

Article

Development of a Hybrid Support Vector Machine with Grey Wolf Optimization Algorithm for Detection of the Solar Power Plants Anomalies

Qais Ibrahim Ahmed ¹, Hani Attar ^{2,*} , Ayman Amer ², Mohanad A. Deif ³  and Ahmed A. A. Solyman ⁴

¹ Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Istanbul Gelisim University, Istanbul 34310, Turkey

² Department of Energy Engineer, Zarqa University, Zarqa 13133, Jordan

³ Department of Bioelectronics, Modern University of Technology and Information (MTI) University, Cairo 11728, Egypt

⁴ Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Nişantaşı University, Istanbul 34398, Turkey

* Correspondence: hattar@zu.edu.jo

Abstract: Solar energy utilization in the industry has grown substantially, resulting in heightened recognition of renewable energy sources from power plants and intelligent grid systems. One of the most important challenges in the solar energy field is detecting anomalies in photovoltaic systems. This paper aims to address this by using various machine learning algorithms and regression models to identify internal and external abnormalities in PV components. The goal is to determine which models can most accurately distinguish between normal and abnormal behavior of PV systems. Three different approaches have been investigated for detecting anomalies in solar power plants in India. The first model is based on a physical model, the second on a support vector machine (SVM) regression model, and the third on an SVM classification model. Grey wolf optimizer was used for tuning the hyper model for all models. Our findings will clarify that the SVM classification model is the best model for anomaly identification in solar power plants by classifying inverter states into two categories (normal and fault).

Keywords: solar energy; intelligent grid system; power plant anomalies; PV



Citation: Ahmed, Q.I.; Attar, H.; Amer, A.; Deif, M.A.; Solyman, A.A.A. Development of a Hybrid Support Vector Machine with Grey Wolf Optimization Algorithm for Detection of the Solar Power Plants Anomalies. *Systems* **2023**, *11*, 237. <https://doi.org/10.3390/systems11050237>

Academic Editors: Paolo Visconti and Vladimír Bureš

Received: 15 February 2023

Revised: 29 April 2023

Accepted: 30 April 2023

Published: 8 May 2023



Copyright: © 2023 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/>).

1. Introduction

In recent years, there has been rapid growth in renewable energy, including power plants, which is expected to improve the production of clean, low-cost energy and drive economic growth [1]. Concerns about global warming and well-being objectives enabled the definition of fundamental energy goals to be realized [2], such as the nearly exclusive usage of renewable energy sources (RESs) and, especially, the expansion of local RESs exploitation [3]. The target of a 100% renewable energy production landscape cannot ignore adaptable city policies to improve their worldwide competitiveness and sustainability [4,5].

In addition, the World Health Organization and other nations have developed pertinent policies to address environmental issues that affect all humans, decrease emissions of greenhouse gases, and prevent the catastrophic effects of climate change [6]. Moreover, the increasing demand for attention to environmental challenges has become a crucial indicator in the energy sector. The words “carbon neutral” and “carbon peak” are crucial environmental and energy measures to deal with the global warming crisis. Solar photovoltaic (PV) power generation, which represents renewable energy, is unquestionably the driving force behind the “double carbon” goal [7,8].

Notwithstanding the overall negative effects of the COVID-19 epidemic and many price rises impacting the materials and techniques involved in manufacturing PV modules, PV has continued to see amazing growth in 2020 and the first part of 2021. As a result,

worldwide cumulative solar capacity exceeded a terawatt (TW) in 2020 [8]. In addition, solar power generation once again had the highest net installed capacity during this year.

PV systems often experience a range of abnormalities that can affect their performance [9,10], including internal and external faults. Internal faults can produce zero power during daylight hours, including component failure, shading, inverter shutdown, system isolation, and inverter maximum power point issues. External factors can hinder power generation even if not caused by the PV system. Dust, shading, temperature, and humidity are the most influential exterior factors in PV system manufacture.

Identifying PV panel faults play a crucial role in the growth of the PV sector and is crucial for advancing the advancement of PV energy [11]. The intelligent identification of PV panel defects is becoming a workable and promising option as artificial intelligence advances. Developing intelligent PV inspection systems now requires machine vision methods to detect surface flaws in PV panels [8,12]. Motivated by the accomplishments of deep learning in data mining [13], computer vision [14,15], and speech processing, deep learning algorithms have the potential to significantly improve detection accuracy, provide solutions for competent PV power plant inspection, and guide the maintenance and operation of power plants [16].

Monitoring tools are needed to operate photovoltaic (PV) systems effectively, as they can help identify and address issues affecting the system's performance and safety. Online monitoring can provide plant operators with important information for managing and integrating the plant into an intelligent grid. Failure to detect problems with PV arrays can lead to reduced power generation and fire hazards. Early detection of anomalies on solar panels can help prevent power loss and ensure the panels' continued performance and safety. Therefore, it is essential to have efficient and accurate methods for detecting anomalies in PV systems. Some of this paper's significant contributions include the following:

1. This study will investigate various anomaly detection models and conduct comparison tests to evaluate their precision and performance with optimized hyperparameters.
2. This study aims to identify and categorize external and internal factors (AC and DC powers were an example of internal factors that may result in anomalies, while the ambient, module, and irradiation temperatures were examples of external factors) that affect anomalies in PV power plants.
3. The impact of external and internal factors on model accuracy and the correlation between these factors and anomaly detection.

The remainder of this paper is structured as follows: Related work is provided in Section 2, Background information and methodology are covered in Section 3, experimental findings and parameter optimization are shown in Section 4, and the conclusion is provided in Section 5. We summarize our results and make recommendations for further research in the decision.

2. Related Work

Studies have examined various methods for spotting irregularities in photovoltaic (PV) power networks. To monitor flaws in panel modules, S. Aouat et al. [17] employed grayscale cogeneration matrices to extract picture characteristics produced by infrared imaging methods. Support vector machines (SVM) and random forest algorithms (RBF kernel) were employed to build detection models and achieve the necessary detection accuracy in the electroluminescence dataset [18]. Jumaboev et al. used numerous deep-learning models to confirm the viability of deep-learning approaches in solar inspection [19]. F. Lu et al. [20] suggested a semi-supervised anomaly detection approach based on adversarial generative networks to identify defects in PV panels. Using texture analysis and supervised learning, an automated technique for detecting optoelectronic components was used to process infrared pictures [15]. For processing and fault detection of solar panel thermographic sequences, Chiwu Bu et al. [16] employed supervised learning methods for quadratic discriminant analysis (QDA) and linear discriminant analysis (LDA).

One study used a long short-term memory (LSTM) NN algorithm to predict the yield of solar stills, which can recall patterns and anticipate long-term time-series behaviors [21]. Another study provided methods based on artificial intelligence for predicting the water output of solar distillers, including an LSTM model with a moth-flame optimizer [22]. Compared to the solitary LSTM model, the optimized LSTM model outperformed it.

Convolutional LSTM (ConvLSTM), an effective hybrid model that combines LSTM and a convolutional neural network (CNN), was proposed in [23,24]. It exhibited accurate prediction results with reduced latency, hidden neurons, and computational complexity. The deep learning techniques have many applications in various industries, such as data mining, medicine, agriculture, and wind and solar energy production, which has been examined in other recent research [25,26].

For instance, P. Branco et al. [27] investigated different techniques for locating and categorizing abnormalities in PV systems, including the k-nearest-neighbors classification, support vector machines, neural networks, and auto-regressive integrated moving average model (ARIMA) [28].

M. De Benedetti et al. [29] developed a strategy for detecting and predicting abnormalities in PV systems and performing maintenance. The model is based on an artificial neural network (ANN) that predicts AC power generation using solar irradiance and temperature data from PV panels. K. Natarajan et al. [30] present a new method for identifying faults using thermal image processing and an SVM tool that classifies characteristics as non-defective or defective.

A model-based method for identifying irregularities in the DC section of PV plants and brief shadowing [31]. The process entails creating a model based on the one-diode model to describe the typical PV system behavior and produce residuals for defect identification. The residuals are then subjected to a one-class support vector machine (OCSVM) in order to detect errors. Sundown has been described as a sensorless technique for identifying individual panel failures in solar arrays [32]. Sundown employs a model-driven methodology to find departures from predicted behavior by analyzing interactions between the power produced by nearby solar panels. The model can handle multiple issues across various panels simultaneously and categorizes anomalies to determine probable causes, such as snow accumulation, leaf accumulation, dirt buildup, and electrical malfunctions. A brand-new tool named ISDIPV can locate and examine problems in PV solar power systems [33]. It comprises three primary parts data gathering, anomaly detection, and performance deviation diagnostics.

Y. Zhao et al. [34] used multilayer perceptron neural network models and linear transfer functions (LTF) to model average performance. In order to effectively detect and categorize anomalies in PV systems, the study suggested a data-driven method employing PV string currents as indicators. The proposed method for anomaly detection featured two steps: local context-aware detection (LCAD) and global context-aware anomaly detection, and it leveraged unsupervised machine learning techniques (GCAD).

J. Mulongo et al. [35] studied generators as power sources for TeleInfra base stations, and anomalies in fuel consumption were identified using reported data. Four classification approaches—k-nearest neighbors (KNN), multilayer perceptron (MLP), logistic regression (LR), and SVM—were used to identify anomalies in gasoline consumption patterns by learning the practices using pattern recognition. The findings demonstrated that MLP was the most successful measuring and interpretation method.

