IJRCS
International Journal of Robotics and Control Systems
Vol. 4, No. 3, 2024, pp. 1337-1352
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs

Photovoltaic Energy Anomaly Detection using Transformer Based Machine Learning

I Made Wirawan [a,1,*], Aji Prasetya Wibawa [a,2], Triyanna Widiyanintyas [a,3]

[a] Department of Electrical Engineering and Informatics, Universitas Negeri Malang, Jl Semarang 5, Malang and 65145, Indonesia
[1] made.wirawan.ft@um.ac.id; [2] aji.prasetya.ft@um.ac.id; [3] triyannaw.ft@um.ac.id

ARTICLE INFO

Article history
Received December 16, 2023
Revised July 15, 2024
Accepted August 08, 2024

Keywords
Photovoltaic Energy;
Anomaly Detection;
Time Series;
Transformers Anomaly

ABSTRACT

This study uses the Anomaly Transformer model to find anomalies in photovoltaic energy generation in Malang, Indonesia. The main background of this study is the lack of satellite monitoring in this region and the importance of annual data for electricity generation forecasting. Temperature scattered direct solar radiation, and hourly electricity production are all part of the dataset used which is only available since 2019. Anomalies were detected at 05.00 and 16.00 WIB, indicating instability in the energy supply due to high temperatures in the morning and heavy rain in the afternoon. Detection of these anomalies is important to improve the efficiency and reliability of photovoltaic systems, reduce operational costs, and reduce the risk of system failure. Indonesia has many challenges for photovoltaic energy generation due to its unique location, with many islands located close to the equator. The use of the Anomaly Transformer algorithm improves the accuracy of anomaly detection over conventional methods. This algorithm helps to find complex patterns in very large time series. The results show that the anomaly transformer model can effectively detect anomalous patterns. It offers ideas to improve the stability and efficiency of photovoltaic systems in Malang and other areas with comparable environmental conditions. Improved energy efficiency and environmental sustainability are the results of anomaly pattern detection.

1. Introduction

The recent rise in the popularity of generative models can be attributed to the extensive adoption of deep learning in both academic and industrial sectors [1], [2]. Generative models are especially useful in situations where the original data is insufficient or unavailable [3], [4], where privacy concerns restrict data usage, where there is a requirement to simulate events that have never occurred or to generate datasets for specific testing scenarios, such as the hourly presence of photovoltaic-generated electricity in Indonesia.

In the context of renewable energy, generative models play an important role as they can provide solutions to challenges [5], [6]. For example, conditions and to predict future energy system performance, simulate rare conditions, and optimize renewable energy utilization strategies. They can also help in overcoming challenges such as variability and uncertainty in energy production, which are often major obstacles in the integration of renewable energy into the existing power grid [7].

For this dataset-based research project, we employed the photovoltaic electric energy dataset [8], [9] for locations in Malang, Indonesia. The photovoltaic electrical energy data set in Malang, Indonesia, is relevant and interesting because it reflects the challenges and opportunities faced in managing renewable energy in tropical areas [10], [11]. Malang has a unique climate, offering rich data to understand variations in photovoltaic energy production throughout the year [12], [13]. This data set is suitable for generative modeling because it contains high variability and complexity, which are ideal conditions for testing the ability of generative models to generate realistic and useful data [14], [15]. Energy production patterns under various weather conditions, predictions of energy system performance in the long term, and optimization of renewable energy utilization strategies are some examples of important information that can be obtained by using generative models on these data sets [16], [17]. This has significant benefits for scientific research and the development of sustainable energy policies in Indonesia. The dataset comprises the following elements: hourly temperature, direct and diffuse irradiance, electricity, and local time as shown in Fig. 1. These elements have an important role in the generative modeling process, namely hourly temperature, variations in photovoltaic panel efficiency can be measured by modeling hourly temperature, which creates more realistic energy production data [18], [19]. Next is direct and scattered radiation, scattered radiation is light that is reflected or scattered in the atmosphere [20], [21], while direct radiation is light that reaches the panel directly [22], [23]. These two types of radiation affect the amount of energy that photovoltaic panels can produce. By including these two components, the model can replicate a wide range of lighting conditions and produce accurate energy production data [24], [25]. The next element is the electricity produced [26], [27], data regarding the electricity produced is the main output we want to model. This element is very important because it is the basis for evaluating the performance and validity of the generative model that we develop. The final element is local time [28]-[30], Local time information is important for correlating other elements with daily and seasonal cycles. This helps the model capture temporal patterns in energy production, which is very useful for long-term analysis and predictions.

