

Article **Machine Learning for Long Cycle Maintenance Prediction of Wind Turbine**

Chia-Hung Yeh 1,2,*, Min-Hui Lin ¹ , Chien-Hung Lin ³ [,](https://orcid.org/0000-0003-3833-2572) Cheng-En Yu ¹ and Mei-Juan Chen ⁴

- ¹ Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan; d063010005@student.nsysu.edu.tw (M.-H.L.); alleneatpie@gmail.com (C.-E.Y.)
- ² Department of Electrical Engineering, National Taiwan Normal University, Taipei 10610, Taiwan
- ³ Taiwan Power Company, Kaohsiung 80446, Taiwan; u621363@taipower.com.tw
- Department of Electrical Engineering, National Dong Hwa University, Hualien 97401, Taiwan; cmj@gms.ndhu.edu.tw
- ***** Correspondence: yeh@mail.ee.nsysu.edu.tw

Received: 27 February 2019; Accepted: 3 April 2019; Published: 8 April 2019

Abstract: Within Internet of Things (IoT) sensors, the challenge is how to dig out the potentially valuable information from the collected data to support decision making. This paper proposes a method based on machine learning to predict long cycle maintenance time of wind turbines for efficient management in the power company. Long cycle maintenance time prediction makes the power company operate wind turbines as cost-effectively as possible to maximize the profit. Sensor data including operation data, maintenance time data, and event codes are collected from 31 wind turbines in two wind farms. Data aggregation is performed to filter out some errors and get significant information from the data. Then, the hybrid network is built to train the predictive model based on the convolutional neural network (CNN) and support vector machine (SVM). The experimental results show that the prediction of the proposed method reaches high accuracy, which helps drive up the efficiency of wind turbine maintenance.

Keywords: Internet of Things (IoT); sensors; deep learning; data mining; long cycle maintenance; convolutional neural network; wind turbine; conditional monitoring

1. Introduction

Internet of Things (IoT) sensors can communicate and interact over the Internet, and be monitored and controlled remotely [\[1,](#page-8-0)[2\]](#page-8-1). Data collected from sensors are to be processed to generate valuable information for management. Predictive maintenance techniques are designed to help evaluate the equipment condition to determine the maintenance schedule [\[3,](#page-8-2)[4\]](#page-8-3). We aim to predict when the wind turbines would fail and how long they would last before their failure.

Wind power, different from burning fossil fuels, is plentiful, clean, and renewable. People have made use of wind power for hundreds of years. Wind turbines are usually operated in tough environments and need to endure high wind speed and extreme temperatures. The maintenance cost of wind turbine components is usually higher than the procurement cost. A wind farm would encounter many difficulties such as drivetrain failures, spalled bearings, and fractured gears, so its operation and maintenance have become an important research field.

The research fields regarding maintenance cost reduction include periodic inspections, corrective maintenance, and condition monitoring. [\[5\]](#page-8-4). Condition monitoring is normally used on rotating machinery, so it is very suitable for monitoring wind turbines [\[6\]](#page-8-5). Since condition monitoring is often used to maintain the wind turbines in the wind power industry, data collected from sensors are analyzed to predict the equipment condition. When an unusual event is detected, sensors record the

event using event codes, which helps the operator to quickly understand the wind turbines' status at that time. Because there are erroneous and irrelevant data in the large amount of data collected, efficient data pre-processing and analyses techniques are used to extract valuable information for maintenance. However, most of the methods are rule-based, and it is difficult for these methods to reach high accuracy of maintenance cycle prediction. This paper tries to analyze the collected data to improve the efficiency of wind turbine maintenance. Among all maintenance categories, the most time-consuming task is the long cycle maintenance, which requires a great deal of human and material resources. Therefore, the efficient and effective prediction of the occurrence of long cycle maintenance is very important in the management of wind turbines.

