buildings

MDPI

Article

Predictive Maintenance 4.0 for Chilled Water System at Commercial Buildings: A Methodological Framework

Malek Almobarek [1,*], Kepa Mendibil [1] and Abdalla Alrashdan [2]

[1] Department of Design, Manufacturing and Engineering Management, Faculty of Engineering, University of Strathclyde, Glasgow G1 1XQ, UK
[2] Industrial Engineering Department, College of Engineering, Alfaisal University, Riyadh 50927, Saudi Arabia
* Correspondence: malek.almobarek@strath.ac.uk

Abstract: Predictive maintenance is considered as one of the most important strategies for managing the utility systems of commercial buildings. This research focused on chilled water system (CWS) components and proposed a methodological framework to build a comprehensive predictive maintenance program in line with Industry 4.0/Quality 4.0 (PdM 4.0). This research followed a systematic literature review (SLR) study that addressed two research questions about the mechanism for handling CWS faults, as well as fault prediction methods. This research rectified the associated research gaps found in the SLR study, which were related to three points; namely fault handling, fault frequencies, and fault solutions. A framework was built based on the outcome of an industry survey study and contained three parts: setup, machine learning, and quality control. The first part explained the three arrangements required for preparing the framework. The second part proposed a decision tree (DT) model to predict CWS faults and listed the steps for building and training the model. In this part, two DT algorithms were proposed, C4.5 and CART. The last part, quality control, suggested managerial steps for controlling the maintenance program. The framework was implemented in a university, with encouraging outcomes, as the prediction accuracy of the presented prediction model was more than 98% for each CWS component. The DT model improved the fault prediction by more than 20% in all CWS components when compared to the existing control system at the university.

Keywords: predictive maintenance; Industry 4.0; Quality 4.0; decision tree algorithm; chilled water system; HVAC; commercial buildings; industrial engineering; engineering management

check for updates

Citation: Almobarek, M.; Mendibil, K.; Alrashdan, A. Predictive Maintenance 4.0 for Chilled Water System at Commercial Buildings: A Methodological Framework. *Buildings* **2023**, *13*, 497. https://doi.org/10.3390/buildings13020497

Academic Editor: Etienne Saloux

Received: 12 January 2023
Revised: 7 February 2023
Accepted: 10 February 2023
Published: 12 February 2023

1. Introduction

1.1. Overview

In the past decades and especially today, the downtowns of large cities have been mainly made up of commercial buildings, and the owners or the caretakers of these buildings make efforts to develop strategies and plans for their upkeep and to control their equipment. One of the said strategies is predictive maintenance (PdM), which is defined as a strategical monitoring approach that optimizes the usability of a particular equipment/system [1]. On a related note, PdM 4.0, which is in line with Industry 4.0/Quality 4.0, can determine the best time to detect equipment/system faults using machine learning (ML) models or artificial intelligence (AI) [2]. Bousdekis and others have outlined the benefits of developing the said strategies, especially PdM 4.0, and indicated that they have shown a positive impact for improving many aspects related to the organizations, such as maintenance and operation costs, replacement costs, repair downtime and verifications, machine failures, spare part stock, part service life, production, operator safety, and overall profit [3]. Using the outputs of a novel AI approach [4], PdM 4.0 can be considered a control task that maintains buildings efficiently. Moreover, PdM 4.0 ensures the sustainability of the buildings, as it allows the human and the machine to be harmonized [5,6]. In contrast, Achouch and others discussed the challenges of PdM 4.0 regarding four aspects, which are financial and organizational limitations, the limitations of data sources, activity limitations for repairing machines, and the limitations in the deployment of industrial PdM models [7].

This article focuses on one of the most important building utility systems, which is the chilled water system (CWS). It is part of the heat, ventilation, and air conditioning (HVAC) system and contains four main components, which are chillers, cooling towers, pumps, and terminal units, which are operated in an interactive way [8]. It plays a significant role in controlling the ambient temperature, which should meet the satisfaction of the buildings' occupants [9]. Furthermore, efficiently maintaining a CWS will prevent premature replacement of its components and save energy [8]. Therefore, the goal of this research was to present a comprehensive PdM 4.0 program for CWS via a methodological framework. There are a number of studies that have presented PdM 4.0 programs for CWS, focusing on either on one, two, or three components of the system to predict its faults. These faults were predicted through ML techniques such as the decision tree (DT) algorithm [10–12], artificial neural network algorithm [13–15], and support vector machine algorithm [16–18]. Moreover, a systematic literature review (SLR) addressed PdM 4.0 applications in 168 studies on CWS [19]. This research followed this SLR study, which was underpinned by two research questions, responded to three research gaps, and proposed a route for PdM 4.0. Table 1 shows the research questions and the research gaps, while Figure 1 visualizes the proposed PdM 4.0 route.

Table 1. Research Questions and Gaps.

Research Question	Research Gap
(1) How can faults be identified, in order to predict them?	(1) The literature did not consider the same faults and only concentrated on selected faults, as some faults were either not stated/mentioned or were not fully described.
(2) What are the methods that can be used to predict the faults?	(2) The current literature does not specify how data were collected or justify the period or the frequency of the collected data, as well as being limited to testing the model and not controlling it. (3) The suggested programs/frameworks/models did not contain, or contained inconclusive, solutions for the mentioned faults from a management point of view, as they ended at how to detect/predict the faults. Moreover, these programs did not comprehensively study/cover the whole system.

Figure 1. PdM 4.0 Route.

In addition, this research utilized the outcomes of an industry survey (IS) study [20] in building the framework. The IS study provided two tools that were used in this research, which were the fault frequencies and fault solutions. The utilization of these tools are explained in the next sections. Following Jebreen [21], and as the IS study collected quantitative and qualitative data from a number of professional participants [20], this research employed an inductive approach that proposed a methodological framework for a PdM 4.0 program for CWS in commercial buildings. From a philosophical point of view, this research follows the pragmatic paradigm, which was introduced by William James in 1898 [22]. This is defined as a philosophical tradition that considers ideologies as instruments for prediction as well as for problem solving [23]. According to Sakib and Wuest, the aforementioned research paradigm is ideal for PdM research [24]. Therefore, this research is an extension of the previously published SLR and IS studies [19,20]. It rectifies the research gaps that were identified by the SLR study and applies the recommendations of the IS study.

From an ML point of view, this research utilized the DT algorithm within the proposed framework; as the SLR study indicated that DT typically shows a high accuracy for predicting faults that affect the condition of a CWS over time [19]. The next subsection gives an overview of the mentioned algorithm.

1.2. Decision Tree Algorithm

DT is a common ML algorithm that is mainly used for classification, prediction, and regression applications. It has many benefits, and Sharma and Kumar argued that it can be used to predict continuous and discrete values [25]. They also indicated that it can capture nonlinear relationships, as well as being easier to use than other ML algorithms for understanding, interpretation, and visualization [25].

The DT has a tree-like structure, with a root node and intermediate nodes that split into branches. The last intermediate node is split into leaves and is terminated with an end node. Each node represents a classification or prediction feature. A branch or a leaf represents the possible value of the feature. The path from the root node to the end node is labeled using the predicted outcome or target classification, which is assigned using an existing training dataset. Using supervised training algorithms, the features are split recursively from top down according to certain criteria. Figure 2 depicts the general structure of the tree [26].

