can obtain the time complexity of each user's autoencoder model is approximately $O(2^{36})$.



Figure 4. Learning curves of model loss in the five autoencoder models.

The anomaly detection model was then trained by the one-class SVM using the highlevel and low-level features of five residences. The dates of all detected anomalies were labelled for subsequent visualization analysis. The study used the one-class SVM algorithm in the scikit-learn, version 1.0.2, during implementation. The kernel parameter of the model was set to linear, the gamma to auto, and all nu values to 0.02 (that is, anomalies accounted for 2% of the sample dataset). The remaining parameters were the default value. Next, the k-means of the cluster analysis, silhouette coefficient, and quartiles were calculated to plot the primary electricity load behavior of individual users. This was concluded by plotting the load profiles of the abnormal electricity consumption behaviors of users detected by high-level and low-level features as well as their main load behaviors to compare the performances of the two models.

Figures 5–9 show the anomaly power consumption events for the five tested residential users detected using high-level and low-level features. In the graphs, the green curve denotes the central value of the primary load behavior of users, the light green and grass green areas denote the Q1–Q4 range of the corresponding electricity consumption feature, and the red curve denotes the abnormal electricity consumption load. The plots on the left are anomalies detected using low-level features, whereas those on the right are anomalies detected using high-level features. The figures of the individual anomaly event are given in Appendix A.



Figure 5. Load profiles of the main and abnormal electricity consumption events of User 01.



Figure 6. Load profiles of the main and abnormal electricity consumption events of User 02.



Figure 7. Load profiles of the main and abnormal electricity consumption events of User 03.



Figure 8. Load profiles of the main and abnormal electricity consumption events of User 04.



Figure 9. Load profiles of the main and abnormal electricity consumption events of User 05.

Note that the electricity consumption behaviors of the five residents can be divided into two groups. An analysis of the primary difference between the two groups indicated that the major contributing factor was temperature, because high-load anomalies mostly occurred in summer months, and low-load anomalies in non-summer months. Here, the definition of summer and non-summer months is the same as that used by power companies in Taiwan, that is, summer months span from the start of June to the end of September, and the remaining months are non-summer months.

Visualization analysis clearly indicated that anomalies detected using high-level features were often extreme power consumption behaviors displayed as sharp rises or falls. These anomalies demonstrated significantly more distinctive features than those of the main load behaviors of the user, making it highly likely that they were real anomalies. In contrast, anomalies detected using low-level features were predominantly load behaviors lower than the main load behaviors of the user. However, such events were likely normal electricity consumption behaviors created when the user was not at home that day. In addition, the direct use of low-value features could not effectively identify obvious and rapidly changing load behaviors (i.e., anomalies detected using high-level features). It is noteworthy that there is a minor performance difference between using low-level and high-level features in the experiment of User 05. To classify the point, we analyze the load profiles of User 05 clearly and find that the load profile of User 05 has a property compared with the other users; that is, User 05 has a more regular power consumption behavior than others. Therefore, this study can infer that the anomaly detection model using high-level features has a better performance than using low-level features under the situation in which the user has irregular power consumption behavior. Due to the randomness of power consumption behaviors in most users, the performance of anomaly detection using high-level features in general is better than using low-level features.

5. Conclusions

This study trains an autoencoder model and uses the network model to extract the lowlevel features (96-dimension features) of the residential power consumption data to be the high-level features (two-dimensional features) for improving the performance of abnormal power consumption behavior detection models for residential users. The experiments are implemented to prove that the anomaly detection model using the high-level features is more performance than the model using the low-level features in terms of identifying the abnormal power consumption behaviors of residential users. If the proposed technology can be integrated into the home energy management system (HEMS), HEMS can provide the appropriate energy-saving suggestions at a suitable timing due to the more accurate rate of anomaly behavior detection.