Using KNN and OCSVM for anomaly detection, a unique method of PV system monitoring is introduced [36]. These self-learning algorithms greatly minimize measurement labor while enhancing the accuracy of defect monitoring. In connection [37], a multilayer perceptron and the k-nearest-neighbors method are used to analyze data from a DC sensor and distinguish between different electrical current characteristics. A sensorless fault detection method in PV plants is proposed in [38] based on the sharp fall in current between two MPPT sampling points.

Simulations demonstrated that anomalies could be detected in diverse conditions, regardless of brightness and irradiance levels. J. Balzategui et al. [39] proposed a framework for anomaly detection in monocrystalline solar cells. The system is comprised of two parts. The first part involves constructing an anomaly detection model based on the generative adversarial network (GAN) that can detect abnormal patterns using only ideal training examples. The generated features are then used to train a fully convolutional network for supervised learning. Q. Wang et al. [40] introduced a method for real-time analysis of raw video streams obtained from aerial thermography. They combine statistical machine learning and image processing techniques to identify and contain abnormalities in PV images using resilient principal component analysis (RPCA). The method also includes post-processing techniques for image noise reduction and segmentation. Energy yield data are evaluated using a variety of models [41], and it is discovered that proximity-based models, linear models, anomalous ensembles, probabilistic models, and NN have the highest detection rates.

Authors in [42] introduce SolarClique, a data-driven method for identifying anomalies in the electricity generation of solar power facilities, which does not require sensor equipment for fault/anomaly detection and relies only on the array's output and the output of adjacent arrays for operational anomaly detection. An anomaly detection method using a semi-supervised learning model was suggested in [43] to predict solar panel settings to avoid situations where the solar panel cannot produce electricity due to equipment degradation. This method uses a clustering model to filter normal behaviors and an Autoencoder model for neural network classification.

J. Pereira et al. [44] present a comprehensive, unsupervised, and scalable approach for detecting anomalies in offline and real-time time-series data. They employed a variational autoencoder model with an encoder and decoder parameterized by recurrent neural networks to capture the temporal dependencies in the data. The results demonstrate that the model effectively identifies unusual configurations using probabilistic restoration metrics such as anomaly scores. Reference [45] describes a unique ensemble anomaly detection method for detecting cyber-physical attacks using nonlinear regression models and anomaly scores in intelligent grids.

B. Rossi et al. [46] implemented an unsupervised, contextual, and collective detection approach to monitoring the data flow of a large energy dispenser in the Czech Republic. It detects anomalies using clustering silhouette thresholding combined with typical item-set mining and category clustering algorithms. A. Toshniwal et al. [47] overview various anomaly detection methods, including graph analysis, nearest neighbor, clustering, and statistical analysis, spectral, and information-theoretic ways. The choice of the best AD algorithm depends on factors such as the input data, the nature of anomalies, the desired output data, and domain-specific knowledge.

3. Materials and Methods

3.1. Theoretical Background

A. Support Vector Machine

The supervised learning algorithm, SVMs, can be applied to classification or regression applications. They perform regressions with the slightest error or locate the hyperplane in a high-dimensional space that maximum separates various classes. When the number of features is substantially higher than the number of samples, and the data is noise-free or has minimal noise, SVMs are extremely helpful. They work best when the classes are well-separated and in high-dimensional spaces. SVMs can be employed with various kernel functions, which enables them to adapt to complex nonlinear relationships in the data, which is one advantage of SVMs. The "kernel trick" can also be utilized to handle high-dimensional data. The algorithm can function in a higher-dimensional feature space without explicitly calculating the data coordinates in that space.

Vapnik and his colleagues presented the support vector regression (SMV) method [48], and the development route was established by expanding the SVM classification

algorithm [49]. As a supervised machine learning technology, SMV permits the prediction of continuous real-valued variables and is a machine learning regression technique for regression analysis [50].

B. Support Vector Machine Regression

The model is first trained using a sample data set in the support vector machine method and then used to make predictions. Let the sample data be (x_i, y_i) , where $i = 1, 2, \dots, n$. Here, x_i represents the input factor that influences the value of the cyber safety situation, and y_i represents the predicted situation value. The regression function can be expressed as [51]:

$$y_i = \omega \cdot x_i + b \tag{1}$$

The formula incorporates the weighting matrix ω , the bias direction b , the input vector x_i , and the predicted regression value y_i . It also uses an ϵ -insensitive loss function.

$$|y - y_i|_\epsilon = \begin{cases} 0, & |y - y_i| \leq \epsilon \\ |y - y_i| - \epsilon, & \text{other} \end{cases} \tag{2}$$

If it is permissible for the fitting error to exceed ϵ , the constraint in Equation (1) can be modified by adding a relaxation factor, resulting in the following:

$$\begin{aligned} \min \frac{1}{2} &= \|\omega\|^2 + \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{s.t. } &y_i - \omega \cdot x_i - b \leq \epsilon + \xi_i^* \\ &\omega \cdot x_i + b - y_i \leq \epsilon + \xi_i \\ &\xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, n \end{aligned} \tag{3}$$

In the formula, C is the penalty factor, and ξ_i and ξ_i^* are the relaxation variables. By taking the derivative of the optimization problem given in Equation (3), we can obtain the dual problem, as shown below:

$$\begin{aligned} \min W(a, a^*) &= -\epsilon \sum_{i=1}^l (a_i^* + a_i) + \sum_{i=1}^l y_i (a_i^* - a_i) \\ &\quad - \frac{1}{2} \sum_{ij=1}^l (a_i^* - a_i)(a_j^* - a_j)K(x_i, x_j) \\ \text{s.t. } &\sum_{i=1}^l a_i^* = \sum_{i=1}^l a_i \\ &0 \leq \alpha_i^* \leq C, 0 \leq \alpha_i \leq C, i = 1, 2, \dots, l \end{aligned} \tag{4}$$

In the formula, $K(x_i, x_j)$ is the kernel function, and α_i, α_i^* are Lagrange multipliers. The regression function that results is:

$$y_i = \sum_{i=1}^n (a_i^* - a_i)K(x_i, x_j) + b \tag{5}$$

C. Support Vector Machine Classification

In this section, a brief overview of SVM classification is provided. SVMs are mainly used for binary variety, meaning they can differentiate between two classes. Based on statistical learning theory, the fundamental objective of SVM classification is to identify a hyperplane that acts as a decision boundary and effectively separates the positive (+1) and negative (−1) classes.

Consider the task of classifying a set of training data (x_i, y_i) , where $i = 1, 2, \dots, m$, into two classes. The feature vector x_i is in the R^n space and has n dimensions, while the class label y_i belongs to the set $\in \{+1, -1\}$. The objective of the generalized linear SVM is to

find the optimal decision boundary, represented by the hyperplane $f(x) = \langle w \times x \rangle + b$, by solving the following optimization equation:

$$\text{Minimize}_{w,b} \frac{1}{2} \|w\|^2 \text{ subject to } y_i(w \cdot x_i + b) - 1 \geq 0 \quad (6)$$

SVMs aim to find the hyperplane in a high-dimensional space that maximally separates the positive and negative examples by solving a constrained optimization problem. The optimal orientation and distance of the hyperplane are determined by the weight vector w and the bias term b , respectively. To minimize the magnitude of w and maintain the minimum distance between the hyperplane and the nearest examples (the margin), the Lagrange multipliers $\lambda_i (i = 1, 2, \dots, m)$, are used to enforce the constraint, where m is the number of examples. The hyperplane that results from this process has the lowest training errors.

$$f(x) = \text{sign} \left(\sum_{i=1}^m y_i \lambda_i \langle x_i \times x_j \rangle + b^* \right) \quad (7)$$

The function $\text{sign}(\cdot)$ is indicated, and the x_i vectors for which $\lambda_i > 0$ are referred to as support vectors. If linear equations cannot define the hyperplane, the nonlinear mapping function $\phi(x)$ can express the data x in a higher-dimensional Euclidean space H . The optimization problem in space H can then be solved by replacing $\langle x_i \times x_j \rangle$ with $\langle \phi(x_i) \times \phi(x_j) \rangle$, using a kernel function [14]. The frequently employed kernel functions are sigmoid, radial basis function (RBF), polynomial, and linear. The RBF is often used in SVM classification due to its exceptional nonlinear classification abilities. The following is the equation for the RBF:

$$k(x_i, x_j) = \exp \left(-\frac{\|x_i - x_j\|^2}{\sigma^2} \right) \quad (8)$$

The nonlinear SVM classifier can take the following forms after including the kernel function:

$$f(x) = \text{sign} \left(\sum_{i=1}^m y_i \lambda_i k(x_i, x_j) + b \right) \quad (9)$$

D. Grey Wolf Optimizer (GWO)

The grey wolf optimization (GWO) algorithm is a novel meta-heuristic group intelligence optimization technique introduced in 2014 by Mirjalili et al. [49,52,53]. Grey wolves' natural leadership structure and hunting strategy influenced the meta-heuristic optimization technique known as grey wolf optimization (GWO) [54]. The algorithm resembles the wolf pack's leadership structure, in which the alpha wolves take the lead, the beta wolves assist, and the delta wolves search for food. Each wolf in the pack is assigned a position and is rewarded based on their performance. GWO is based on these principles and uses a population of candidate solutions, called wolves, to search for the optimal solution. The wolves are assigned different positions in a pack and their reward is based on their position in the hierarchy. The GWO algorithm is used to solve optimization problems with a large number of variables and complex constraints. It is a fast, reliable, and efficient optimization algorithm that can be applied to various types of optimization problems.