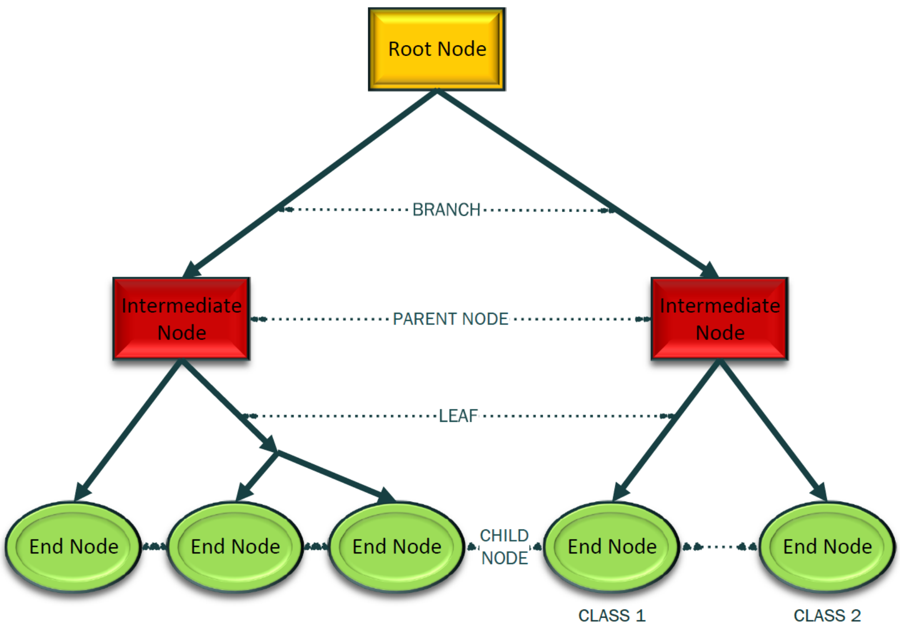


Figure 2. General Structure of the DT.

2. Methodological Framework

Constructing management frameworks for projects, continuous activities, or any other core program of building facility management gives structure to the program and allows corrective measures that can achieve related goals [27]. The framework in this research was built from an engineering management point of view, where each part of the framework contains multiple managerial steps. Table 2 describes the parts of the proposed framework, as well as the objective of each part.

Table 2. Framework Structure.

Part	Objective
Setup	• To understand the CWS at the building under study, in order to identify the numbers of each component, as well as their location at the site; • To ensure that the data reading tools are in the right locations; • To prepare the data collection plan, which includes data collection tools, determining the schedule of data collection, and formation of the team who will collect the data.
Machine Learning	• To formulate the algorithm, train the prediction model, and test it.
Quality Control	• To make a control plan for the maintenance program and evaluate the prediction model.

The above parts should be followed in the same logical order as shown and detailed in the next subsections.

2.1. Setup Part

In order to prepare the framework, three stages are suggested, to be gone through in the same order as they are listed in the following subsubsections.

2.1.1. CWS Drawing

As recommended by SLR [19], the first step in preparing the framework is to understand the as-built drawing of CWS at the building under study, in order to determine the number of CWS components installed there and to determine their locations around the site; and then to study the whole system accordingly. Such drawings show the actual building layout and are normally handed over to the facility management after completion of the building construction [28]. Following the standard in [29], this research made a simplified schematic CWS drawing; in order to easily identify the numbers of each component, as shown in Figure 3.

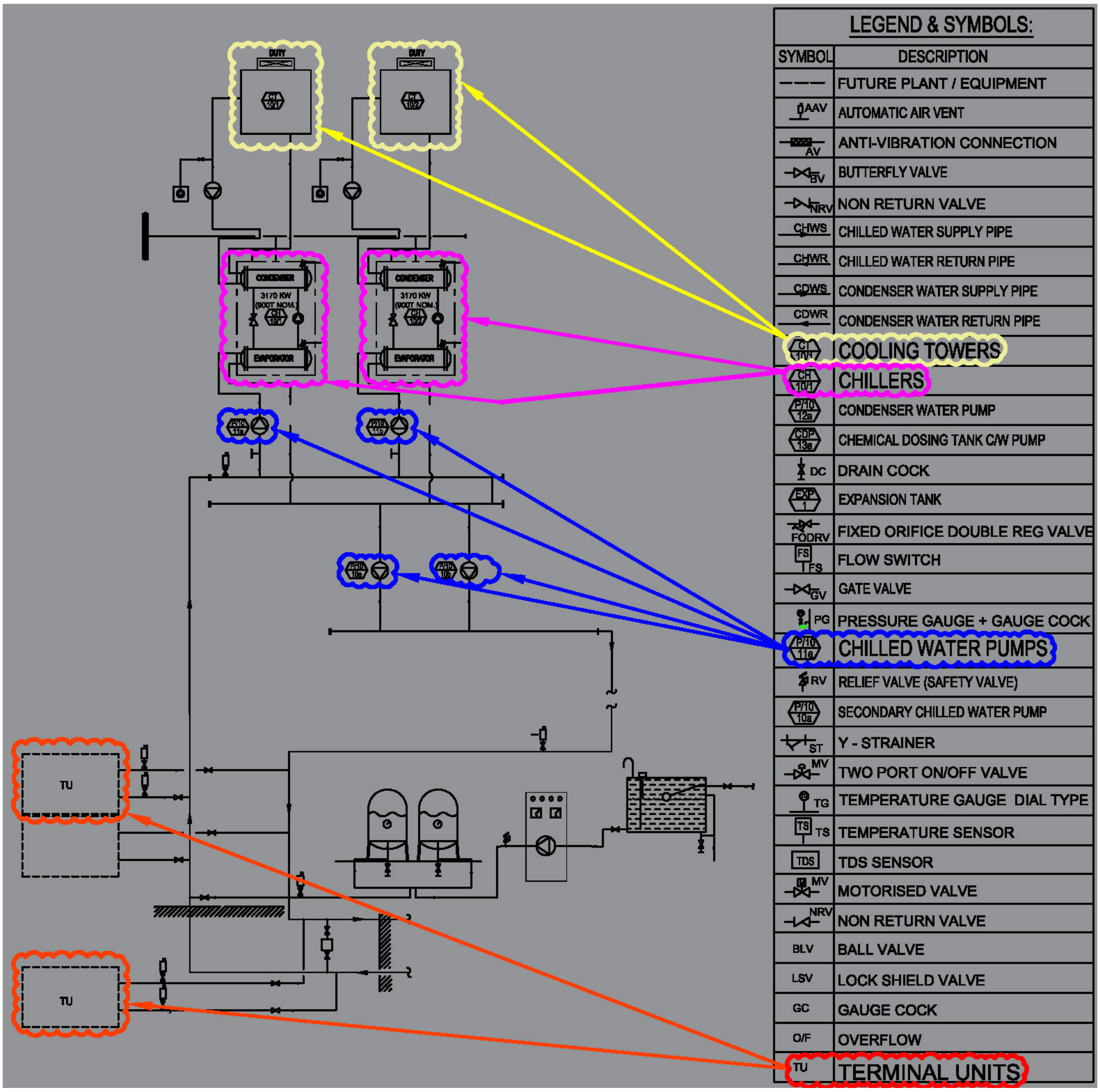


Figure 3. CWS As-Built Drawing.