However, because the study adopted an unsupervised learning method to establish the anomaly detection mechanism, its explanatory power of abnormal electricity consumption behaviors remains insufficient. In addition, this study only uses visualization analysis to evaluate the performance of the anomaly detection models, and only the basic autoencoder model is used and evaluated. Therefore, how to improve the explanatory power of the output results of the anomaly detection model, how to find the explore quantitative indicators and methods that can more accurately compare the performances between the models trained using high-level and low-level features, and further using the more advanced deep learning model to improve the performance of anomaly detection models will be our future work.

Author Contributions: Conceptualization, C.-W.T. and K.-C.C.; methodology, C.-W.T., J.L. and C.-W.Y.; investigation, H.-Y.H.; formal analysis, C.-W.T. and Y.-C.C.; writing—original draft, C.-W.T. and K.-C.C.; writing—review and editing, J.L.; and project administration, Y.-C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Ministry of Science and Technology, Taiwan, R.O.C. (Grant Nos. MOST 110-2221-E-143-003, MOST 110-2221-E-259-001, MOST 110-2221-E-143-004, MOST 110-2221-E-039-004, MOST 110-2222-E-005-006, MOST 110-2634-F-005-006, and MOST 111-2218-E-005-007-MBK), Bureau of Energy, Ministry of Economic Affairs, Taiwan (Grant No. 111-E0208), and China Medical University, Taiwan (Grant No. CMU110-S-21).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

To clearly represent each anomaly power consumption event indicated by the models using the low-level and high-level features, the individual load profiles of each user are shown in this appendix. Here, the green curve also denotes the central value of the primary load behavior of users, the light green and grass green areas indicate the Q1–Q4 range of the corresponding electricity consumption feature, and the red curve is the abnormal load profile.



Figure A1. Load profiles of each abnormal electricity consumption events of User 01 using the low-level features.



Figure A2. Load profiles of each abnormal electricity consumption events of User 01 using the high-level features.



Figure A3. Load profiles of each abnormal electricity consumption events of User 02 using the low-level features.



Figure A4. Load profiles of each abnormal electricity consumption events of User 02 using the high-level features.



Figure A5. Load profiles of each abnormal electricity consumption events of User 03 using the low-level features.



Figure A6. Load profiles of each abnormal electricity consumption events of User 03 using the high-level features.



Figure A7. Load profiles of each abnormal electricity consumption events of User 04 using the low-level features.



Figure A8. Load profiles of each abnormal electricity consumption events of User 04 using the high-level features.



Figure A9. Load profiles of each abnormal electricity consumption events of User 05 using the low-level features.



Figure A10. Load profiles of each abnormal electricity consumption events of User 05 using the high-level features.