In Changhua Coastal Industrial Park, there are two wind farms where all 31 wind turbines are installed with the V80 module [\[7\]](#page-8-6). The sensors include an anemometer, electricity meter, gearbox thermometer, and high-speed bearings. A large amount of data is collected by the sensors when monitoring the wind turbines, and data mining techniques [\[8\]](#page-8-7) are then employed to process the massive amount of data and dig out useful information for maintenance. In our method, all maintenance periods are firstly divided into four categories and the relations of the maintenance periods of these categories are analyzed. The Apriori algorithm [\[9\]](#page-8-8) is used to find the key event codes, which occur before the maintenance periods. Also, linear regression is applied to filter some errors from the operation data. Six features: wind speed, the power output of wind turbines, the oil temperature of the wind turbine gearbox, the temperature of the high-speed bearings, the number of consecutive short under-maintenance periods, and the frequency of the event codes, are extracted from the results above. Then a hybrid network consisting of a convolutional neural network (CNN) and support vector machine (SVM) is constructed to predict the long cycle maintenance time. Experimental results show that the proposed method achieves high accuracy in predicting the period of long cycle maintenance.

The rest of this paper is organized as follows. In Section [2,](#page-1-0) we review the background knowledge. Section [3](#page-2-0) focuses on the proposed method of long cycle maintenance. In Section [4,](#page-7-0) the experimental results demonstrate the efficiency of the proposed method. Concluding remarks and future work are given in Section [5.](#page-8-9)

2. Background Review

2.1. Condition Monitoring of Wind Turbines

Condition monitoring is the process of monitoring the condition of machinery to identify a significant change which is implicative of an operational problem in order to prevent its failure and the consequences associated with it. Condition monitoring uses sensors and signal-processing equipment to continuously monitor wind turbine components regarding vibration analysis [\[10\]](#page-8-10), oil analysis [\[11\]](#page-8-11), performance monitoring [\[12\]](#page-8-12), strain measurement [\[13\]](#page-9-0), acoustic emission [\[14\]](#page-9-1), and ultrasonic testing techniques [\[15\]](#page-9-2). Through data mining techniques, the data collected during the operation of major components can be used to reduce major component failure. Maintenance schedules can be planned more efficiently, and the costs of maintenance are reduced. Condition monitoring is commonly used in most offshore wind farms because of their high costs of maintenance.

2.2. Convolutional Neural Network (CNN)

Deep learning is a technique required for implementing machine learning. CNNs are one of the most popular architectures in deep learning. CNNs are hierarchical neural networks and have been successfully used for data analyses, speech recognition, image classification, pattern recognition, and computer vision. A typical CNN structure contains subsampling layers between each convolutional layer and fully connected layers. In order to implement the CNN, there are several parameters to be set, such as input size, kernel (filter) dimensions, number of neurons in each layer, and the topology of the CNN. In this paper, CNNs are trained to learn the features that are related to wind turbine maintenance. A combination of the CNN and SVM (Support Vector Machine), where the output of the CNN is replaced by an SVM, could achieve great performance for classification.

2.3. Simple Linear Regression

In statistics, simple linear regression is a method which allows us to summarize and study the relationships between two continuous variables. Consider the model function for the slope β and y-intercept α in the simple linear regression:

$$
y = \alpha + \beta x + \mu_i,\tag{1}
$$

where μ_i , named residuals, are the distance between the points of the data and the regression line. The advantage of simple linear regression is that it is efficient and intuitive to find the nature of the relationship between two variables. In this paper, the simple linear regression is used to model the relationship between the operation data.

2.4. Apriori Algorithm

Data mining is the computational process of finding patterns in large data sets. Some techniques have been proposed in previous research regarding data mining on wind turbines, including prediction models for wind turbine states [\[16\]](#page-9-3), fault diagnosis [\[17](#page-9-4)[,18\]](#page-9-5), and power curve monitoring [\[19\]](#page-9-6). The Apriori algorithm is one of the most popular algorithms in discovering association rules. The Apriori algorithm was proposed by Srikant et al. in 1994 and is designed to work on transactional databases. The goal of this algorithm is to extract information from a data set and transform it into a comprehensible structure for further use. The simple description of the Apriori algorithm is as follows. Assume that there are two itemsets *X* and *Y* in the database. If there is an association rule between *X* and *Y*, the rule is denoted as $X \Rightarrow Y$ called association rules.