2.1.2. Reading Tools for Operational Parameters

Following the SLR and IS studies [19,20], one of the fundamentals of PdM 4.0 is the datasets that contain the readings of the CWS operational parameters. Here, operational parameters are defined as quantifiable factors that give numerical data about the performance of the CWS [19]. In this research, the operational parameters chosen were the temperature of water leaving the chillers and cooling towers, pressure for pumps, and the space temperature for terminal units; as they are the best for showing the health condition of these components [20].

In order to collect the readings of these parameters, the associated tools were assumed to be available at the building under study. The measurement tools can be meters, gauges, sensors, thermostats, or any other agent, such as the building management system (BMS). In case of the unavailability of these reading tools, the authors in [30,31] outlined procedures on how to install such tools. Following the standard operating procedure in [29], Table 3 shows the best location for installing the reading tool for each CWS component in the building under study.

Table 3. Best Location for Reading Tools.

CWS Component	Location
Chiller	Chilled water supply header
Cooling Tower	Straight pipeline entering the condenser
Pump	Discharge pipeline
Terminal Unit	1.5 m above the floor level in a space or in the return air duct

Once the reading tools are installed, they have to be connected to the computer unit (CU) that will be used in the PdM program. Kayastha and others outlined a procedure for how to connect such tools to computers [32]. This course of action was utilized and is explained in the third subsection of the methodological framework section (Quality Control).

2.1.3. Data Collection

After determining the numbers of each component and finalizing the reading tools, the last stage of the setup is data collection. The IS study proposed time frequencies to collect data in a building [20]. Following these proposals, the readings of water temperature leaving a particular chiller should be taken every thirty minutes over a study period of twelve weeks. The same should be applied for cooling towers, but over a study period of sixteen weeks. With regard to pumps, the readings of pressure should be taken every hour over a study period of twenty-four weeks. For terminal units, the readings of space temperature should be taken every forty-five minutes over a study period of eight weeks. The SLR and IS studies suggested utilizing a check sheet to collect the data for each component, which should contain the readings as well as the inspection results. The inspection results will be either '1' in case of fault or '0' in case of no fault. As recommended by the IS study, the check sheet must be filled out by experienced technicians or users [20]. Each check sheet should be recorded by two team members, one for the morning and part of the afternoon shift, and one for the evening and the second part of the afternoon shift. Appendix A shows a proposed check sheet for terminal units, and the same was applied for other CWS components, taking into consideration the differences in the time intervals and the unit of the operational parameters between the components. After collecting the data, a file for each particular component should be created in Excel, and then the information from the related check sheet should be logged. Thus, each file should contain two columns, one for the readings and the another for the inspection results [20], and then it should be saved in the CU in csv format. Therefore, these files present the required datasets, and at to this point, the setup is completed, and accordingly the ML part can be started, as explained in the next subsection.

2.2. Machine Learning Part

In this part, two DT algorithms are recommended for use, which are the C4.5, a successor of the iterative dichotomiser 3 (ID3), and the classification and regression tree (CART) algorithm, as they are efficient for splitting the trees [33]. The basic principle of the splitting mechanism is to select a root node from the 'N' features and subsequently decide which attribute should be used next as the intermediate node. Different statistical criteria should be used to make these decisions, such as the Gain Ratio and the Gini Index. According to Grąbczewski [34], the Gain Ratio criterion is mainly used in the C4.5 algorithm, while the Gini Index is used in the CART algorithm. The Gain Ratio is calculated as in Equation (1):

$$Gain\ Ratio_{(A)} = \frac{Information\ Gain}{SplitInfo} = \frac{Entropy(parent) - \sum_{j=1}^{k} Entropy(j, child)}{\sum_{j=1}^{k} \frac{D_j}{D} log_2 \frac{D_j}{D}} \quad (1)$$

In information theory, entropy measures the uncertainty in data. The entropy(parent) measures the amount of randomness (impurity) in the parent node before it splits. D is the number of instances in the parent node and Dj is the number of instances in the child j, and k is the number of discrete values of an attribute A, which is tested at the parent node for splitting. The entropy at each child node is found using Equation (2):

$$Entropy = -\sum_{i=1}^{n} p_i \, log_2 \, p_i \tag{2}$$

where p_i is the probability of selecting an instance in class i, and n is the number of classes. The attribute that is selected for splitting the parent node is the one with the highest Gain Ratio. Similarly, the Gini Index for the CART algorithm can be found by Equation (3):

$$Gini\ Index_{(A)} = \sum_{j=1}^{k} \frac{D_j}{D} Gin(j, child) \tag{3}$$

Similarly to the Entropy, the Gini Index measures the impurity at the parent node. The Gini of a child node is found using Equation (4):

$$Gini = 1 - \sum_{i=1}^{n} p_i^2 \tag{4}$$

The attribute that is selected for splitting at the parent node is the one with the smallest Gini Index. In this research, this attribute is the operational parameter of each CWS component. Many programming languages/software can read collected data and train the prediction model, such as Python [35]. The software should be installed in the CU and the required codes should be written in a way that allows reading the files (datasets) for each CWS component, which were mentioned in the data collection stage of the setup part, and then to train and test the model. The next section of this article (Implementation and Discussion) gives a case study on how a prediction model is trained and tested.

2.3. Quality Control Part

This is the last part of the proposed framework, and its goal is to ensure the prediction model is working correctly, as well as to rectify the faults immediately. To do this, this research suggests making a control plan, which should contain monitoring and response actions [36]. Table 4 clarifies the descriptions of these two actions, as well as who is responsible for executing each action.

Table 4. Control Plan.

Quality Control Action	Description	Responsible
Monitoring	The prediction model should be connected to the reading tools, which were connected to the CU during the setup part. This is to ensure that the CU shows a continuous reading for each CWS component.	Information Technology (IT) Department or Programming Supplier
Response	When the prediction model shows a fault, which is a "1" as a result of a particular reading, the related component should be inspected and then to be rectified as per the solutions tabulated in the IS article [20].	Facility Department Officer/technician

After that, the response actions should be documented as follows:

- Listing the lessons learned from the proposed PdM program, such as focusing on the faults that occurred, and then brainstorming permanent solutions to avoid the reoccurrence of such faults;

- Tracking the spare part stock;
- Ensuring that the CU is working efficiently;
- Training more technicians to be familiar with the prediction model;
- Making regular reports about the performance of the proposed PdM program for future improvements.

3. Implementation and Results

This section presents a case study on the proposed framework. The case study was performed at a university in Riyadh city, Kingdom of Saudi Arabia. Implementation of the framework was carried out as per the three parts proposed in the previous section (Methodological Framework).