This method optimizes the search by simulating grey wolves' predatory behavior, including tracking, encirclement, chase, and assault. It has a rigid social structure where the top three wolves are regarded as the best, and the bottom three are referred to as ω . These wolves are positioned around the best wolves (α , β , and δ). Further definitions related to the GWO algorithm are provided next.

1. The distance separating a grey wolf and its quarry:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_q(t) - \vec{X}(t) \right| \quad (10)$$

At each iteration t , the value of \vec{C} is set to 2 times the random number \vec{r}_1 (where \vec{r}_1 is between 0 and 1). \vec{X}_q stands for the position vector of the quarry and \vec{X} represents the position vector of the grey wolf.

2. Gray wolf location update:

$$\vec{X}(t+1) = \vec{X}_q(t) - \vec{A} \cdot \vec{D} \quad (11)$$

The formula $\vec{A} = 2\vec{a} \cdot \vec{r}_2 - \vec{a}$, where \vec{a} assigns a value that is determined by the convergence gene, \vec{a} , which decreases from 2 to 0, and a randomly generated number, r_2 in the range of [0,1].

3. Prey position positioning:

$$\vec{D}_\alpha = \left| \vec{C}_1 \cdot \vec{X}_\alpha - \vec{X} \right|, \vec{D}_\beta = \left| \vec{C}_2 \cdot \vec{X}_\beta - \vec{X} \right|, \vec{D}_\delta = \left| \vec{C}_3 \cdot \vec{X}_\delta - \vec{X} \right| \quad (12)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \left(\vec{D}_\alpha \right), \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \left(\vec{D}_\beta \right), \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \left(\vec{D}_\delta \right) \quad (13)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (14)$$

The locations of α , β , and δ are represented by \vec{X}_α , \vec{X}_β , and \vec{X}_δ , respectively. \vec{C}_1 , \vec{C}_2 , and \vec{C}_3 are random vectors, and \vec{X} is the location of the current solution.

The key stages in the grey wolf optimization process include:

Step 1: Initialization of the population, where each member of the population represents a solution to the optimization problem, a , A , C , \vec{X}_α , \vec{X}_β , \vec{X}_δ , and \vec{X}_δ values.

Step 2: Calculate the fitness value for each individual in the population.

Step 3: To identify the suboptimal solution, the optimal solution, and the current optimal solution, compare the fitness values of intelligent individuals with the fitness values of \vec{X}_α , \vec{X}_β and \vec{X}_δ .

Step 4: Compute the values of a , A , and C .

Step 5: Update the current position of each intelligent individual according to Equation (12).

Step 6: If the maximum number of iterations is met, it will return to Step 2; otherwise, it will terminate.

3.2. Materials

Data from solar power plants in India (near Gandikota, Andhra Pradesh) were gathered over 34 days at 15-min intervals. Twenty-two inverter sensors were connected to both inverters and plants to monitor the generation rate (an internal factor that may lead to anomalies). The inverters monitored meteorological conditions at the plant level (an external factor that can produce anomalies). The description of the variables employed in the study has shown in Table 1. These data are publicly available, licensed, and accessible in accordance with [55].

Table 1. Explanation of the variables employed in the study.

Variable Type	Variable Name	Variable Abbreviation (Unit)	Variable Description
Internal factor	DC power	Power_DC (kW)	The DC power produced by the inverter
	AC power	Power_AC (kW)	The AC power produced by the inverter
	Total yield	Total_power_DC (kW)	The total DC power output from the inverter over a period of time.
External factors	Solar irradiance	IRR (kW/m ²)	The intensity of the electromagnetic radiation emitted by the sun per unit area
	Ambient temperature	Amb_Temp (°C)	The temperature around the solar power plant
	Solar panel temperature	Module_Temp (°C)	The temperature indication for the solar module is measured by attaching a sensor to the panel.

3.3. Methodology

This research investigates three different approaches for detecting anomalies in solar power facilities. The first methodology is based on a physical model, the second on a hybrid GWO SVM regression model, and the third on a hybrid GWO SVM classification model. The overall process is represented in Figure 1 and involves the following steps: data preparation, feature selection, prediction phase, and performance evaluation. These are explained in more detail in the subsequent sections.

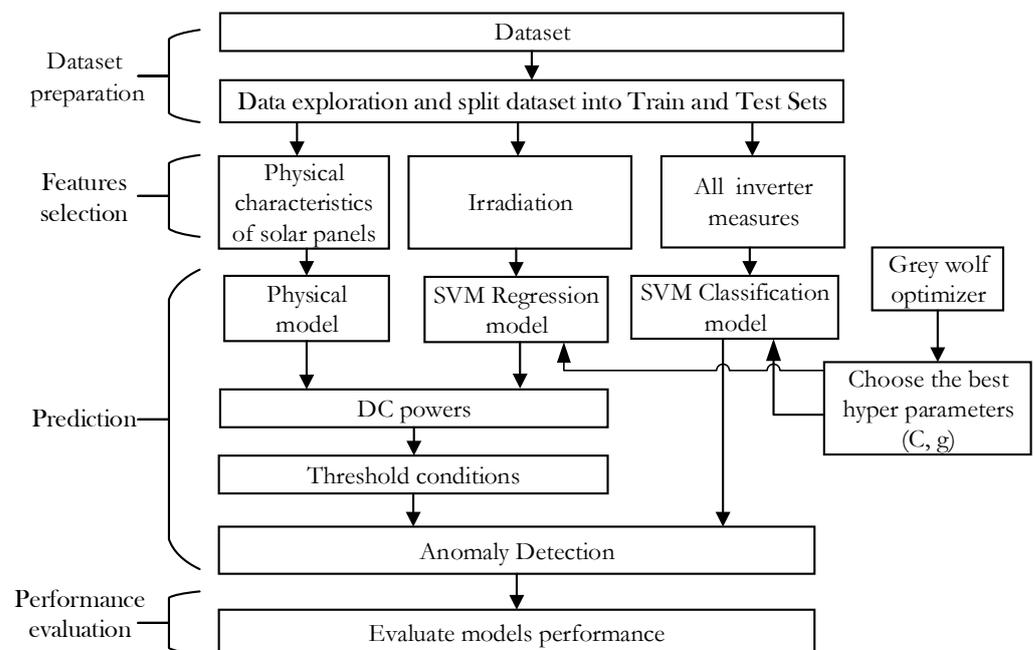


Figure 1. Summary of the proposed methodology steps.

3.3.1. Data Preparation

The main dataset was preprocessed by combining the power generation and weather data into a single dataset, removing any measures with missing values, and looking at the

scale of the variables. Then, 80% of the dataset was picked randomly to train the prediction model, and the other 20% was used to test the model's accuracy.

3.3.2. Features Selection

The main objective of feature selection is to improve the performance of a predictive model by utilizing only pertinent data and diminishing the computational cost of modeling by decreasing the input variable to be modeled. Spearman's rank correlation (ρ) was utilized to quantify the correlation rank between variables using the following equation:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (15)$$

The formula for d_i the difference between the two ranks for each observation is based on n being the total number of observations.

3.3.3. Prediction Phase

Physical Model

According to Hooda et al. (2018), the nonlinear equation can be used to model the Power_{DC} output of a photovoltaic cell.

$$P(t) = a * E(t)(1 - b * (T(t) + (E(t)/800) * (c - 20) - 25) - d * \ln(E(t))) \quad (16)$$

where $P(t)$ is Power_{DC}, $E(t)$ is irradiance, Temperature $T(t)$, and coefficients a, b, c, d [13].

GWO-SVM Classification/Regression Model

GWO and SVM were combined to create GWO-SVM C and GWO-SVM R, which were used to predict Power_{DC}. While support vector regression models usually perform well in linear and nonlinear relationships, SMV accuracy depends on properly selecting parameters C (penalty term) and g (kernel width). These two factors are known to have a vast range of variations and significantly influence the accuracy of SMV. As there is no set procedure for choosing these values, finding the proper parameters can be computationally demanding and seen as an optimization problem. To tackle this problem, we have used the hybrid optimization technique GWO. The GWO method optimizes the SVM regression prediction procedure by performing the following:

Step 1 determines the independent and dependent variables according to the model's assumptions and provides data for the support vector machine's training and testing sets.

Step 2: Initialize the values of g and c as components of input for the GWO algorithm and determine the values of a, A , and C ;

Step 3: Each intelligent individual location carrier must have the letters c and g to initialize the grey wolf population;

Step 4: Determine each grey wolf's fitness value by learning the training set's data using the SVM's initial c and g values.

Step 5: The grey wolf group is classified into four levels based on their fitness value: a, b, d , and x .

Step 6: Update the position of each grey wolf according to Equation (12) of the algorithm.

Step 7: The GWO algorithm then modifies each intelligent person's position following the fitness value, keeping the location with the highest fitness value;

Step 8: The training of the model is completed, and the optimal values of c and g are produced once the iteration count reaches the iteration maximum number, known as max iteration;

Step 9: Constructing the SVM regression forecasting model involves using the optimal c and g values from the grey wolf approach described earlier. The model's performance is then evaluated, and prediction is carried out using the pre-partitioned test data set. A flowchart illustrating the design steps of the algorithm can be found in Figure 2, which

encompasses the selection of the best parameters, g and c , for the SVM using the grey wolf method.