Support: a representation of how frequently the itemset appears in the database, "*Support*" is defined as:

$$
Support{X \Rightarrow Y} = P(X \cap Y). \tag{2}
$$

Confidence: the transaction that contains *X* would also contain *Y*, "*Con f idence*" is defined as:

$$
Confidence\{X \Rightarrow Y\} = P(Y|X) = \frac{P(X \cap Y)}{P(X)}.
$$
\n(3)

Lift: the ratio of the confidence of the association rule to the probability of γ , "Lift" is defined as:

$$
Lift\{X \Rightarrow Y\} = \frac{P(Y|X)}{P(Y)} = \frac{P(X \cap Y)}{P(X)P(Y)}.
$$
\n(4)

If $Lift{X \Rightarrow Y} > 1$, *X* and *Y* are dependent, which makes the rule potentially useful in predicting. The Apriori algorithm generates candidate item sets of the length *k* from item sets of the length *k* − 1. Then it reduces the candidates which have infrequent sub-patterns. After that, it scans the database to determine frequent item sets among the candidates. Apriori utilizes a "bottom-up" approach to extend frequent subsets of one item at a time (a step known as candidate generation), and groups of candidates are tested against the data. The algorithm ends when no further successful extensions are found. In this paper, we employ the Apriori algorithm to find the key event codes of wind turbines.

3. Proposed Method

In this section, maintenance time data, event codes, and operation data are introduced first, followed by the explanations on the proposed methods: Data aggregation and prediction model construction.

3.1. Data Aggregation

Figure [1](#page-3-1) illustrates the concept of the proposed method and Table 1 shows the specifications of the V80 module wind turbines in Changhua Coastal Industrial Park. Firstly, all the data are aggregated to filter errors and extract significant features. Data aggregation is a process in which information is gathered and expressed in a summary form for the purpose of the follow-up statistical analysis. The results of data aggregation will be applied to deep learning models to predict the long cycle maintenance time. cycle maintenance time.

Figure 1. The architecture of the long cycle maintenance time prediction. **Figure 1.** The architecture of the long cycle maintenance time prediction.

	Rated power	2000 kW				
Operating Data	Cut-in wind speed	4.0 m/s				
	Rated wind speed	15 m/s				
Rotor	Cut-out wind speed	25 m/s				
	Rotor diameter	80 _m				
	Operational interval	$9-19$ rpm				
	Air brake	Full blade feathering with 3 pitch cylinders				
Gearbox	Two planetary stages and one helical stage					

Table 1. Specification of the V80 module wind turbine.

3.1.1. Maintenance Time and Event Codes

3.1.1. Maintenance Time and Event Codes status. The maintenance time data are collected for 941 consecutive days, and the frequency is updated every ten minutes. The re[co](#page-4-0)rded maintenance time data is shown in Figure 2, where the rows represent the time of the 10-min cycles and the columns represent the wind turbine number. If the recorded time data equal "NAN", it means that the wind turbine is under maintenance. To better describe the maintenance time data, attributes such as NAN time, time condition, and NAN period are defined. Figure 3 illustrates that the maint[ena](#page-4-1)nce data contains three attributes: NAN time, time condition, and NAN period. "NAN time" is the maintenance time of the wind turbine, shown as a green double-headed arrow. "Time condition" is the interval from the end of the previous NAN to the start of the current NAN, shown as a blue double-headed arrow. "NAN period" is the interval from the end of the previous NAN to the end of the current NAN, shown as an orange double-headed arrow. Figure 4 shows an example of these attributes being extracted through data aggregation. Starting from the left are the columns of wind turbine number, the previous NAN end time, the current NAN start time, the current NAN end time, and the interval of the NAN period. In Table [2,](#page-4-3) sensors would record the fault message represented as "event code" when the abnormal operation of a wind turbine occurs. The Vestas supervisory control and data acquisition (SCADA) system would collect and record the event code that occurs first in that 10-min cycle time. Maintenance time data archives the records of accumulated maintenance hours and maintenance

Machine Time	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9	No. 10	No. 11			
2015/10/06 11:40:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 11:50:00'	1381	1036	1121	912	807	484	763	1289	602	183 NaN		NAN time		
'2015/10/06 12:00:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 12:10:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 12:20:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 12:30:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			
'2015/10/06 12:40:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			
'2015/10/06 12:50:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			
'2015/10/06 13:00:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114		Time	
'2015/10/06 13:10:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			
'2015/10/06 13:20:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114		con	
'2015/10/06 13:30:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			
2015/10/06 13:40:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114		ndition	
'2015/10/06 13:50:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			MAN
'2015/10/06 14:00:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			
'2015/10/06 14:10:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			period
'2015/10/06 14:20:00'	1381	1036	1121	912	807	484	763	1289	602	183	1114			
2015/10/06 14:30:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 14:40:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 14:50:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 15:00:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN	NAN time		
'2015/10/06 15:10:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 15:20:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 15:30:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 15:40:00'	1381	1036	1121	912	807	484	763	1289	602		183 NaN			
'2015/10/06 15:50:00'	1381	1036	1121	912	807	484	763	1289	602	183	1115			
'2015/10/06 16:00:00'	1381	1036	1121	912	807	484	763	1289	602	183	1115			