3.1. Implementaion of Setup Part

3.1.1. CWS Drawing

The main goal of the proposed framework is to make a PdM 4.0 program that considers the whole CWS (i.e., all CWS components). Therefore, to start implementing the framework, the CWS as-built drawing was collected and then, following Figure 3, the numbers of each CWS component were determined, as shown in Table 5, as well as their locations around the site.

Table 5. Number of CWS Components.

CWS Component	Quantity
Chiller	5
Cooling Tower	7
Pump	19
Terminal Unit	72

3.1.2. Reading Tools

At this stage, the standard shown in Table 3 was followed, and it was ensured that the reading tools for the operational parameters of each CWS component were in the best location. Figures 4–7 show the reading tool location for each CWS component. As stated in the previous section, these tools read the temperature for water leaving each chiller and cooling tower, the pressure for pumps, and the space temperature for terminal units.

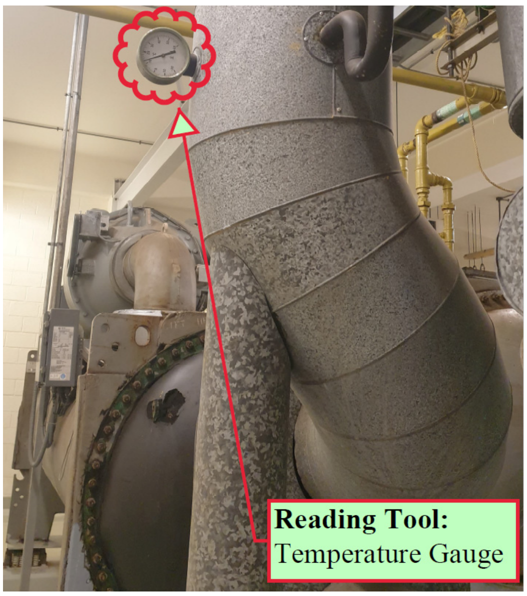


Figure 4. Chiller Reading Tool.

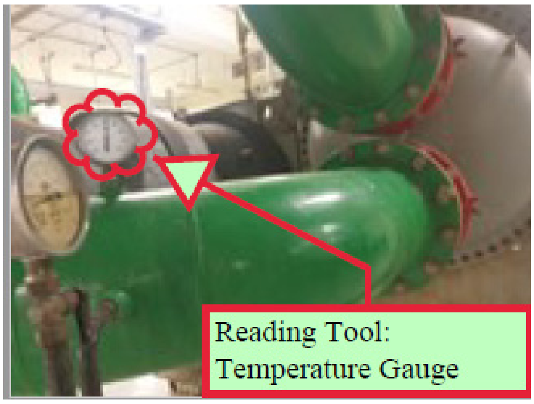


Figure 5. Cooling Tower Reading Tool.

Figure 6. Pump Reading Tool.

Figure 7. Terminal Unit Reading Tool.

Through the IT department of the university, these reading tools were connected via sensors to a CU, to be ready for the quality control part.

3.1.3. Data Collection

The most important stage in setting up the PdM 4.0 program was the datasets required to build the prediction model. As the previous two stages had been finalized, data collection was started as per the recommendations of the IS study [20]. This research used the recommended minimum frequencies as time intervals when collecting the data, and the recommended maximum frequencies were used as study periods for each CWS component.

Twelve qualified technicians from the university were assigned for the subject matter. Two operational units for each CWS component were selected as subjects. The readings for the water temperatures of each chiller and cooling tower were collected using check sheets. The same was performed for the pressures for each pump and the space temperatures for each terminal unit. In addition, the inspection result, which was either a fault "1" or fault free "0", was included for each check sheet. Appendix B illustrates a fully filled one day check sheet for a particular pump, and Table 6 shows the data collection plan.

Table 6. Data Collection Plan.

CWS Component	Time Interval for Reading and Inspection (Minutes)	Study Time (Weeks)	Study Period
Chiller	30	12	From 29 May 2022 to 20 August 2022
Cooling Tower	30	16	From 29 May 2022 to 17 September 2022
Pump	60	24	From 29 May 2022 to 12 November 2022
Terminal Unit	45	8	From 29 May 2022 to 23 July 2022

After that, an Excel file was created for each component, and the information in all related check sheets was transferred to the associated Excel file. Following the procedure proposed in the methodological framework section, each Excel file represented a dataset that contained two cells, one for the readings and another for the inspection results, as shown in Appendix C for a one of the cooling towers. After that, each file was named and saved in csv format. For example, for a particular pump, the file was named and saved as "pu.csv"; so it could be read when training the prediction model, as shown in the next section.

3.2. Implementation of the Machine Learning Part

A DT was built for each CWS component selected. As stated in the methodological framework section, the faults of each component were predicted using the related attributes. Table 7 shows the attribute and the training data size for each unit of the selected CWS components that were in operation.

Table 7. Main Inputs of The Prediction Model.

CWS Component	Attribute	Data Size
Chiller	Water Leaving Temperature (°C)	2688
Cooling Tower	Water Leaving Temperature (°C)	3584
Pump	Pressure (Bar)	2688
Terminal Unit	Space Temperature (°C)	1288

The C4.5 and CART algorithms were used to train the tree. Different training parameters were used to optimize the tree accuracy. The parameters included the training to testing ratio and the level of pruning. The model was executed in Python, with the script shown in Figure 8 being for a particular pump. The same was done with other CWS components, taking into consideration the changes in file reading/loading.

```
# Load libraries
!pip install pydotplus
import csv
import numpy as np
import pandas as pd
from sklearn.tree import DecisionTreeClassifier # Import Decision Tree Classifier
from sklearn.model_selection import train_test_split # Import train_test_split function
from sklearn import metrics #Import scikit-learn metrics module for accuracy calculation
# load training dataset
load_file= pd.read_csv("pu.csv")
#split dataset in features and target variable
feature_cols = ['Pressure']
X = load_file.iloc[:, :-1].values  # Features
y = load_file.iloc[:, -1].values  # Target variable
# Split dataset into training set and test set
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=1) # 70% training and 30% test
# Create Decision Tree classifer object
clf = DecisionTreeClassifier(criterion='gini',max_depth = 1)

# Train Decision Tree Classifer
clf = clf.fit(X_train,y_train)

#Predict the response for test dataset
y_pred = clf.predict(X_test)
# Model Accuracy, how often is the classifier correct?
print("Accuracy:")
#print("Predicted values:")
out1=metrics.accuracy_score(y_test, y_pred)
print(out1)
#write output to file CSV
f=open('results.csv', 'w')
writer=csv.writer(f, delimiter=',')

for i in range(0, len(y_pred)):
    float_list = [X_test[i],y_pred[i]]
    #print(len(y_pred),i,float_list)
    writer.writerow(float_list)

f.close()
import matplotlib.pyplot as plt
from sklearn.tree import plot_tree
plt.figure(figsize=(10,8), dpi=70)
plot_tree(clf, feature_names=feature_cols, filled=True);
```

Figure 8. DT Python Code.

The initial run of the training stage was performed without pruning, which led to prediction overfitting, as can be seen for the chiller tree in Figure 9. The same was done for the other CWS components.