References

- Koolen, D.; Sadat-Razavi, N.; Ketter, W. Machine Learning for Identifying Demand Patterns of Home Energy Management Systems with Dynamic Electricity Pricing. *Appl. Sci.* 2017, 7, 1160. [CrossRef]
- Seyedzadeh, S.; Rahimian, F.; Glesk, I.; Roper, M. Machine learning for estimation of building energy consumption and performance: A review. *Vis. Eng.* 2018, 6, 5. [CrossRef]
- Al Tarhuni, B.; Naji, A.; Brodrick, P.G.; Hallinan, K.P.; Brecha, R.J.; Yao, Z. Large scale residential energy efficiency prioritization enabled by machine learning. *Energy Effic.* 2019, 12, 2055–2078. [CrossRef]
- 4. Yu, L.; Xie, W.; Xie, D.; Zou, Y.; Zhang, D.; Sun, Z.; Jiang, T. Deep reinforcement learning for smart home energy management. *IEEE Internet Things J.* **2019**, *7*, 2751–2762. [CrossRef]
- 5. Wang, A.; Lam, J.C.K.; Song, S.; Li, V.O.K.; Guo, P. Can smart energy information interventions help householders save electricity? a svr machine learning approach. *Environ. Sci. Policy* **2020**, *112*, 381–393. [CrossRef]
- Machorro-Cano, I.; Alor-Hernández, G.; Paredes-Valverde, M.A.; Rodríguez-Mazahua, L.; Sánchez-Cervantes, J.L.; Olmedo-Aguirre, J.O. HEMS-IoT: A big data and machine learning-based smart home system for energy saving. *Energies* 2020, 13, 1097. [CrossRef]
- Kumar, M.; Zhang, X.; Liu, L.; Wang, Y.; Shi, W. Energy-efficient machine learning on the edges. In Proceedings of the 2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), New Orleans, LA, USA, 18–22 May 2020; pp. 912–921.
- 8. Naji, A.; Al Tarhuni, B.; Choi, J.; Alshatshati, S.; Ajena, S. Toward cost-effective residential energy reduction and community impacts: A data-based machine learning approach. *Energy AI* **2021**, *4*, 100068. [CrossRef]
- Aslam, M.S.; Ghazal, T.M.; Fatima, A.; Said, R.A.; Abbas, S.; Khan, M.A.; Siddiqui, S.Y.; Ahmad, M. Energy-Efficiency Model for Residential Buildings Using Supervised Machine Learning Algorithm. *Intell. Autom. Soft Comput.* 2021, 30, 881–888. [CrossRef]
- Revati, G.; Hozefa, J.; Shadab, S.; Sheikh, A.; Wagh, S.R.; Singh, N.M. Smart building energy management: Load profile prediction using machine learning. In Proceedings of the 2021 29th Mediterranean Conference on Control and Automation (MED), Bari, Italy, 22–25 June 2021; pp. 380–385.
- 11. Anastasiadou, M.; Vítor, S.; Miguel, S.D. Machine Learning Techniques Focusing on the Energy Performance of Buildings: A Dimensions and Methods Analysis. *Buildings* **2021**, *28*, 12. [CrossRef]
- 12. Zekić-Sušac, M.; Mitrović, S.; Has, A. Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities. *Int. J. Inf. Manag.* 2020, *58*, 102074. [CrossRef]
- 13. Shapi, M.K.M.; Ramli, N.A.; Awalin, L.J. Energy consumption prediction by using machine learning for smart building: Case study in Malaysia. *Dev. Built Environ.* 2020, *5*, 100037. [CrossRef]
- 14. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly detection: A survey. ACM Comput. Surv. 2009, 41, 1–58. [CrossRef]
- 15. Wang, X.; Ahn, S.-H. Real-time prediction and anomaly detection of electrical load in a residential community. *Appl. Energy* **2019**, 259, 114145. [CrossRef]
- 16. Zhang, Y.; Chen, W.W.; Black, J. Anomaly Detection in Premise Energy Consumption Data. In Proceedings of the 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 24–28 July 2011.
- 17. Chou, J.S.; Telaga, A.S. Real-time detection of anomalous power consumption. *Renew. Sustain. Energy Rev.* 2014, 33, 400–411. [CrossRef]
- Luo, J.; Hong, T.; Yue, M. Real-time anomaly detection for very short-term load forecasting. J. Mod. Power Syst. Clean Energy 2018, 6, 235–243. [CrossRef]
- 19. Fenza, G.; Gallo, M.; Loia, V. Drift-aware methodology for anomaly detection in smart grid. *IEEE Access* **2019**, *7*, 9645–9657. [CrossRef]
- 20. Inayah, N.; Wijaya, M.Y.; Fitriyati, N. Energy Saving Potential Prediction and Anomaly Detection in College Buildings; In Proceedings of the International Conference on Mathematics and Islam (ICMIs 2018), Mataram, Indonesia, 3–5 August 2020.
- 21. Zhang, Y.; Lin, F.; Wang, K. Robustness of short-term wind power forecasting against false data injection attacks. *Energies* **2020**, 13, 3780. [CrossRef]
- Jokar, P.; Arianpoo, N.; Leung, V.C.M. Electricity theft detection in AMI using customers' consumption patterns. *IEEE Trans.* Smart Grid 2016, 7, 216–226. [CrossRef]
- Pinceti, A.; Sankar, L.; Kosut, O. Load Redistribution Attack Detection Using Machine Learning: A Data-driven Approach. In Proceedings of the 2018 IEEE Power & Energy Society General Meeting (PESGM), Portland, OR, USA, 5–10 August 2018.
- Fang, Z.; Cheng, Q.; Mou, L.; Qin, H.; Zhou, H.; Cao, J. Abnormal electricity consumption detection based on ensemble learning. In Proceedings of the 2019 9th International Conference on Information Science and Technology (ICIST), Inner Mongolia, China, 2–5 August 2019; pp. 175–182.
- 25. Wang, X.; Yang, I.; Ahn, S.-H. Sample Efficient Home Power Anomaly Detection in Real Time Using Semi-Supervised Learning. *IEEE Access* 2019, 7, 139712–139725. [CrossRef]
- Welikala, S.; Dinesh, C.; Ekanayake, M.P.B.; Godaliyadda, R.I.; Ekanayake, J. Incorporating Appliance Usage Patterns for Non-Intrusive Load Monitoring and Load Forecasting. *IEEE Trans. Smart Grid* 2017, 10, 448–461. [CrossRef]
- 27. Hosseini, S.S.; Agbossou, K.; Kelouwani, S.; Cardenas, A.; Henao, N. A Practical Approach to Residential Appliances on-Line Anomaly Detection: A Case Study of Standard and Smart Refrigerators. *IEEE Access* **2020**, *8*, 57905–57922. [CrossRef]