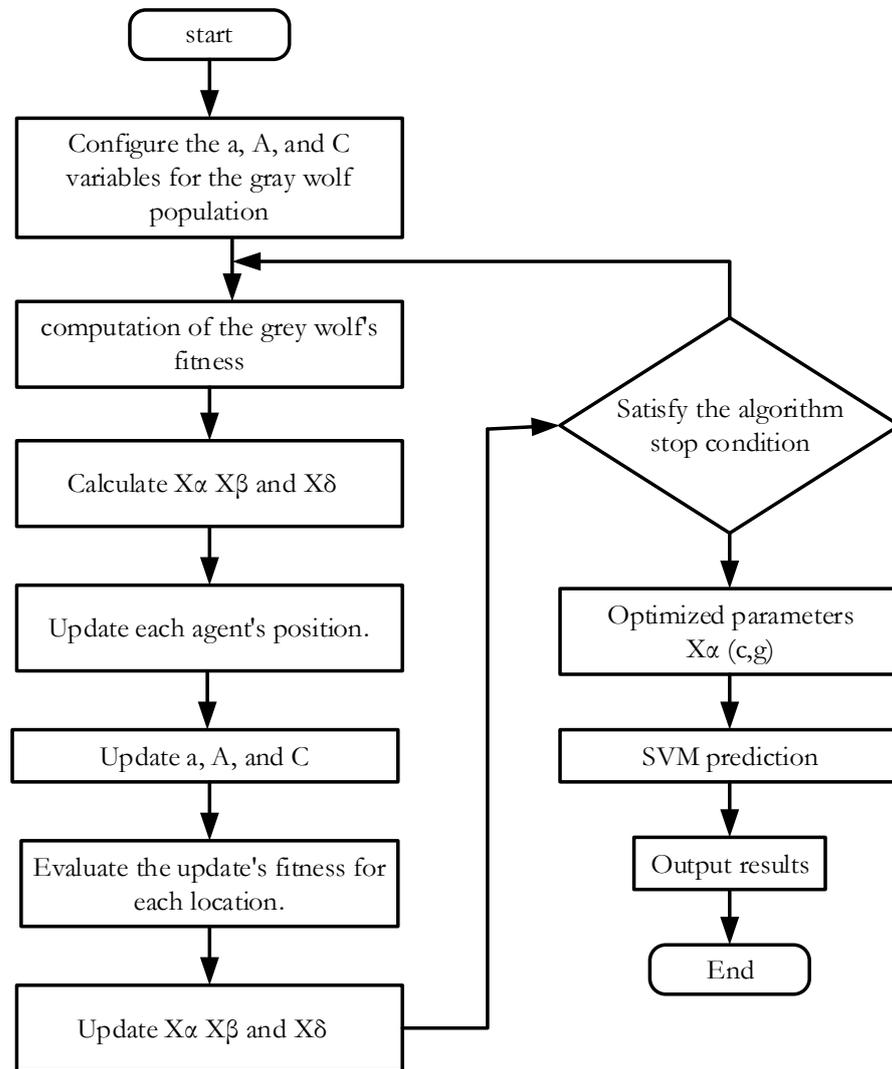


Figure 2. The proposed architecture of the GWO optimizer with SMV.

3.3.4. Anomaly Prediction Decision for Models

If the value of $P(t)$ derived from the GWO-SVM R and physical models is over the threshold, it indicates that the inverter is functioning normally; however, if the value is below the threshold, it means that the inverter has failed. The GWO-SVM C model further distinguishes the two categories, classifying solar panels into three categories (normal, defective, and marginal) based on the $P(t)$ value.

$$P(t) = \begin{cases} Normal, & P(t) \geq \theta \\ Fault, & P(t) < \theta \end{cases} \quad (17)$$

In our experiment, we tested a range of threshold levels. The optimal limit was 215 kW.

3.3.5. Performance Evaluation

The effectiveness of the physical and GWO-SVM R models developed in this article is evaluated using root mean square error (MSE) metrics.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2} \quad (18)$$

The formula uses the sample size, denoted by n , and the projected value, denoted by x_i , and the actual scenario value, denoted by y_i .

The classification accuracy of the GWO-SVM_C model can be calculated using Equations (19)–(21).

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{19}$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \tag{20}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{21}$$

The numbers TP (true positive) and TN (true negative) reflect the number of the inverter’s properly identified normal/fault states, whereas FN (false negative) and FP (false positive) stand for the number of the inverter’s falsely identified Normal/Fault states.

4. Experimental and Results

This research aims to produce a prediction model to classify the internal and external elements that lead to PV power plant faults. To do this, we conducted three experiments. First, the data characteristics’ correlation were studied to find the most important factors for training our predictive models. Second, various optimization models were tested to establish the best hyperparameters for the proposed prediction models. Third, three methodologies were employed for identifying anomalies in PV plants. Finally, the most accurate prediction model was used to determine which days and inverters had the highest failure rate.

The correlation between the input factors and the solar panel output power DC is illustrated in Figure 3 using the correlation coefficients. The correlation coefficients range from -1 to 1 and are a normalized covariance representation. A negative correlation of -1 indicates that a rise in one variable results in a decrease in the other. A positive correlation of 1 shows that a rise in one variable leads to an increase in the other. The diagonal values, which represent the correlation between a variable and itself (autocorrelation), are all equal to 1 , as a variable is perfectly correlated with itself.

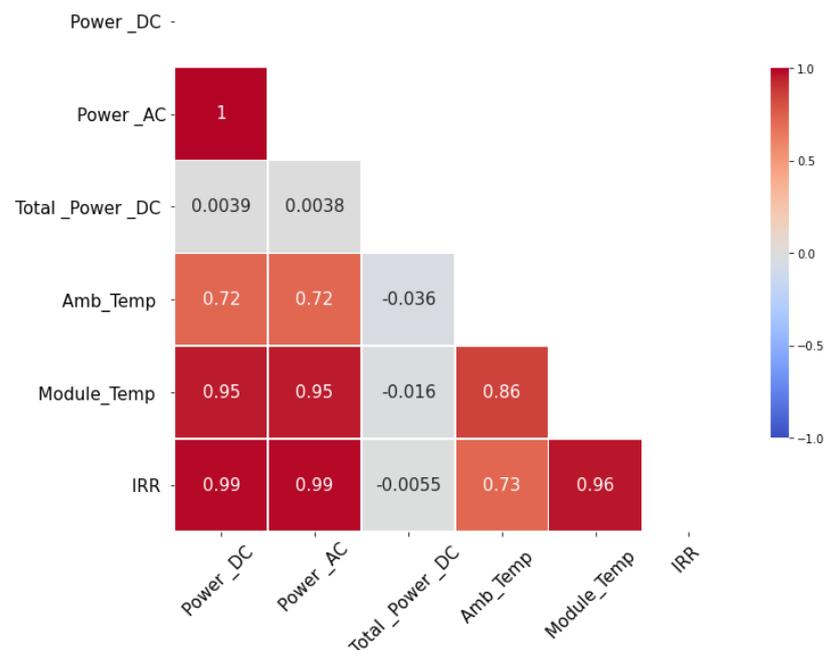


Figure 3. Correlation matrix for dataset variables.

A clear correlation between power DC output and power AC inputs, module temperature, and irradiation can be seen. In contrast, there is less of a connection between daily yield, total yield, ambient temperature, and power DC. The inputs of our experiment are power AC, module temperature, and irradiation, while daily yield, total yield, and ambient temperature are treated as outputs.

As seen in Figure 4, the variables of our dataset have varied scales due to an error in the prediction phase. The datasets were denoised, and feature scaling was performed using conventional scaling approaches to ensure the correctness of our models. Consequently, the data’s mean value was set to 0 and its variance value to 1. The formula for this is feature scaling: $\text{mean value of a feature} - (\text{original value} / \text{standard deviation})$.

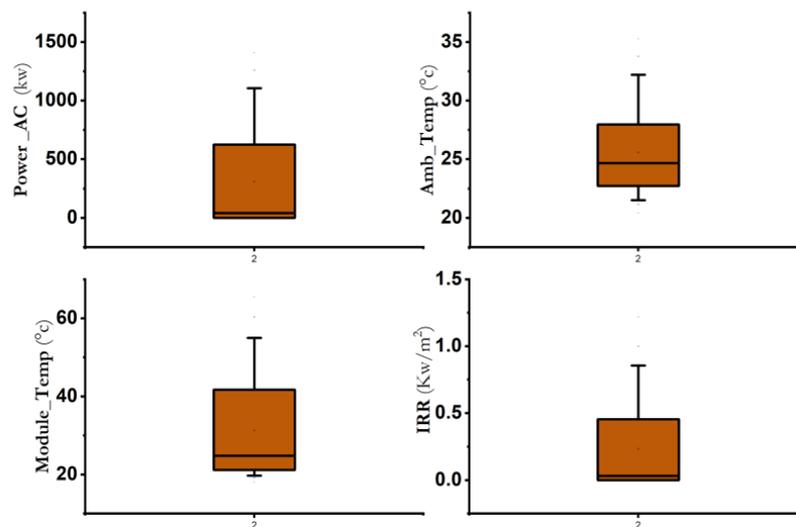


Figure 4. Box plots for selected input variables.

Figure 5’s scatter plot of DC power and irradiance reveals outliers, demonstrating that some inverters did not receive DC power despite enough sunlight to generate electricity. This highlights the efficiency of our solar panel lines in converting sunlight to DC power.