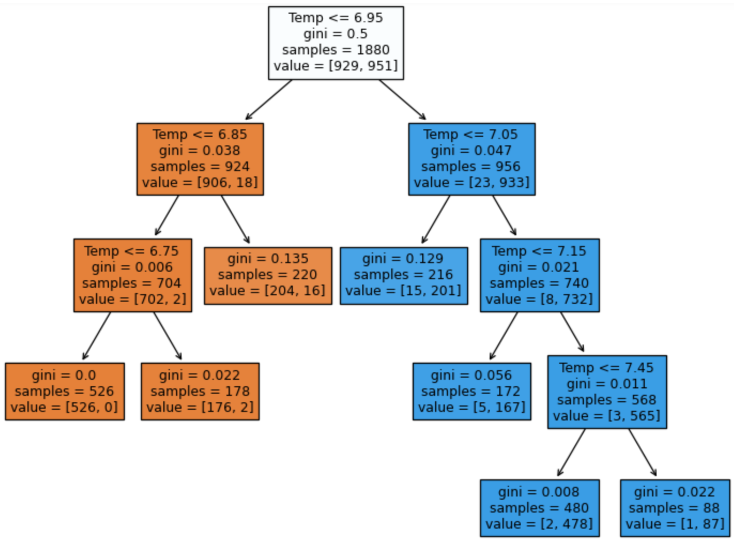


Figure 9. Chiller DT without Pruning.

Examining the different pruning methods, the optimally trained trees for each CWS component were found, as shown in Figures 10–13.

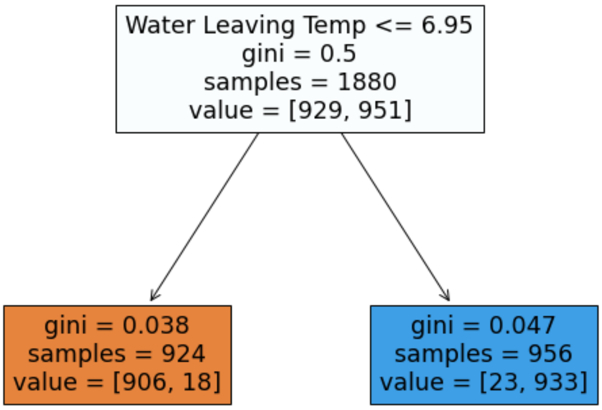


Figure 10. Chiller DT.

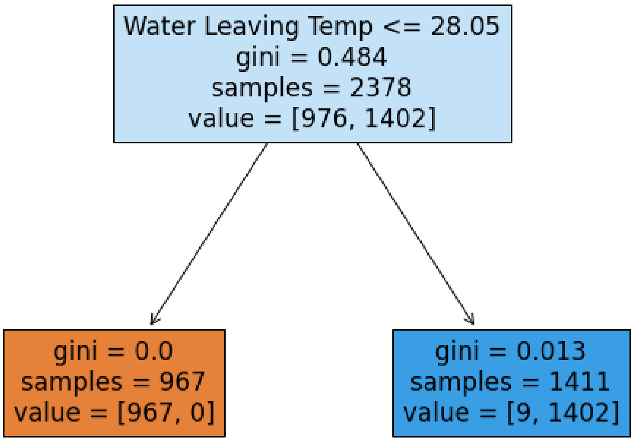


Figure 11. Cooling Tower DT.

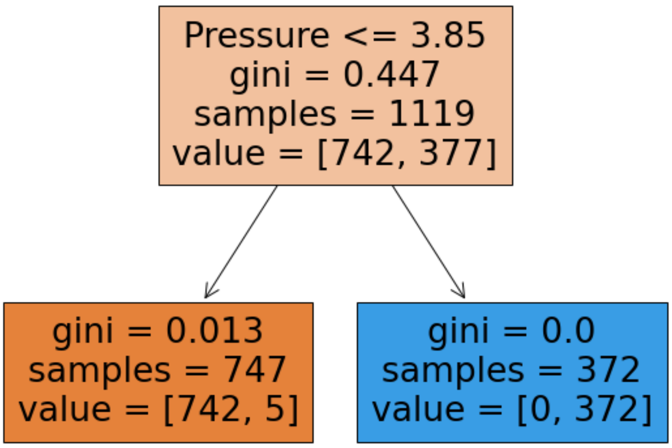


Figure 12. Pump DT.

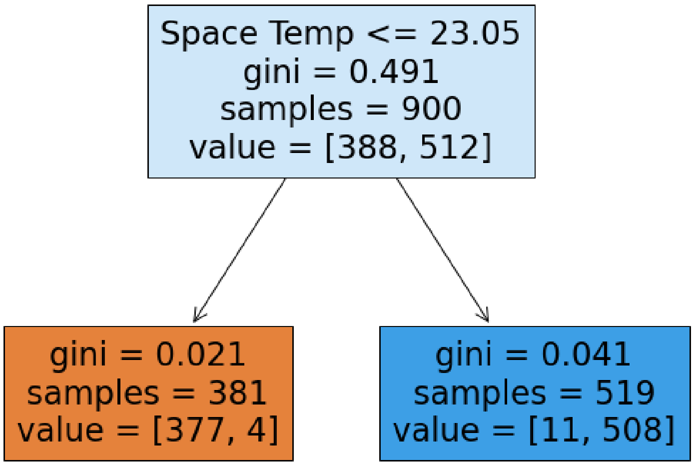


Figure 13. Terminal Unit DT.

Changing the training to testing ratio and the training algorithms had a very small impact on the prediction accuracy. A 70 to 30 percent training to testing ratio was adopted using the CART training algorithm. The prediction accuracies of each component at the optimal DT setting are presented in Table 8.

Table 8. CWS Component Prediction Accuracies.

CWS Component	Prediction Accuracy (%)
Chiller	98.50
Cooling Tower	99.60
Pump	99.80
Terminal Unit	99.20

3.3. Implementation of the Quality Control Part

After successfully building the prediction model, the control plan mentioned in Table 4 was actioned. In the monitoring of the control plan, the prediction model was connected to the CU, in order to begin the second stage of the plan (Response). After that, the readings of all CWS components at the university, which are shown in Table 5, were observed daily, as per the minimum frequencies mentioned in Table 6, for a month time, excluding weekends. For example, for a particular terminal unit, the reading of space temperature was observed every 45 min. During this period, the DT model predicted 16 different faults in the chillers, 11 different faults in the cooling towers, 12 different faults in the pumps, and 19 different faults in the terminal units, noting that the occurrence of some faults was repeated. Table 9 lists how many faults in total were predicted by the DT model within the said period, as well as the fault that occurred most for each CWS component. All fault signals from the CU, which were displayed as “1”, led to real faults around the site. Thereafter, all the faults within this period were solved immediately after their appearance, by following the solutions mentioned in [20].

Table 9. Number of CWS Faults.

CWS Component	Number of Faults	Most Occurred Fault
Chiller	101	Refrigeration Leak
Cooling Tower	113	Malfunctioning Blowdown System
Pump	79	Noisy Non-Return Valve
Terminal Unit	138	Low Static Pressure

Furthermore, the facility department at the university was advised to keep observing the readings as was convenient and to inspect the site in case of a fault “1”. In addition, they were advised to document the response action of the control plan, as per the steps proposed in the previous section (Methodological Framework). On a related note, the existing monitoring system implemented by the department was a BMS, and they were asked to give a report for the same period that was used for observing the whole CWS via the DT model. The report contained the total number of faults that were predicted by the BMS for each CWS component. Figure 14 shows a comparison between the DT model and the BMS in predicting the faults within the same period.