Figure 2. Recorded maintenance time data. **Figure 2.** Recorded maintenance time data. **Figure 2.** Recorded maintenance time data.

Figure 3. An illustration of the relationships between NAN time, time condition, and NAN period.

\blacksquare 1 2015/02/17 21:10:00 2015/02/23 09:40:00 2015/02/23 16:40:00 1		
2015/02/23 16:40:00 2015/03/17 14:30:00 2015/03/21 17:00:00 3		
2015/03/21 17:00:00 2015/03/23 13:40:00 2015/03/26 17:30:00 4		
2015/03/26 17:30:00 2015/03/27 09:40:00 2015/03/27 11:00:00 1		
2015/03/27 11:10:00 2015/04/09 10:50:00 2015/04/09 11:50:00 1		

Figure 4. Example of data after the data aggregation step. **Figure 4.** Example of data after the data aggregation step.

Table 2. Event codes of failure messages in common use. **Table 2.** Event codes of failure messages in common use.

Event Codes	Failure Message
900	Pause pressed on the keyboard
276	Start auto-outyawing CCW
329	DC Overvolt: 880 V state 1
144	High wind speed: 25.1 m/s
315	ExEx low voltage L1: 282 V
340	OVP active UDC 0V state 1
335	ExtHighIRotorInv phase: 3
147	High gear temp
154	Max rotor RPM: 21.8 RPM
214	Low oil level
309	Pause over RCS 13

The NAN period is classified into four maintenance categories according to the length of its NAN maintenance hours. The NAN period classification rule is defined as follows. The length of the NAN time in the NAN period ranging from 1–10 is classified as the first class, 11–50 as the second class, 51–100 as the third class, and those above 100 as the fourth class. Figure 5 shows the relationship between the maintenance categories and the number of event codes. Apparently, the first class is the converse are this manufacture composite above the relationship of the fourth context supply and the relationsh
most frequently observed one, so it exhibits the highest number of event codes. between the maintenance categories and the number of event codes. Approximately, the first codes. time, and the fourth class is defined as the long cycle maintenance time because it has the longest most frequently observed one, so it exhibits the highest number of event codes.

Figure 5. The number of event codes with regard to maintenance categories. $(x$ -axis indicates four categories of NAN periods and their corresponding length of NAN times; *y*-axis is the number of event codes).

means that its maintenance cost is the highest. If the long cycle maintenance time can be predicted, the cost could be significantly reduced. We analyze the event codes of the NAN period that occurred predicted and ϵ control the control of the control of the coupled be significantly reduced. We are due to the control of the control of the control of the control of the NAN period on d the before the fourt[h c](#page-5-1)lass. Figure 6 is an example of the time graph of the target NAN period and the
NAN period of the fourth also NAN period of the fourth class. The fourth class has the longest maintenance hours among the four maintenance categories, which The fourth class has the longest maintenance hours among the four maintenance categories,

Figure 6. Time graph of the fourth between the target NAN period of the fourth between the NAN period of the NA period of the fourth class. period of the fourth class. period of the fourth class. **Figure 6.** Time graph of the relationship between the target NAN period occurring before the NAN **Figure 6.** Time graph of the relationship between the target NAN period occurring before the NAN

maintenance. We use the Apriori algorithm to find association rules of the event codes of the NAN period occurring before the fourth class. Because the operation performance of two wind farms is different, we analyze the maintenance data of two wind farms separately. From the analysis results, we found that two event codes, Code 154 and Code 309, happened along with the long cycle maintenance time. Based on knowledge in this domain and the experience of the Taiwan Power Company, another two event codes, Code 147 and Code 214, are also considered to be related to the long cycle maintenance two event codes, Code 117 and Code 211, are also considered to be related to the long cycle maintenance time. So, the frequency of Code 154, Code 309, Code 147, and Code 214 are chosen to be the features. for predicting the long cycle maintenance time. time. So, the frequency of Code 154, Code 309, Code 147, and Code 214 are chosen to be the features It is quite reasonable to assume that several event codes exist, which lead to long cycle