	local_time	electricity	irradiance_direct	irradiance_diffuse	temperature
0	2019-01-01 07:00:00	0.275	0.121	0.212	25.464
1	2019-01-01 08:00:00	0.437	0.214	0.326	26.661
2	2019-01-01 09:00:00	0.535	0.268	0.409	27.723
3	2019-01-01 10:00:00	0.565	0.253	0.469	28.472
4	2019-01-01 11:00:00	0.592	0.276	0.487	29.039
5	2019-01-01 12:00:00	0.545	0.218	0.478	29.195
6	2019-01-01 13:00:00	0.483	0.186	0.422	29.145
7	2019-01-01 14:00:00	0.415	0.181	0.335	28.948
8	2019-01-01 15:00:00	0.251	0.093	0.216	28.307
9	2019-01-01 16:00:00	0.087	0.029	0.088	27.252

Fig. 1. Sample dataset photovoltaic (www.renewables.ninja)

Fig. 1 shows the daily variation in photovoltaic electricity production in Malang, Indonesia taken on January 1, 2019, at the time interval of 07:00 to 16:00. From this figure, peak electricity production occurs around midday when solar radiation is most intense. In addition, seasonal patterns are also evident, with higher electricity production during the dry season compared to the rainy season. This data includes hourly temperature, direct and scattered radiation, and local time, all of which contribute to an understanding of the factors that affect the efficiency and output of photovoltaic systems. At 07:00, the electricity value was 0.275 with an irradiance direct of 0.121 and an irradiance diffuse of 0.212. The temperature at that time was 25.464°C. The highest electricity value was at 11:00, which was 0.592, with irradiance direct of 0.276 and irradiance diffuse of 0.487. The temperature at that time was 29.039°C. The irradiance values (both direct and diffuse) tend to increase from morning to midday and then decrease in the afternoon. The air temperature also increased from morning to noon, reaching a peak around 12:00 to 14:00, then decreased slightly in the afternoon.

The main objective of this research is to develop and apply a generative model that can effectively detect anomalies in photovoltaic energy data. By leveraging the generative model, we aim to generate realistic simulations of energy production that can highlight anomalous patterns. Understanding and identifying these anomalies will improve the efficiency and reliability of photovoltaic energy systems in the region. Additionally, this research aims to contribute to the development of more advanced anomaly detection algorithms that can be applied to similar datasets in other locations, thereby enhancing environmental sustainability and energy efficiency globally.

2. Method

To address the challenge of anomaly detection in photovoltaic energy time series data, this research utilizes an innovative machine learning-based approach. Anomaly detection is a crucial element in ensuring the reliability and efficiency of photovoltaic energy systems [31], [32]. Significant variations in sunlight intensity and weather conditions affect the performance of photovoltaic panels [33], [34], making anomaly detection a complex yet important task. This research utilizes a modified Transformer model, known as the Anomaly Transformer [35], [36], to detect anomalous patterns in photovoltaic energy data [37], [38]. This model was chosen for its ability to handle the complexity of time series data with multi-level features [39], [40]. By integrating the Anomaly-Attention mechanism, the model is expected to identify latent relationships in the data and provide more accurate results than conventional methods. This research focuses on anomaly detection in photovoltaic energy production data in Malang, Indonesia. Major challenges faced in photovoltaic energy production in Indonesia include significant temperature variations and heavy rainfall that can affect the efficiency and stability of photovoltaic systems. Thus, anomaly detection becomes highly relevant to understanding and addressing the instability that occurs in these energy systems. The Anomaly Transformer model [41], [42] was chosen for its ability to handle the complexity of time series data with various features present in photovoltaic data such as temperature, direct and scattered solar radiation, and hourly electricity production. The identification of anomalies at 05:00 and 16:00 WIB, for example, shows the instability of the energy supply caused by high temperatures in the morning and heavy rainfall in the afternoon. This is crucial for improving the efficiency and reliability of photovoltaic systems in Malang and other regions in Indonesia with similar environmental conditions. The anomaly detection process involves several important stages, from data collection and preprocessing to model application and evaluation. The following diagram illustrates the Anomaly Transformer architecture used in this study shown in Fig. 2.