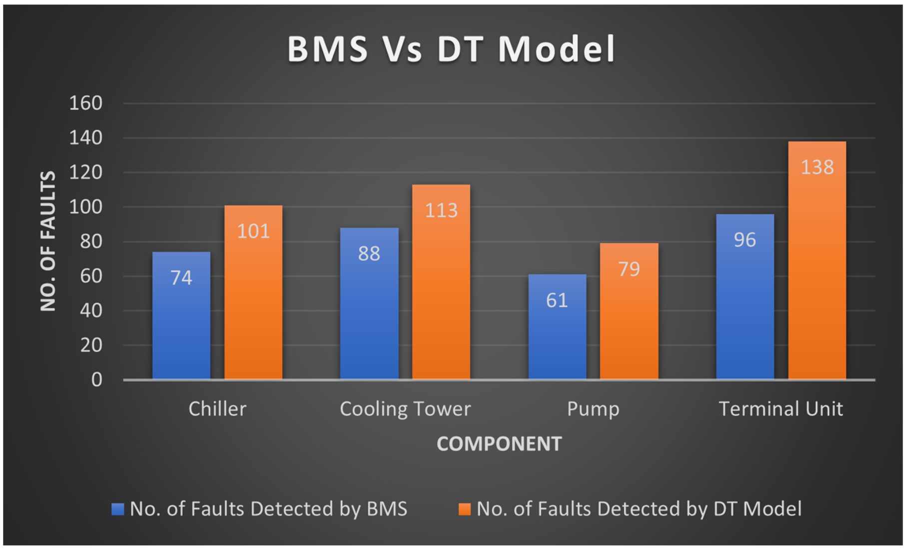


Figure 14. Comparison of Prediction Performance.

4. Discussion

The case study applied a methodological framework proposing three parts to build an efficient PdM 4.0 program. During the setup part, the first stage gave an overview of the building CWS, by determining the number of units of each component, as well as identifying their locations on site. This action was easily carried out using the schematic shown in Figure 3. The second stage was to give a clear picture about the best location for the reading tools for each CWS component. The third stage was to produce a clear map on how to collect the data, as the main objective of the setup part was to create a dataset for each CWS component. These datasets were essential to allow beginning the second part of the proposed framework. To recall what was mentioned in the previous sections, each row of a dataset contained a reading of the operational parameter in one column and its associated inspection result in another column.

In the second part, the datasets were used in building the ML model. The results were encouraging, as the DT model showed a very high prediction accuracy for each CWS component, as shown in Table 8. This confirmed that the fault frequencies proposed in [20], which were used while collecting the data of this research, are valid for tracking faults. In the third part of the proposed framework, the aforementioned control plan in Table 4 facilitated the execution of the prediction model. The empirical period of this part provided the following findings:

- The C4.5 and CART algorithms had a similar prediction accuracy for each CWS component.
- The DT model had a better performance than the BMS in predicting the faults for all CWS components, as shown in Figure 14. This fulfilled the requirements of the facility department, who manage the CWS at the university.
- The most frequent fault in chillers was refrigeration leaks. This was also confirmed by the SLR study [19], as well as the IS study [20], which reported this fault as the most common in chillers;

- A malfunctioning blowdown system was the most common fault in the cooling towers. This finding matches what was found in the IS study [20]. The IS study stated that the majority of the survey's participants suffered from this fault;
- With regard to the pumps, a noisy non-return valve occurred most often. This also matches the information provided by the IS study [20], where the majority of the survey's participants faced this fault continuously;
- Low static pressure in the terminal units occurred more than twice a day. The IS study [20] had already confirmed that the most of the survey's participants were finding this fault on a regular basis while operating the associated terminal unit;
- The solutions provided in the IS study [20] gave practical actions to rectifying the predicted faults. In this regard, one of the research gaps listed by the SLR study was that the previous 168 studies considered did not cover the whole CWS (i.e., all for components) and ended their PdM programs once detecting the faults [19]. However, the SLR study recommended having control measures, including fault solution, after completing the prediction model, which will allow a comprehensive PdM program, such as the proposed framework.

As stated in the first section of this article (Introduction), this research is a response to the mentioned SLR study [19] as well as the IS study [20]. Considering the gaps in Table 1, this study covered the first gap and prepared tools to rectify the second and the third gaps [20]. The first tool in the IS study was the frequencies, which were used in collecting the data and in controlling the whole CWS. The second tool was the solutions to faults, which were used in the quality control part. Therefore, this research has contributed to building a framework that will provide a comprehensive PdM 4.0 program for the whole CWS in commercial buildings. On a related note, this framework was implemented at another site for external validity purposes. The site is a hotel that is related to the same foundation that manages the university. The DT's prediction accuracy for each CWS component was similar to those at the university. Although this framework has obtained encouraging results, it has some challenges from a research point of view, as follows:

- The availability of the data source;
- The experience of the team who collect the data;
- The organizational culture at the building, which may not be cooperative;
- The associated costs, such as arranging the reading tools, the CU, and the labor.

5. Conclusions

This research proposed a methodological framework for a PdM 4.0 program for commercial buildings. A framework was made for one of the most important utility systems of the commercial buildings, the CWS, which has four components. These are the chillers, cooling towers, pumps, and terminal units. The framework contains three parts, which are the setup, ML, and quality control. Each part of this framework has multiple managerial stages or steps to build the maintenance program.

The setup part of this framework contained three stages. The first stage allowed efficiently understanding the building through analyzing its as-built drawing. By doing so, it was possible to determine the unit numbers of each CWS component in the building, as well as to know their locations on site. A schematic was made in this regard, to make a simplified view of such drawings. The second stage of the setup part focused on the reading tools for the CWS operational parameters. How to make the reading tools available and the best location for each tool was discussed. The readings of the operational parameters were essential for creating the datasets that are were used in the second part of the framework, ML. The operational parameters chosen in this research were the water temperatures for chillers and cooling towers, the pressures for pumps, and the space temperatures for terminal units. The third and last stage of this part addressed the data collection. It presented the data required and proposed a complete plan for collecting them. Therefore, the main goal of the setup part was to provide the datasets that were required to build a prediction model, which was explained in the second part of this framework, ML. As this research was intended to implement a PdM 4.0 program, the second part of the framework utilized ML. The DT technique was chosen to build a model for predicting the CWS faults, as recommended by the SLR study. Two DT algorithms, C4.5 and CART, were proposed to build, train, and test the model. The last part of the framework, which was quality control, proposed a control plan to evaluate the prediction model. The control plan required two actions, which were monitoring and response. Both actions proposed executing the DT model while operating the CWS, and then to control the system via many aspects, such as solving the faults predicted and documenting the outcomes of the predictive model from an engineering management point of view.