To find other features, we designed a rule to predict the long cycle maintenance time. If there are more than eight consecutive short NAN periods before the fourth class, and the time between the the more than eight consecutive short twist periods before the fourth class, and the time between the fourth class and the previous NAN period is shorter than the predefined threshold a large maintenance cycle is likely to occur. Figure $\frac{7}{5}$ shows that for the wind turbine 6, a total of twenty-six consecutive $\frac{1}{2}$ short NAN periods (15—first class, one—third class, one—second class, and nine-one class) occur. before the fourth class. In Table 3, time thresholds are set to be less than 14 days, 30 days, or 45 days and the prediction accuracy can reach 30%. The result is satisfying when we use the rule to predict the long cycle maintenance time only. It means that there is a high correlation between the number of consecutive short NAN periods and the long cycle maintenance time. Therefore, the number of $\frac{1}{2}$ consecutive short NAN periods is chosen to be the feature to predict long cycle maintenance time. \mathbf{r} and \mathbf{r} is chosen to be the feature to predict long cycle fourth class and the previous NAN period is shorter than the predefined threshold, a large maintenance
 $\frac{1}{2}$ cycle is likely to occur. Figure [7](#page-6-0) shows that for the wind turbine 6, a total of twenty-six consecutive $\frac{1}{100}$ short NAN periods (15—first class, one—third class, one—second class, and nine–one class) occur
https://www.andrail.org/www.andrail.org/www.andrail.org/www.andrail.org/www.andrail.org/www.andrail.org/www.an before the fourth class. In Table [3,](#page-6-1) time thresholds are set to be less than 14 days, 30 days, or 45 days, because the fourth class. In Table 3, time thresholds are set to be less than 14 days, 30 days, or 45 days, and the prediction accuracy can reach 30%. The result is satisfying when we use the rule to predict
 the long cycle maintenance time only. It means that there is a high correlation between the number of consecutive short NAN periods and the long cycle maintenance time. Therefore, the number of consecutive short NAN periods is chosen to be the feature to predict long cycle maintenance time.

Figure 7. Histogram of the number of occurrences with regard to maintenance categories. $(x$ - and y -axes are maintenance categories that happened sequentially, and the number of occurrences, respectively).

Time Condition	Less than 14 Days Less than 30 Days Less than 45 Days		
Rule	29.79%	31.91%	31.91%

Table 3. The accuracy of the long cycle maintenance time prediction with the rule.

3.1.2. Operation Data accuracy control and data is control operation data is collected for Ω

Operation data including wind speed, the power output of wind turbines, the oil temperature of the wind turbine gearbox, and the temperature of the high-speed bearings, are collected by the Vestas supervisory control and data acquisition (SCADA) system. The operation data is collected for 941 consecutive days, and the updating frequency is every ten minutes, which is the same time length as the available maintenance data. There are some errors in the wind speed and the power output of the wind turbine's data due to the bug in the data collection system. The relation of wind speed and output power could be easily modeled by regression analysis. According to the characteristics of linear regression, linear regression is a suitable model to fit this formula. To filter the errors in the wind speed and the power output of the wind turbines data, a residual threshold is set. Then, the after this step. We calculate the standard deviation from the predicted moment to the last long cycle maintenance. The standard deviation is chosen to be the feature of the long cycle maintenance time. data whose residual is greater than the threshold is filtered. Most of the error data could be filtered

combines the CNN and SVM. CNNS are quite good at learning invariant features, but not always in \mathcal{L} optimal for construction (most of the middle parameters are in the middle layers). On the middle layers are in the other middle layers are in the other middle layers of the other middle layers. On the other middle layers i *3.2. Prediction Model Construction*

To predict the long cycle maintenance time precisely, we proposed a hybrid network which combines the CNN and SVM. CNNs are quite good at learning invariant features, but not always optimal for classification (most of the trainable parameters are in the middle layers). On the other hand, SVM with a fixed kernel function, cannot learn complicated invariances but can produce good decision surfaces when applied to well-behaved feature vectors. It is interesting to investigate a hybrid system in which the convolutional net is trained to extract features that are relatively invariant to irrelevant variations of the input. Thus, an SVM with a simple Gaussian kernel can perform well at separating the categories in the learned feature space. SVM is added to the top layer, instead of the softmax layer.