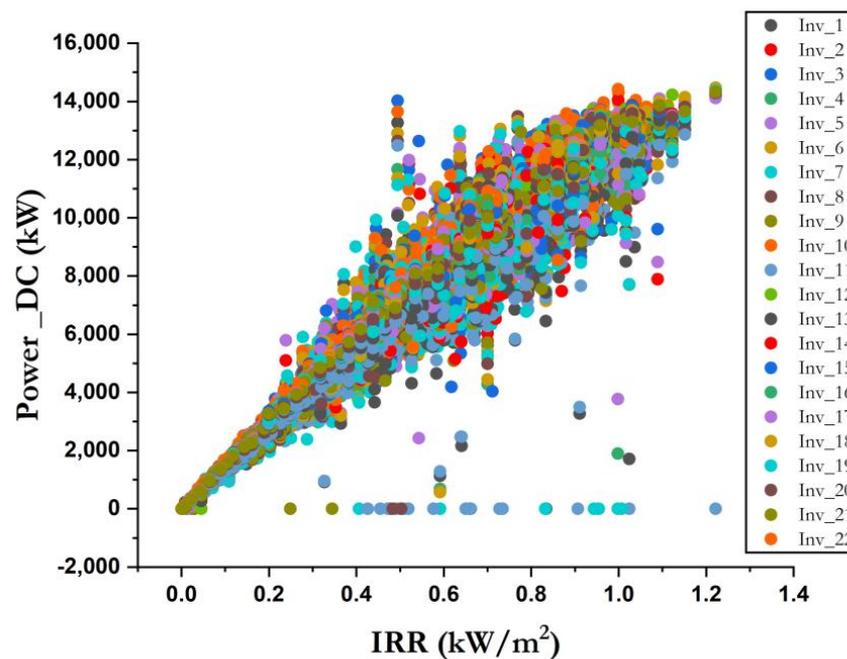


Figure 5. Distribution of Power_{DC} and IRR for each inverter.

With the addition of inverter states, Figure 6’s scatter plot between DC power and irradiance has been created to identify any anomalies in the solar energy-to-electricity conversion that would indicate faulty photovoltaic panel lines. Equipment failure can be assumed if no power is detected at the inverter during normal daylight operation.

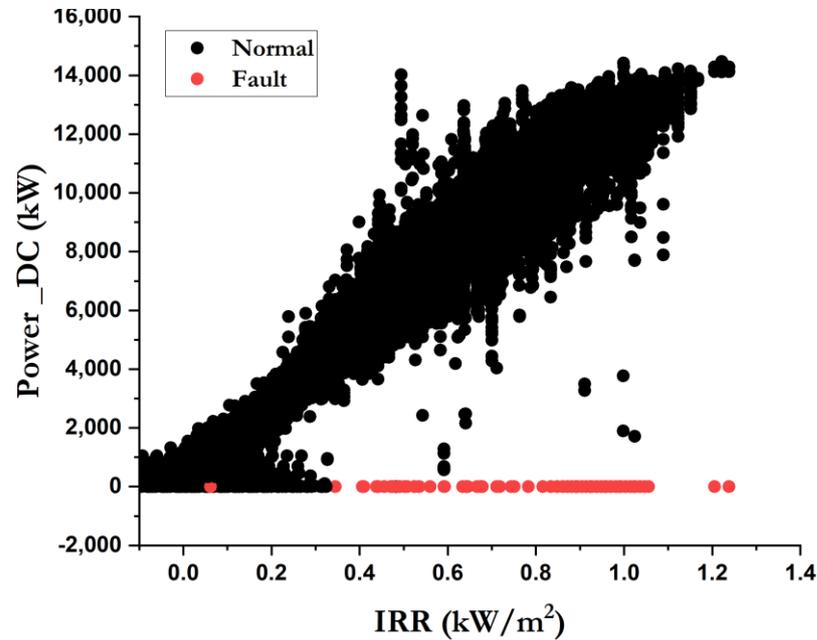


Figure 6. Distribution of Power_DC and IRR for each inverter after determining inverter status.

Before deploying SVM classification and regression models to forecast the amount of DC power generated by the inverter for solar power plants, the GWO method is utilized to optimize the SVM model hyperparameters to improve the prediction accuracy. To demonstrate the advantages of the GWO algorithm, two benchmark optimizer models have been created for comparison: grid search (GS) and random search (RS). The philosophy behind these optimizers and the rationale for their selection are discussed in reference [42].

The GWO-SVM classification model uses all dataset attributes to predict power DC. On the other hand, the physical model only employs irradiation and module temp. Table 2 displays the initial parameters for the GWO algorithm, and the SMV model was created utilizing the Jupiter Programming Language, which supports Python. Irradiation was applied in the GWO-SVM regression model to power DC.

Table 2. Initial parameters of the GWO, GS, and RS.

Optimizer Name	Parameter	Value
GWO	A	Min = 0 and max = 2
	Number of agents	100
	Iterations number	50
GS and RS	C = Linear	Min = 0.001 and max = 10,000
	G = Linear, RBF, sigmoid	Min = 0.001 and max = 10,000

Figure 7 depicts the prediction accuracy results of the proposed models SVM_R and GWO_R with the selected benchmark models. All models were trained using the dataset from 6 June to 21 June 2020, and the RMSE values for all models are displayed in Table 3.

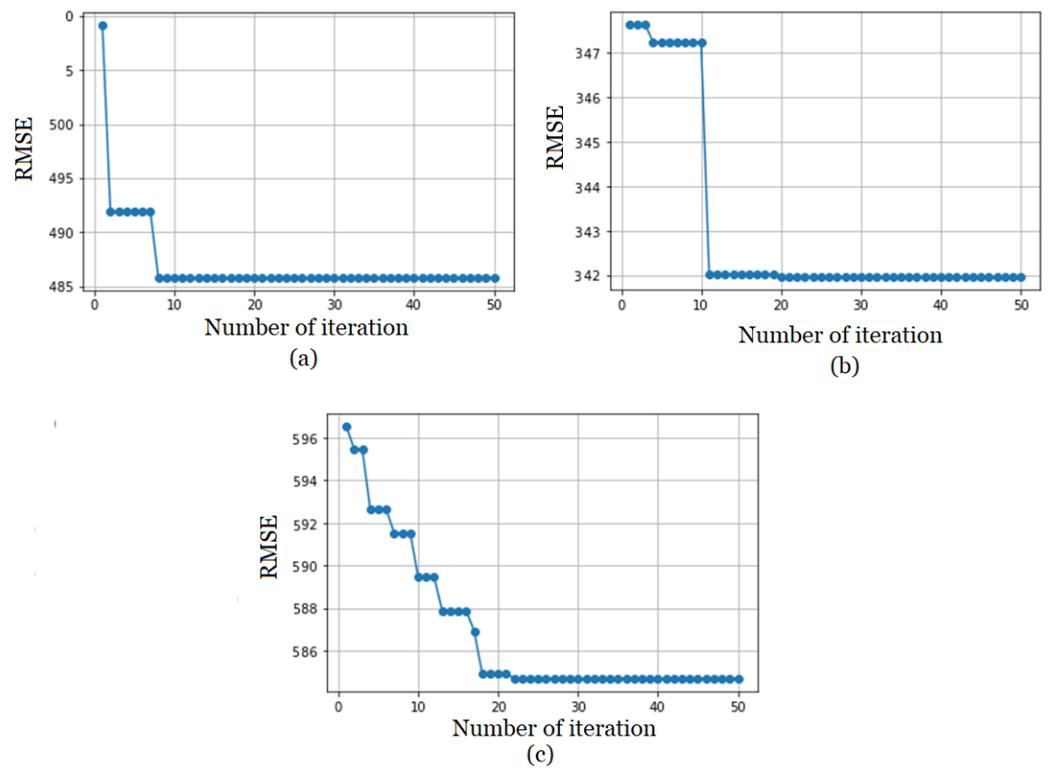


Figure 7. Convergence curves for (a) GS optimizer, (b) GW optimizer, and (c) RS.

Table 3. The impact of different optimization methods on Power_Dc prediction accuracy.

Model	RMSE
RS_SVM	532.47
SVM_R	415.98
GS_SVM	400.83
GWO-SVM_R	318.04

In general, the performance of the SVM_R prediction model increased when GS and GWO were included. The RMSE for SVM_R is reduced by 3% when using GS and 23% when using SVM-GW_R. When SVM_R is combined with SVM_RS, the RMSE of Power Dc prediction rises from (415.98) to (532.47), indicating a reduction in performance. These results demonstrate that the GWO optimizer enhances the SVM’s regression performance.

Figure 7 compares the convergence curves of the GWO, GS, and RS optimization algorithms. The physical capacity is depicted as the average physical capacity obtained. The worst, mean, and best fitness values decrease as performance increases. The results show that the GWO model performs best, as evidenced by its superior results.

Three models are tested for anomaly identification in solar power plants to categorize the inverters into two classes based on their performance (normal and fault). The first model employs the physical model, whereas the second model employs the hybrid GWO_SVM_R model. The hybrid GWO_SVM_C model is the third model. The result of this comparison is depicted in Figure 8 as a confusion matrix.