This research implemented the proposed framework in a university in Riyadh city, Kingdom of Saudi Arabia. The DT model produced encouraging results, where the prediction accuracy was 98.50 percent for chillers, 99.60 percent for cooling towers, 99.80 percent for pumps, and 99.20 percent for terminal units. Furthermore, the DT model was evaluated over an empirical period. The model gave outstanding performance in predicting the faults of all CWS components, especially when it was compared to the BMS, which was the existing control system at the university. During the said period, the DT made a 27 percent improvement in predicting the faults for chillers, 22 percent for cooling towers, 23 percent for pumps, and 31 percent for terminal units. On a separate note, refrigeration leaks, malfunctioning blowdown systems, noisy non-return valves, and low static pressure faults occurred often during this period in chillers, cooling towers, pumps, and terminal units, respectively. This confirmed the information provided by the IS study with regard to the most common faults.

Though this research, along with the SLR and IS studies, provided significant outcomes towards implementing PdM 4.0 for CWS in commercial buildings, future research agendas could explore further insights about this topic, as follows:

- To discuss how to integrate the ML models with the building automation and management systems such as BMS, for a more efficient prediction model;
- To propose an intelligent system for updating the datasets, which are required to build the prediction model, in order to rise the control efficiency of commercial buildings;
- To investigate and give more focus to the repeated occurrence of faults, especially the aforementioned four faults, which are refrigeration leaks in chillers, malfunctioning blowdown systems in cooling towers, noisy non-return valves in pumps, and low static pressure in terminal units;
- To use the ideas of this research, which built the framework, and extend them to other HVAC systems such as heating systems, as well as for other utility systems, such as the electrical system.

Author Contributions: Conceptualization, M.A. and K.M.; methodology, M.A.; investigation and analysis, M.A. and A.A.; writing—original draft, M.A.; writing—review and editing, A.A.; data curation, M.A.; visualization, M.A.; supervision, K.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to privacy restrictions.

Acknowledgments: Many thanks to the DMEM department at the University of Strathclyde, especially Anja Maier, Jorn Mehnen, Xichun Luo, Ian Whitfield, Laura Hay, and Gillian Eadie for their guidance, advice, and support. Many thanks also to Alfaisal University President (Mohammed Alhayaza), VP External relation and advancement (H.R.H. Maha Bint Mishari Bin Abdulaziz AlSaud), VP Admin and Finance (Khaled AlKattan), and Acting Dean of College of Engineering (Muhammad Anan) for their unlimited support. A special thanks goes to the Associate Dean of Alfaisal's College of Engineering for Academic and Student Affairs (Sobhi Mejjaouli) for his perpetual cooperation and support. In addition, thanks go to the operation and maintenance manager at Alfaisal University (Kaleemoddin Ahmed) and his team for their great assistance during this research.

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

Appendix A

Check Sheet (Building Name: ....)

Component: Terminal Unit #.......

Day & Date: ....

Time	Space Temperature (°C)	Fault Free (0) Fault (1)	Time	Space Temperature (°C)	Fault Free (0) Fault (1)
6:30			14:45		
7:15			15:30		
8:00			16:15		
8:45			17:00		
9:30			17:45		
10:15			18:30		
11:00			19:15		
11:45			20:00		
12:30			20:45		
13:15			21:30		
14:00			22:15		
			23:00		
Inspector Name:			Inspector Name:		
Signature:			Signature:		

Appendix B

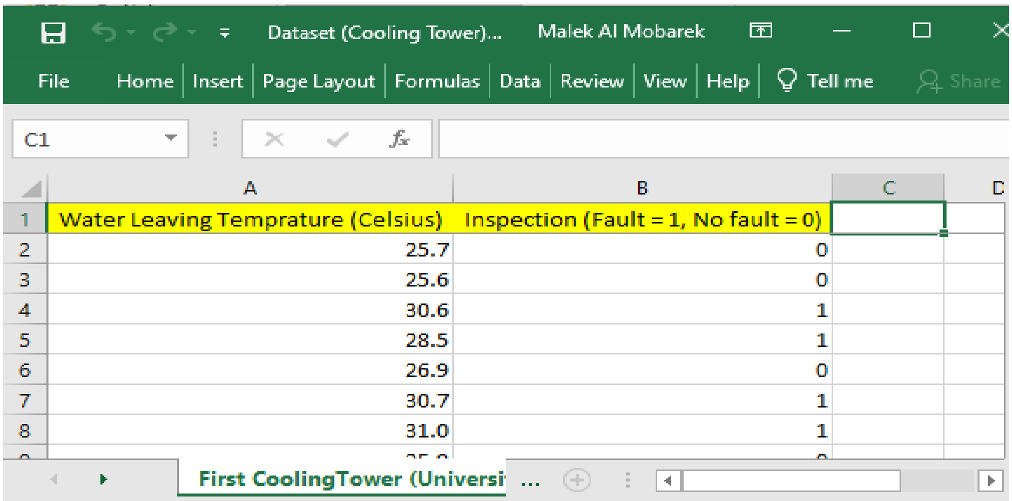
Check Sheet (Building Name: Students Center)

Component: Pump #1

Day & Date: Thursday November 10, 2022

Time	P (bar)	Fault Free (0) Fault (1)	Time	P (bar)	Fault Free (0) Fault (1)
7:00	4.0	1	15:00	3.8	0
8:00	4.0	1	16:00	3.9	1
9:00	3.8	0	17:00	3.7	1
10:00	3.7	0	18:00	4.0	1
11:00	3.8	0	19:00	4.3	1
12:00	4.4	1	20:00	4.1	1
13:00	4.3	1	21:00	3.7	0
14:00	4.5	1	22:00	3.6	0
Inspector Name:	SABIR AHMED		Inspector Name:	Mohd Yonus	
Signature:			Signature:		

Appendix C



	Water Leaving Temprature (Celsius)	Inspection (Fault = 1, No fault = 0)
2	25.7	0
3	25.6	0
4	30.6	1
5	28.5	1
6	26.9	0
7	30.7	1
8	31.0	1