The architecture of the hybrid network is shown in Figure 8, where two-layer CNNs are used to extract features, and the features are then used as the input to train an RBF-SVM (Radial Basis Function). Six features: wind speed, power output of the wind turbines, oil temperature for the wind turbine gearbox, temperature of the high-speed bearings, the number of consecutive short NAN periods, and the frequency of event codes (Code 154, Code 309, Code 147, and Code 214) from the data aggregation are used to train the hybrid network. In the following experiments, a two-layered 1-D CNN and the fully connected layer are used to extract features, and the features are then used as input to train the RBF-SVM. The model is trained by the features extracted from the data aggregation.

Figure 8. The architecture of the hybrid network. **Figure 8.** The architecture of the hybrid network.

4. Experimental Results 4. Experimental Results

Data aggregation and model construction are implemented in Python with Keras [20]. We used Data aggregation and model construction are implemented in Python with Keras [\[20\]](#page-9-7). We used two 1-D CNNs in all experiments with SVM in the output layer. The experiments conducted in the two 1-D CNNs in all experiments with SVM in the output layer. The experiments conducted in the proposed method used the dataset collected from sensors in 31 wind turbines for 941 consecutive proposed method used the dataset collected from sensors in 31 wind turbines for 941 consecutive days. This dataset is a collection of wind turbine operation data, maintenance time data, and event code data. There are 4,205,119 records in each of the operation data of the four categories: wind speed, the power $\frac{1}{2}$ of the wind turbines, the oil temperature for the wind turbine gearbox, and the temperature output of the wind turbines, the oil temperature for the wind turbine gearbox, and the temperature of the wind turbines of the temperature the high-speed bearings. The maintenance time data also contains 4,205,119 samples and the event code data contain 6251 samples. We used data aggregation to extract the input data of size 6×135 ,649 from six features, including the above four categories, the number of consecutive short NAN periods, and the frequency of event codes. The dataset is randomly split into two parts, 80% as training data and 20% as test data. Both the first 1×2 CNN and the second 1×3 CNN consisted of a rectified linear unit (ReLU) and a dropout rate of 0.3. The maximum pooling layer with a stride of two is placed after each CNN, so that the features are downscaled twice to reduce the computational cost. The size of the fully-connected layer is 1 \times 128. The output layer used RBF-SVM for classification. The learning rate $t \geq 0.001$ T second the similar attention and slightly higher than the training loss. This step can be step can ε is set at 0.001. To prevent the underestimation or overestimation, the validation loss is ensured to be similar and slightly higher than the training loss. This step can assure that the prediction model is robust not only for the training data. Based on the above settings, a well-trained hybrid network can λ is the class corresponding to accuracy. In this experiment, we assume that λ be obtained.

Classification accuracy is chosen as the machine learning metric to evaluate our supervised learning model. Table 4 is the class corresponding to accuracy. In this experiment, we assume that the maximum long cycle maintenance time interval is 900 days. The classes are the number of segments that 900 days are divided into. For example, we set the prediction system output to binary classification. The 900 days will be divided into two segments (Class 2): the first segment is 0–450 days and the second is 451–900 days. If we get an output which is the first one, it means that the long cycle maintenance $\frac{1}{300}$ days. $\frac{1}{300}$ degree $\frac{1}{300}$ days $\frac{1}{300}$ da time is predicted within 0–450 days. If the output is the second, the long cycle maintenance time is predicted within 451–900 days. For Class 3, the 900 days are divided into three segments, 0–300, 301–600, and 601–900 days; the same method applies to Class 4–Class 9. It is clear that the more the our method is higher than 70% for the long cycle maintenance within 100 days, which can enhance the **Class Accuracy** wind turbine management for power companies. classes, the lower the accuracy. The accuracy of Class 9 reaches 73%, e.g., the prediction accuracy of

Table 4. The prediction accuracy of long cycle maintenance time.