- Fan, C.; Xiao, F.; Zhao, Y.; Wang, J.Y. Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. *Appl. Energy* 2018, 211, 1123–1135. [CrossRef]
- 29. Pereira, J.; Margarida, S. Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Autoencoders with Attention. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018.
- Weinberg, D. Our Machines Now Have Knowledge We Will Never Understand. Available online: https://www.wired.com/ story/our-machines-now-have-knowledge-well-never-understand/ (accessed on 25 March 2022).
- 31. Janetzko, H.; Stoffel, F.; Mittelstädt, S.; Keim, D.A. Anomaly detection for visual analytics of power consumption data. *Comput. Graph.* 2014, *38*, 27–37. [CrossRef]
- 32. Naganathan, H.; Chong, W.K.; Huang, Z.; Cheng, Y. A Non-stationary Analysis Using Ensemble Empirical Mode Decomposition to Detect Anomalies in Building Energy Consumption. *Procedia Eng.* **2016**, *145*, 1059–1065. [CrossRef]
- Cabrera, D.F.M.; Zareipour, H. Data association mining for identifying lighting energy waste patterns in educational institutes. Energy Build. 2013, 62, 210–216. [CrossRef]
- Li, Z.; Zareipour, H. Abnormal Energy Consumption Analysis Based on Big Data Mining Technology. In Proceedings of the 2020 Asia Energy and Electrical Engineering Symposium (AEEES), Chengdu, China, 29–31 May 2020.
- Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Chap. Learning Internal Representations by Error Propagation; MIT Press: Cambridge, MA, USA, 1986; Volume 1, pp. 318–362. Available online: http://dl.acm.org/citation.cfm?id=104279.104293 (accessed on 25 March 2022).
- 36. Baldi, P. Autoencoders, unsupervised learning, and deep architectures. In Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, JMLR Workshop and Conference Proceedings, Edinburgh, Scotland, 26 June–1 July 2012.
- 37. Schölkopf, B.; Williamson, R.; Smola, A.; Shawe-Taylor, J.; Platt, J. Support Vector Method for Novelty Detection. In Proceedings of the 12th International Conference on Neural Information Processing Systems, Denver, CO, USA, 29 November–4 December 1999.
- Li, K.L.; Huang, H.K.; Tian, S.F.; Xu, W. Improving one-class SVM for anomaly detection. In Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No. 03EX693), Xi'an, China, 2–5 November 2003; Volume 5, pp. 3077–3081.
- Wang, Y.; Johnny, W.; Andrew, M. Anomaly intrusion detection using one class SVM. In Proceedings of the Fifth Annual IEEE SMC Information Assurance Workshop, IEEE, New York, NY, USA, 10–11 June 2004.