References

1. Kullu, O.; Cinar, E. A Deep-Learning-Based Multi-Modal Sensor Fusion Approach for Detection of Equipment Faults. *Machines* **2022**, *10*, 1105. [CrossRef]
2. Sahal, R.; Breslin, J.G.; Ali, M.I. Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. *J. Manuf. Syst.* **2020**, *54*, 138–151. [CrossRef]
3. Bousdekis, A.; Apostolou, D.; Mentzas, G. Predictive maintenance in the 4th industrial revolution: Benefits, business opportunities, and managerial implications. *IEEE Eng. Manag. Rev.* **2019**, *48*, 57–62. [CrossRef]
4. Cotrufo, N.; Saloux, E.; Hardy, J.M.; Candanedo, J.A.; Platon, R. A practical artificial intelligence-based approach for predictive control in commercial and institutional buildings. *Energy Build.* **2020**, *206*, 109563. [CrossRef]
5. Simon, L.; Rauffet, P.; Guérin, C.; Seguin, C. Trust in an autonomous agent for predictive maintenance: How agent transparency could impact compliance. In Proceedings of the 13th AHFE Conference (Applied Human Factors and Ergonomics), New York, NY, USA, 24–28 July 2022; HAL Open Science: Lyon, France, 2022; pp. 61–67.
6. Villa, V.; Bruno, G.; Aliev, K.; Piantanida, P.; Corneli, A.; Antonelli, D. Machine Learning Framework for the Sustainable Maintenance of Building Facilities. *Sustainability* **2022**, *14*, 681. [CrossRef]
7. Achouch, M.; Dimitrova, M.; Ziane, K.; Sattarpanah Karganroudi, S.; Dhouib, R.; Ibrahim, H.; Adda, M. On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges. *Appl. Sci.* **2022**, *12*, 8081. [CrossRef]
8. Almobarek, M.; Mendibil, K.; Alrashdan, A. Faults handling in chilled water system maintenance program. In Proceedings of the 12th International Conference on Industrial Engineering and Operations Management, Istanbul, Turkey, 7–10 March 2022; IEOM Society International: Southfield, MI, USA, 2022; pp. 1616–1625.
9. Almobarek, M.; Mendibil, K.; Alrshdan, A. Study of Factors Influencing Room Ambient Temperature Using Design of Experiments. In Proceedings of the 11th International Conference on Industrial Engineering and Operations Management, Singapore, Singapore, 7–11 March 2021; IEOM Society International: Southfield, MI, USA, 2021; pp. 1509–1516.
10. Li, D.; Zhou, Y.; Hu, G.; Spanos, C.J. Fusing system configuration information for building cooling plant Fault Detection and severity level identification. In Proceedings of the 2016 IEEE International Conference on Automation Science and Engineering (CASE), Fort Worth, TX, USA, 21–25 August 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1319–1325.
11. Yan, R.; Ma, Z.; Kokogiannakis, G.; Zhao, Y. A sensor fault detection strategy for air handling units using cluster analysis. *Autom. Constr.* **2016**, *70*, 77–88. [CrossRef]
12. Tehrani, M.M.; Beauregard, Y.; Rioux, M.; Kenne, J.P.; Ouellet, R. A Predictive Preference Model for Maintenance of a Heating Ventilating and Air Conditioning System. *IFAC-Pap* **2015**, *48*, 130–135. [CrossRef]
13. Rueda, E.; Tassou, S.A.; Grace, I.N. Fault detection and diagnosis in liquid chillers. *Proc. Inst. Mech. Eng. Part E J. Process. Mech. Eng.* **2005**, *219*, 117–125. [CrossRef]
14. Zhou, Q.; Wang, S.; Xiao, F. A novel strategy for the fault detection and diagnosis of centrifugal chiller systems. *HVACR Res.* **2009**, *15*, 57–75. [CrossRef]
15. Dudzik, M.; Romanska-Zapala, A.; Bomberg, M. A neural network for monitoring and characterization of buildings with Environmental Quality Management, Part 1: Verification under steady state conditions. *Energies* **2020**, *13*, 3469. [CrossRef]
16. Montazeri, A.; Kargar, S.M. Fault detection and diagnosis in air handling using data-driven methods. *J. Build. Eng.* **2020**, *31*, 101388. [CrossRef]
17. Liang, J.; Du, R. Model-based fault detection and diagnosis of HVAC systems using support vector machine method. *Int. J. Refrig.* **2007**, *30*, 1104–1114. [CrossRef]

18. Hu, R.; Granderson, J.; Auslander, D.; Agogino, A. Design of machine learning models with domain experts for automated sensor selection for energy fault detection. *Appl. Energy* **2018**, *235*, 117–128. [CrossRef]
19. Almobarek, M.; Mendibil, K.; Alrashdan, A. Predictive Maintenance 4.0 for Chilled Water System at Commercial Buildings: A Systematic Literature Review. *Buildings* **2022**, *12*, 1229. [CrossRef]
20. Almobarek, M.; Mendibil, K.; Alrashdan, A.; Mejjaouli, S. Fault Types and Frequencies in Predictive Maintenance 4.0 for Chilled Water System at Commercial Buildings: An Industry Survey. *Buildings* **2022**, *12*, 1995. [CrossRef]
21. Jebreen, I. Using inductive approach as research strategy in requirements engineering. *Int. J. Comput. Inf. Technol.* **2012**, *1*, 162–173.
22. Malachowski, A. *The New Pragmatism*, 1st ed.; Routledge: London, UK, 2014.
23. Bacon, M. *Pragmatism: An Introduction*; Polity Press: Cambridge, UK, 2012.
24. Sakib, N.; Wuest, T. Challenges and opportunities of condition-based predictive maintenance: A review. *Procedia CIRP* **2018**, *78*, 267–272. [CrossRef]
25. Sharma, H.; Kumar, S. A survey on decision tree algorithms of classification in data mining. *Int. J. Sci. Res.* **2016**, *5*, 2094–2097.
26. Fletcher, S.; Islam, M.Z. Decision tree classification with differential privacy: A survey. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 1–33. [CrossRef]
27. Wildenauer, A.; Mbabu, A.; Underwood, J.; Basl, J. Building-as-a-Service: Theoretical Foundations and Conceptual Framework. *Buildings* **2022**, *12*, 1594. [CrossRef]
28. Ellis, G. What Are as Built Drawings? *Digital Builder Blog, Autodesk Construction Cloud.* 2021. Available online: https://constructionblog.autodesk.com/as-built-drawings/ (accessed on 2 December 2022).
29. ASHRAE Handbook. Available online: www.ashrae.org (accessed on 5 December 2022).
30. Lam, H.F.; Yang, J.H.; Hu, Q. How to install sensors for structural model updating? *Procedia Eng.* **2011**, *14*, 450–459. [CrossRef]
31. Beckmann, C.; Consolvo, S.; LaMarca, A. Some Assembly Required: Supporting End-User Sensor Installation in Domestic Ubiquitous Computing Environments. In *Ubiquitous Computing*; Davies, N., Mynatt, E.D., Siio, I., Eds.; Springer: Berlin/Heidelberg, Germany, 2004.
32. Kayastha, N.; Niyato, D.; Hossain, E.; Han, Z. Smart grid sensor data collection, communication, and networking: A tutorial. *Wirel. Commun. Mob. Comput.* **2014**, *14*, 1055–1087. [CrossRef]
33. Javed Mehedi Shamrat, F.M.; Ranjan, R.; Hasib, K.M.; Yadav, A.; Siddique, A.H. Performance Evaluation Among ID3, C4.5, and CART Decision Tree Algorithm. In *Pervasive Computing and Social Networking*; Ranganathan, G., Bestak, R., Palanisamy, R., Rocha, Á., Eds.; Springer: Singapore, 2022.
34. Grąbczewski, K. *Meta-Learning in Decision Tree Induction*; Springer: New York, NY, USA, 2014.
35. Guttag, J. *Introduction to Computation and Programming Using Python: With Applications to Understanding Data*, 2nd ed.; The MIT Press: Cambridge, MA, USA, 2017.
36. Almobarek, M.; Alrashdan, A. Water budget control using DMAIC in commercial buildings. *Int. J. Six Sigma Compet. Advant.* **2022**, *14*, 86–208. [CrossRef]