Class	Accuracy
2	95%
3	92%
4	87%
5	84%
6	76%
7	77%
8	76%
9	73%

5. Conclusions

This paper proposes a machine learning framework to predict the long cycle maintenance time of wind turbines efficiently. In addition to the sensor data, we also analyze the characteristics of event codes to extract two useful features including the frequency of the event code and the NAN period, which enhances the prediction performance. Different from the traditional CNN, the hybrid network, which is based on two layers of 1-D CNN and the RBF-SVM, benefits from invariant features learned from CNN and an accurate classification output predicted by SVM. The hybrid network achieves better prediction results for the long cycle maintenance time of the wind turbine. Experimental results show that the prediction accuracy of our method is higher than 70% for the long cycle maintenance within 100 days, which can really enhance the wind turbine management for power companies. In the future, we will make use of more data from sensors to discover insightful information for event prediction and further enhance the maintenance of wind turbines.

Author Contributions: C.-H.Y., M.-H.L., C.-H.L., C.-E.Y. and M.-J.C. conceived and designed the experiments; C.-E.Y. conducted an investigation; C.-H.Y. and C.-H.L. provided the resources; M.-H.L., C.-H.L. and C.-E.Y. performed the experiments; C.-H.Y. and C.-H.L. validated the experimental results; M.-H.L. visualized the experimental data; C.-H.Y. and M.-J.C. supervised the research; C.-H.Y. and C.-E.Y. wrote the original draft; C.-H.Y., M.-H.L., C.-H.L., C.-E.Y. and M.-J.C. reviewed and edited the article.

Funding: This work was supported by the Taiwan Power Company, Taiwan, under the Grant TPC-546-030-0135. This work was also financially supported by Ministry of Science and Technology, Taiwan, under the Grants MOST 105-2221-E-110-094-MY3, 106-2221-E-110-083-MY2, 107-2218-E-003-003- and 107-2218-E-110-004-.