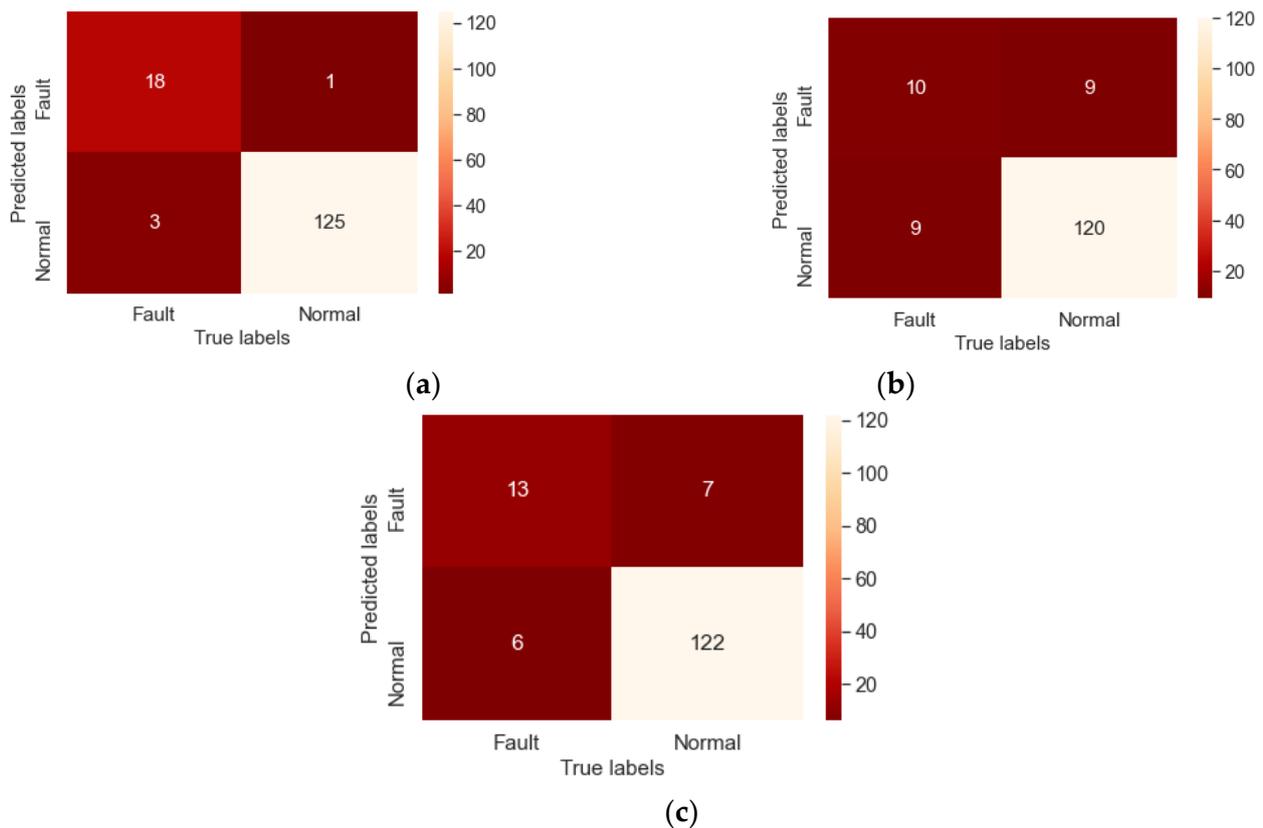


Figure 8. Confusion matrix for (a) SVM-GW_C model, (b) physical model, and (c) SVM-GW_R.

The confusion matrix reveals that the true classification of the GWO_SVM_C model for all classes was 143. On the other hand, the model was offered incorrectly four times, once as the fault class and twice as the normal class. Therefore, GWO_SVM_C attained an overall accuracy of 97.28%, followed by the SVM_GW_R model (91.22%) and the physical model (87.84%) with the lowest accuracy.

In addition, specificity and sensitivity rates were determined and displayed in Table 4 based on a confusion matrix. Compared to the physical and SVM_GW_R models, the SVM-GW_C has the greatest values for sensitivity and specificity.

Table 4. Sensitivity and specificity rates for predictive models.

Model	Sensitivity	Specificity
SVM-GW_C	85.71%	99.21%
SVM-GW_R	68.42%	94.57%
Physical model	52.63%	93.02%

Based on the previous findings, we can infer that the SVM-GW_C model is the best model for anomaly identification in solar power plants by classifying inverter states into two categories (normal and fault). In addition, SVM-GW_C can discover the day and inverter with the highest failures, as illustrated in Figures 9 and 10. On 7 June, the greatest number of errors were reported. Inverters 18 and 22 saw the most failures.

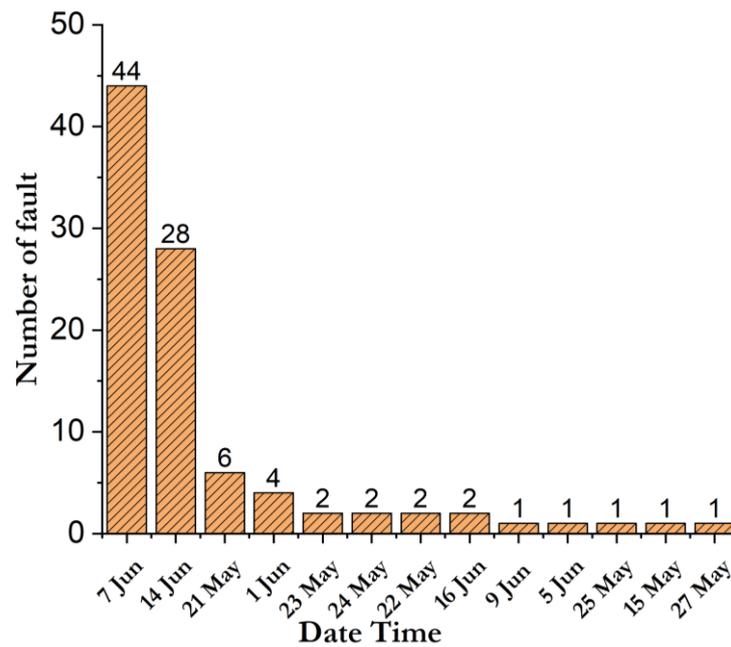


Figure 9. The number of the most failures days in inverters is ranked based on the SVM-GW_C model anomaly detection.

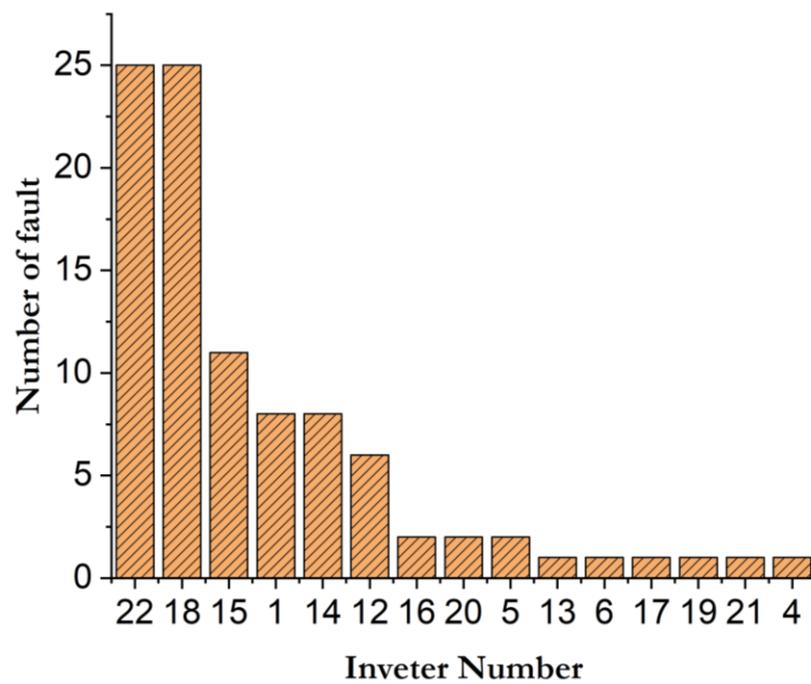


Figure 10. Ranking the number of failures in inverters based on the SVM-GW_C model anomaly detection.

5. Results Discussion

In the first experiment, the researchers tested the effectiveness of three different optimization algorithms, namely the grey wolf optimization (GWO) algorithm, grid search (GS), and random search (RS), on the task of predicting power Dc accurately. The results of the experiment showed that the GWO algorithm outperformed the other two algorithms in terms of accuracy.

The “superior results” of the GWO algorithm imply that it provided the most accurate predictions of power Dc compared to the other two algorithms. This suggests that the GWO algorithm was able to find the optimal set of parameters to minimize the error in the predictions better than the other algorithms. It is worth noting that the superiority of the

GWO algorithm in this experiment does not necessarily imply that it will always be the best algorithm for power Dc prediction. The effectiveness of optimization algorithms can vary depending on the specific problem and data set being analyzed. Therefore, further research may be required to determine the most appropriate algorithm for power Dc prediction in different contexts.

In the second test investigated, the aim was to evaluate the performance of three different algorithms in detecting DC-generated power signal abnormalities. These algorithms are SVM_GW_C, SVM-GW_R, and the physical model.

The SVM_GW_C algorithm is a support vector machine (SVM) based algorithm that uses the gradient weighted class separability (GW_C) approach for feature selection. The SVM-GW_R algorithm is another SVM-based algorithm that uses the gradient-weighted regression (GW_R) approach for feature selection. Finally, the physical model algorithm is a model-based approach that uses the physical characteristics of the power signal to detect abnormalities.