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

References

- 1. Raposo, D.; Rodrigues, A.; Sinche, S.; Silva, J.S.; Boavida, F. Industrial IoT monitoring: Technologies and architecture proposal. *Sensors* **2018**, *18*, 3568. [\[CrossRef\]](http://dx.doi.org/10.3390/s18103568) [\[PubMed\]](http://www.ncbi.nlm.nih.gov/pubmed/30347883)
- 2. Qin, T.; Wang, B.; Chen, R.; Qin, Z.; Wang, L. IMLADS: Intelligent maintenance and lightweight anomaly detection system for internet of things. *Sensors* **2019**, *19*, 958. [\[CrossRef\]](http://dx.doi.org/10.3390/s19040958) [\[PubMed\]](http://www.ncbi.nlm.nih.gov/pubmed/30813486)
- 3. Zhu, M.; Liu, C. A Correlation Driven Approach with Edge Services for Predictive Industrial Maintenance. *Sensors* **2018**, *18*, 1844. [\[CrossRef\]](http://dx.doi.org/10.3390/s18061844) [\[PubMed\]](http://www.ncbi.nlm.nih.gov/pubmed/29874887)
- 4. Wu, Z.; Guo, Y.; Lin, W.; Yu, S.; Ji, Y. A weighted deep representation learning model for imbalanced fault diagnosis in cyber-physical systems. *Sensors* **2018**, *18*, 1096. [\[CrossRef\]](http://dx.doi.org/10.3390/s18041096)
- 5. Sachat, A.E.; Meristoudi, A.; Markos, C.; Sakellariou, A.; Papadopoulos, A.; Katsikas, S.; Riziotis, C. Characterization of industrial coolant fluids and continuous ageing monitoring by wireless node—Enabled fiber optic sensors. *Sensors* **2017**, *17*, 568. [\[CrossRef\]](http://dx.doi.org/10.3390/s17030568) [\[PubMed\]](http://www.ncbi.nlm.nih.gov/pubmed/28287488)
- 6. Santa, J.; Sanchez-Iborra, R.; Rodriguez-Rey, P.; Bernal-Escobedo, L.; Skarmeta, A.F. LPWAN-based vehicular monitoring platform with a generic IP network interface. *Sensors* **2019**, *19*, 264. [\[CrossRef\]](http://dx.doi.org/10.3390/s19020264) [\[PubMed\]](http://www.ncbi.nlm.nih.gov/pubmed/30641877)
- 7. Changhua Coastal Industrial Park. Available online: [https://en.wikipedia.org/wiki/Changhua_Coastal_](https://en.wikipedia.org/wiki/Changhua_Coastal_Industrial_Park) [Industrial_Park](https://en.wikipedia.org/wiki/Changhua_Coastal_Industrial_Park) (accessed on 1 February 2019).
- 8. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*, 3rd ed.; Morgan Kaufmann: Burlington, MA, USA, 2006.
- 9. Agrawal, R.; Srikant, R. Fast algorithms for mining association rules. In Proceedings of the 20th International Conference on Very Large Data Bases, Santiago, Chile, 12–15 September 1994; pp. 487–499.
- 10. Caselitz, P.; Giebhardt, J. Rotor condition monitoring for improved operational safety of offshore wind energy converters. *J. Sol. Energy Eng.* **2005**, *127*, 253–261. [\[CrossRef\]](http://dx.doi.org/10.1115/1.1850485)
- 11. Márquez, F.P.G.; Tobias, A.M.; Pérez, J.M.P.; Papaelias, M. Condition monitoring of wind turbines: Techniques and methods. *Renew. Energy* **2012**, *46*, 169–178. [\[CrossRef\]](http://dx.doi.org/10.1016/j.renene.2012.03.003)
- 12. Muller, H.; Poller, M.; Basteck, A.; Tilscher, M.; Pfister, J. Grid compatibility of variable speed wind turbines with directly coupled synchronous generator and hydro-dynamically controlled gearbox. In Proceedings of the Sixth International Workshop on Large-Scale Integration of Wind Power and Transmission Networks for Offshore Wind Farms, Delft, The Netherlands, 26–28 October 2006; pp. 307–315.
- 13. Schroeder, K.; Ecke, W.; Apitz, J.; Lembke, E.; Lenschow, G. A fibre Bragg grating sensor system monitors operational load in a wind turbine rotor blade. *Meas. Sci. Technol.* **2006**, *17*, 1167. [\[CrossRef\]](http://dx.doi.org/10.1088/0957-0233/17/5/S39)
- 14. Soua, S.; van Lieshout, P.; Perera, A.; Gan, T.H.; Bridge, B. Determination of the combined vibrational and acoustic emission signature of a wind turbine gearbox and generator shaft in service as a pre-requisite for effective condition monitoring. *Renew. Energy* **2013**, *51*, 175–181. [\[CrossRef\]](http://dx.doi.org/10.1016/j.renene.2012.07.004)
- 15. Jasiūnienė, E.; Raišutis, R.; Šliteris, R.; Voleišis, A.; Jakas, M. Ultrasonic NDT of wind turbine blades using contact pulse-echo immersion testing with moving water container. *Ultragarsas* **2008**, *63*, 28–32.
- 16. Kusiak, A.; Verma, A. A data-mining approach to monitoring wind turbines. *IEEE Trans. Sustain. Energy* **2012**, *3*, 150–157. [\[CrossRef\]](http://dx.doi.org/10.1109/TSTE.2011.2163177)
- 17. Kusiak, A.; Verma, A. A data-driven approach for monitoring blade pitch faults in wind turbines. *IEEE Trans. Sustain. Energy* **2011**, *2*, 87–96. [\[CrossRef\]](http://dx.doi.org/10.1109/TSTE.2010.2066585)
- 18. Kim, K.; Parthasarathy, G.; Uluyol, O.; Foslien, W.; Sheng, S.; Fleming, P. Use of SCADA data for failure detection in wind turbines. In Proceedings of the ASME 5th International Conference on Energy Sustainability, Washington, DC, USA, 7–10 August 2011; pp. 2071–2079.
- 19. Schlechtingen, M.; Santos, I.F.; Achiche, S. Using data-mining approaches for wind turbine power curve monitoring: A comparative study. *IEEE Trans. Sustain. Energy* **2013**, *4*, 671–679. [\[CrossRef\]](http://dx.doi.org/10.1109/TSTE.2013.2241797)
- 20. Keras. Available online: <https://github.com/keras-team/keras> (accessed on 1 February 2019).

© 2019 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 [\(http://creativecommons.org/licenses/by/4.0/\)](http://creativecommons.org/licenses/by/4.0/.).