The results of the study indicate that the SVM-GW_C algorithm performs the best, with an overall accuracy of 97.28%. This means that the SVM-GW_C algorithm was able to correctly identify 97.28% of the abnormalities in the DC-generated power signal.

The results you provided refers to a comparative study conducted on various anomaly detection models for photovoltaic components. The study used the same dataset as another study presented by Mariam et al. [26], which investigated three well-known models: Autoencoder LSTM (AE-LSTM), Facebook-Prophet, and Isolation Forest. The comparative study also used two optimization methods, the grid search and genetic algorithm (GA), to tune each algorithm's hyperparameters.

The conclusion drawn from the comparative study was that the AE-LSTM model was the most effective for detecting anomalies in solar power plants, outperforming the other models that were investigated. This conclusion is based on the comparison of the performance metrics obtained from the different models when applied to the same dataset. It is important to note that this conclusion is specific to the dataset used in the study and may not necessarily generalize to other datasets or contexts.

It is important to note that accuracy is not the only measure of performance in machine learning. Other metrics, such as sensitivity and specificity, should also be considered to evaluate the performance of the algorithm. However, without this additional information, it is difficult to draw further conclusions about the effectiveness of the algorithms. Overall, the results (sensitivity and specificity) suggest that the SVM-GW_C algorithm may be a promising approach for detecting abnormalities in DC-generated power signals. However, further research is needed to fully evaluate the performance of the algorithm and its effectiveness in real-world applications.

Therefore, the accuracy, sensitivity, and specificity values were compared for our proposed model (SVM-GW_C) with the reference model (AE-LSTM) in Table 5. The results suggest that the proposed hybrid model, SVM-GW_C, has better prediction ability than the reference model, AE-LSTM, which is a traditional SVM model. This result confirms the findings of the high prediction ability of the presented hybrid model and leads to the conclusion that the (SVM-GW_C) model renders better and more precise predictions than the traditional SVM model, and it can be a reliable method in predicting future forecasts.

Table 5. Accuracy, sensitivity, and specificity rates comparison between the proposed model and the reference model.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)
SVM-GW_C	97.28	85.71%	99.21%
Reference model	89.63	94.32	100%

The results conclude that the SVM-GW_C model can be considered a reliable method for predicting future forecasts. Its superior performance over the traditional SVM model

suggests that the hybrid model approach may be a promising direction for future research in this area.

6. Conclusions

The importance of using data-driven techniques to detect anomalies in modern solar power plants in order to improve efficiency and minimize downtime is crucial. This study compares the efficacy of three distinct models to determine the best machine-learning model for precisely identifying irregularities in photovoltaic (PV) systems. The correlation between the external and internal features of the plants was calculated and used to evaluate the models' ability to recognize anomalies. Three different approaches have been investigated for detecting anomalies in solar power plants in India. The first model is based on a physical model, the second on the SVM regression model, and the third on an SVM classification model. Grey wolf optimizer was used to tune the hyper model for all models. The results showed that GWO-SVM_C was particularly skilled at identifying anomalies and accurately differentiating between healthy signals. Further research will focus on developing intelligent anomaly mitigation strategies. Another promising area for research is the application of distributed machine learning, such as federated learning, in large-scale intelligent solar power grids.

Author Contributions: Conceptualization, Q.I.A. and M.A.D.; methodology, Q.I.A., M.A.D. and H.A.; software, Q.I.A., H.A. and M.A.D.; validation, A.A.; formal analysis A.A.A.S.; investigation, H.A.; resources, A.A. and M.A.D.; data curation, M.A.D.; writing—original draft preparation, Q.I.A.; writing—review and editing, H.A.; visualization, A.A.; supervision, A.A.A.S.; project administration, H.A.; funding acquisition, H.A. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Authors confirm that the data is available at the corresponding author when required.

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

References

1. Vlamincx, M.; Heidbuchel, R.; Philips, W.; Luong, H. Region-based CNN for anomaly detection in PV power plants using aerial imagery. *Sensors* **2022**, *22*, 1244. [[CrossRef](#)] [[PubMed](#)]
2. Mendes, G.; Ioakimidis, C.; Ferrão, P. On the planning and analysis of Integrated Community Energy Systems: A review and survey of available tools. *Renew. Sustain. Energy Rev.* **2011**, *15*, 4836–4854. [[CrossRef](#)]
3. Ceglia, F.; Macaluso, A.; Marrasso, E.; Roselli, C.; Vanoli, L. Energy, environmental, and economic analyses of geothermal polygeneration system using dynamic simulations. *Energies* **2020**, *13*, 4603. [[CrossRef](#)]
4. Lucarelli, A.; Berg, O. City branding: A state-of-the-art review of the research domain. *J. Place Manag. Dev.* **2011**, *4*, 9–27. [[CrossRef](#)]
5. Andersson, I. Placing place branding: An analysis of an emerging research field in human geography. *Geogr. Tidsskr. J. Geogr.* **2014**, *114*, 143–155. [[CrossRef](#)]
6. Ko, C.-C.; Liu, C.-Y.; Zhou, J.; Chen, Z.-Y. Analysis of subsidy strategy for sustainable development of environmental protection policy. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *349*, 012018. [[CrossRef](#)]
7. Wei, Y.-M.; Chen, K.; Kang, J.-N.; Chen, W.; Wang, X.-Y.; Zhang, X. Policy and management of carbon peaking and carbon neutrality: A literature review. *Engineering* **2022**, *14*, 52–63. [[CrossRef](#)]
8. Li, L.; Wang, Z.; Zhang, T. GBH-YOLOv5: Ghost Convolution with BottleneckCSP and Tiny Target Prediction Head Incorporating YOLOv5 for PV Panel Defect Detection. *Electronics* **2023**, *12*, 561. [[CrossRef](#)]
9. Lin, L.-S.; Chen, Z.-Y.; Wang, Y.; Jiang, L.-W. Improving Anomaly Detection in IoT-Based Solar Energy System Using SMOTE-PSO and SVM Model. In *Machine Learning and Artificial Intelligence*; IOS Press: Cambridge, MA, USA, 2022; pp. 123–131.
10. Meribout, M.; Tiwari, V.K.; Herrera, J.; Baobaid, A.N.M.A. Solar Panel Inspection Techniques and Prospects. *Measurement* **2023**, *209*, 112466. [[CrossRef](#)]
11. Almalki, F.A.; Albraikan, A.A.; Soufiene, B.O.; Ali, O. Utilizing artificial intelligence and lotus effect in an emerging intelligent drone for persevering solar panel efficiency. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 7741535. [[CrossRef](#)]
12. Zeng, C.; Ye, J.; Wang, Z.; Zhao, N.; Wu, M. Cascade neural network-based joint sampling and reconstruction for image compressed sensing. *Signal Image Video Process.* **2022**, *16*, 47–54. [[CrossRef](#)]

13. Rahman, M.M.; Khan, I.; Alameh, K. Potential measurement techniques for photovoltaic module failure diagnosis: A review. *Renew. Sustain. Energy Rev.* **2021**, *151*, 111532. [[CrossRef](#)]
14. Hooda, N.; Azad, A.P.; Kumar, P.; Saurav, K.; Arya, V.; Petra, M.I. PV power predictors for condition monitoring. In Proceedings of the 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm), Sydney, Australia, 6–9 November 2016; pp. 212–217. [[CrossRef](#)]
15. Kurukuru, V.S.B.; Haque, A.; Tripathy, A.K.; Khan, M.A. Machine learning framework for photovoltaic module defect detection with infrared images. *Int. J. Syst. Assur. Eng. Manag.* **2022**, *13*, 1771–1787. [[CrossRef](#)]
16. Bu, C.; Liu, T.; Li, R.; Shen, R.; Zhao, B.; Tang, Q. Electrical Pulsed Infrared Thermography and supervised learning for PV cells defects detection. *Sol. Energy Mater. Sol. Cells* **2022**, *237*, 111561. [[CrossRef](#)]
17. Aouat, S.; Ait-hammi, I.; Hamouchene, I. A new approach for texture segmentation based on the Gray Level Co-occurrence Matrix. *Multimed. Tools Appl.* **2021**, *80*, 24027–24052. [[CrossRef](#)]
18. Mantel, C.; Villebro, F.; Benatto, G.A.D.R.; Parikh, H.R.; Wendlandt, S.; Hossain, K.; Poulsen, P.B.; Spataru, S.; Séra, D.; Forchhammer, S. Machine learning prediction of defect types for electroluminescence images of photovoltaic panels. *Appl. Mach. Learn.* **2019**, *11139*, 1113904.
19. Jumaboev, S.; Jurakuziev, D.; Lee, M. Photovoltaics plant fault detection using deep learning techniques. *Remote Sens.* **2022**, *14*, 3728. [[CrossRef](#)]
20. Lu, F.; Niu, R.; Zhang, Z.; Guo, L.; Chen, J. A generative adversarial network-based fault detection approach for photovoltaic panel. *Appl. Sci.* **2022**, *12*, 1789. [[CrossRef](#)]
21. Elsheikh, A.H.; Katekar, V.; Muskens, O.L.; Deshmukh, S.S.; Elaziz, M.A.; Dabour, S.M. Utilization of LSTM neural network for water production forecasting of a stepped solar still with a corrugated absorber plate. *Process Saf. Environ. Prot.* **2021**, *148*, 273–282. [[CrossRef](#)]
22. Elsheikh, A.H.; Panchal, H.; Ahmadein, M.; Mosleh, A.O.; Sadasivuni, K.K.; Alsaleh, N.A. Productivity forecasting of solar distiller integrated with evacuated tubes and external condenser using artificial intelligence model and moth-flame optimizer. *Case Stud. Therm. Eng.* **2021**, *28*, 101671. [[CrossRef](#)]
23. Ibrahim, M.; Alsheikh, A.; Al-Hindawi, Q.; Al-Dahidi, S.; ElMoaqet, H. Short-time wind speed forecast using artificial learning-based algorithms. *Comput. Intell. Neurosci.* **2020**, *2020*, 8439719. [[CrossRef](#)] [[PubMed](#)]
24. Deif, M.A.; Attar, H.; Amer, A.; Elhady, I.A.; Khosravi, M.R.; Solyman, A.A.A. Diagnosis of Oral Squamous Cell Carcinoma Using Deep Neural Networks and Binary Particle Swarm Optimization on Histopathological Images: An AIO/MT Approach. *Comput. Intell. Neurosci.* **2022**, *2022*, 6364102. [[CrossRef](#)] [[PubMed](#)]
25. Aslam, S.; Herodotou, H.; Mohsin, S.M.; Javaid, N.; Ashraf, N.; Aslam, S. A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. *Renew. Sustain. Energy Rev.* **2021**, *144*, 110992. [[CrossRef](#)]
26. Ibrahim, M.; Alsheikh, A.; Awaysheh, F.M.; Alshehri, M.D. Machine learning schemes for anomaly detection in solar power plants. *Energies* **2022**, *15*, 1082. [[CrossRef](#)]
27. Branco, P.; Gonçalves, F.; Costa, A.C. Tailored algorithms for anomaly detection in photovoltaic systems. *Energies* **2020**, *13*, 225. [[CrossRef](#)]
28. Deif, M.A.; Solyman, A.A.A.; Hammam, R.E. ARIMA Model Estimation Based on Genetic Algorithm for COVID-19 Mortality Rates. *Int. J. Inf. Technol. Decis. Mak.* **2021**, *20*, 1775–1798. [[CrossRef](#)]
29. De Benedetti, M.; Leonardi, F.; Messina, F.; Santoro, C.; Vasilakos, A. Anomaly detection and predictive maintenance for photovoltaic systems. *Neurocomputing* **2018**, *310*, 59–68. [[CrossRef](#)]
30. Natarajan, K.; Bala, K.; Sampath, V. Fault detection of solar PV system using SVM and thermal image processing. *Int. J. Renew. Energy Res.* **2020**, *10*, 967–977.
31. Harrou, F.; Dairi, A.; Taghezouit, B.; Sun, Y. An unsupervised monitoring procedure for detecting anomalies in photovoltaic systems using a one-class support vector machine. *Sol. Energy* **2019**, *179*, 48–58. [[CrossRef](#)]
32. Feng, M.; Bashir, N.; Shenoy, P.; Irwin, D.; Kosanovic, D. Sundown: Model-driven per-panel solar anomaly detection for residential arrays. In Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies, Guayaquil, Ecuador, 15–17 June 2020; pp. 291–295.
33. Sanz-Bobi, M.A.; Roque, A.M.S.; De Marcos, A.; Bada, M. Intelligent system for a remote diagnosis of a photovoltaic solar power plant. *J. Phys. Conf. Ser.* **2012**, *364*, 012119. [[CrossRef](#)]
34. Zhao, Y.; Liu, Q.; Li, D.; Kang, D.; Lv, Q.; Shang, L. Hierarchical anomaly detection and multimodal classification in large-scale photovoltaic systems. *IEEE Trans. Sustain. Energy* **2018**, *10*, 1351–1361. [[CrossRef](#)]
35. Mulongo, J.; Atemkeng, M.; Ansah-Narh, T.; Rockefeller, R.; Nguengang, G.M.; Garuti, M.A. Anomaly detection in power generation plants using machine learning and neural networks. *Appl. Artif. Intell.* **2020**, *34*, 64–79. [[CrossRef](#)]
36. Benninger, M.; Hofmann, M.; Liebschner, M. Online Monitoring System for Photovoltaic Systems Using Anomaly Detection with Machine Learning. In Proceedings of the NEIS 2019 Conference on Sustainable Energy Supply and Energy Storage Systems, Hamburg, Germany, 17 February 2020; pp. 1–6.
37. Benninger, M.; Hofmann, M.; Liebschner, M. Anomaly detection by comparing photovoltaic systems with machine learning methods. In Proceedings of the NEIS 2020 Conference on Sustainable Energy Supply and Energy Storage Systems Hamburg, Germany, 1 December 2020; pp. 1–6.

38. Firth, S.K.; Lomas, K.J.; Rees, S.J. A simple model of PV system performance and its use in fault detection. *Sol. Energy* **2010**, *84*, 624–635. [[CrossRef](#)]
39. Balzategui, J.; Eciolaza, L.; Maestro-Watson, D. Anomaly detection and automatic labeling for solar cell quality inspection based on generative adversarial network. *Sensors* **2021**, *21*, 4361. [[CrossRef](#)]
40. Wang, Q.; Paynabar, K.; Pacella, M. Online automatic anomaly detection for photovoltaic systems using thermography imaging and low rank matrix decomposition. *J. Qual. Technol.* **2022**, *54*, 503–516. [[CrossRef](#)]
41. Hempelmann, S.; Feng, L.; Basoglu, C.; Behrens, G.; Diehl, M.; Friedrich, W.; Brandt, S.; Pfeil, T. Evaluation of unsupervised anomaly detection approaches on photovoltaic monitoring data. In Proceedings of the 2020 47th IEEE Photovoltaic Specialists Conference (PVSC), Calgary, ON, Canada, 15 June–21 August 2020; pp. 2671–2674.
42. Iyengar, S.; Lee, S.; Sheldon, D.; Shenoy, P. Solarclique: Detecting anomalies in residential solar arrays. In Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, Menlo Park and San Jose, CA, USA, 20–22 June 2018; pp. 1–10.
43. Tsai, C.-W.; Yang, C.-W.; Hsu, F.-L.; Tang, H.-M.; Fan, N.-C.; Lin, C.-Y. Anomaly Detection Mechanism for Solar Generation using Semi-supervision Learning Model. In Proceedings of the 2020 Indo-Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), Rajpura, India, 7–15 February 2020; pp. 9–13.
44. Pereira, J.; Silveira, M. 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; pp. 1275–1282.
45. Kosek, A.M.; Gehrke, O. Ensemble regression model-based anomaly detection for cyber-physical intrusion detection in smart grids. In Proceedings of the 2016 IEEE Electrical Power and Energy Conference (EPEC), Ottawa, ON, Canada, 12–14 October 2016; pp. 1–7.
46. Rossi, B.; Chren, S.; Buhnova, B.; Pitner, T. Anomaly detection in smart grid data: An experience report. In Proceedings of the 2016 IEEE International Conference on Systems Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; pp. 2313–2318.
47. Toshniwal, A.; Mahesh, K.; Jayashree, R. Overview of anomaly detection techniques in machine learning. In Proceedings of the 2020 Fourth International Conference on I-SMAC (IoT in Social Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 7–9 October 2020; pp. 808–815.
48. Deif, M.A.; Solyman, A.A.A.; Alsharif, M.H.; Jung, S.; Hwang, E. A hybrid multi-objective optimizer-based SVM model for enhancing numerical weather prediction: A study for the Seoul metropolitan area. *Sustainability* **2021**, *14*, 296. [[CrossRef](#)]
49. Hammam, R.E.; Attar, H.; Amer, A.; Issa, H.; Vourganas, I.; Solyman, A.; Venu, P.; Khosravi, M.R.; Deif, M.A. Prediction of Wear Rates of UHMWPE Bearing in Hip Joint Prosthesis with Support Vector Model and Grey Wolf Optimization. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 6548800. [[CrossRef](#)]
50. Deif, M.A.; Hammam, R.E.; Solyman, A.; Alsharif, M.H.; Uthansakul, P. Automated Triage System for Intensive Care Admissions during the COVID-19 Pandemic Using Hybrid XGBoost-AHP Approach. *Sensors* **2021**, *21*, 6379. [[CrossRef](#)]
51. Deif, M.A.; Hammam, R.E. Skin lesions classification based on deep learning approach. *J. Clin. Eng.* **2020**, *45*, 155–161. [[CrossRef](#)]
52. Deif, M.A.; Solyman, A.A.; Kamarposhti, M.A.; Band, S.S.; Hammam, R.E. A deep bidirectional recurrent neural network for identification of SARS-CoV-2 from viral genome sequences. *Math. Biosci. Eng. AIMS-Press* **2021**, *18*, 8933–8950. [[CrossRef](#)]
53. Deif, M.A.; Attar, H.; Amer, A.; Issa, H.; Khosravi, M.R.; Solyman, A.A.A. A New Feature Selection Method Based on Hybrid Approach for Colorectal Cancer Histology Classification. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 7614264. [[CrossRef](#)]
54. Zamfirache, I.A.; Precup, R.-E.; Roman, R.-C.; Petriu, E.M. Policy iteration reinforcement learning-based control using a grey wolf optimizer algorithm. *Inf. Sci.* **2022**, *585*, 162–175. [[CrossRef](#)]
55. Kannal, A. Solar Power Generation Data. Kaggle.com. 2020. Available online: <https://www.kaggle.com/anikannal/solar-powergeneration-data> (accessed on 30 January 2022).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.