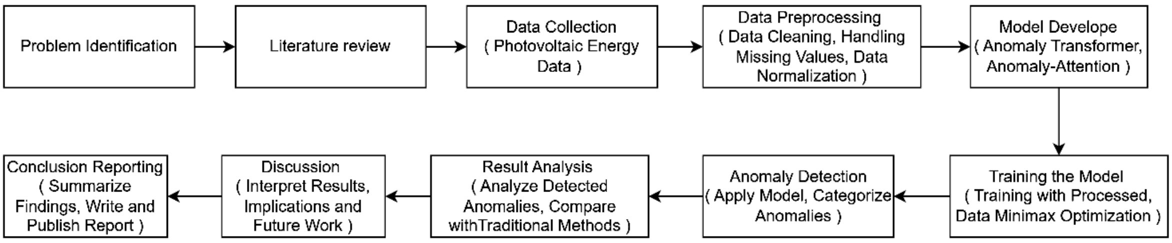


Fig. 2. Research flow diagram

2.1. Anomaly Detection

Anomaly detection is a prevalent application of machine learning and not a novel concept or method [43]-[45]; it has existed for several years. Illustrative instances of its practical applications encompass, yet are not restricted to, the identification of fraudulent insurance claims, fraudulent transactions, cyber assaults, and atypical equipment behaviors.

As a result, the capacity to proactively identify irregularities and mitigate associated risks is a highly advantageous capability that not only prevents unscheduled downtime but also unnecessary maintenance (condition-based as opposed to mandatory maintenance) but also facilitates a more efficient approach to overseeing critical components for such assets [46], [47]. The financial ramifications of unscheduled downtime, unnecessary maintenance expenses, and surplus or deficiency of critical components are of significant magnitude [48], [49].

Anomaly subsequence detection in time series data is a critical undertaking that has wide-ranging implications, spanning from finance applications to health care monitoring to manufacturing processes [50], [51]. An anomaly, which may serve as an indicator for critical occurrences including production malfunctions, delivery delays, system imperfections, or heart flicker, is thus a subject of primary interest [52], [53]. Due to the frequent occurrence of complex patterns and large data series, data scientists have devised a multitude of specialized algorithms to automatically identify these anomalous patterns. In the past, there has been a substantial increase in the quantity and diversity of anomaly detection algorithms. However, because numerous solutions have been developed autonomously and by distinct research communities, there is currently no exhaustive study that systematically assesses and contrasts the various methodologies. Consequently, selecting the optimal detection method for a given anomaly detection task is a challenging endeavor [54].

For example, the financial consequences of unscheduled downtime and unnecessary maintenance costs can be substantial. Research conducted [55], [56] shows that anomaly detection in time-series sensor data can reduce this risk by identifying unusual patterns before they cause significant damage. This demonstrates the importance of proactive anomaly detection capabilities to prevent unplanned downtime and unnecessary maintenance costs, as well as to facilitate a more efficient approach to monitoring critical components.

In the context of this research, anomaly detection aims to identify unusual patterns that may indicate critical issues such as production malfunctions, delivery delays, or system imperfections. For example, anomalies detected at 05:00 AM and 4:00 PM indicate environmental mismatches caused by extreme temperatures and heavy rainfall. By understanding these anomalies, we can take corrective actions to improve the stability and reliability of the photovoltaic energy supply.

The Anomaly Transformer model we developed allows us to overcome the complexity of patterns in photovoltaic data and detect anomalies more effectively compared to traditional methods. The application of this model is expected to not only improve the performance of photovoltaic systems in Malang but also be applied to similar datasets in other regions, ultimately contributing to the improvement of global energy sustainability and efficiency.

Hawkins [57] defined an outlier as an observation that significantly diverges from the rest of the set to raise doubts regarding its origin from a distinct mechanism. Within this framework, an anomaly in time-series data can be defined as one or more data points at a given time step that exhibit unforeseen behaviors that are notably distinct from those of preceding time steps. Following prior scholarly works, we classify the various categories of anomalies associated with time-series data as follows.

2.1.1. Point Anomaly

A point anomaly is an abrupt departure from the norm by a data point or sequence [58], [59]. These anomalies, which frequently manifest as temporal noise, are frequently the result of sensor errors or atypical system operations. In the context of detection, it is customary for operators to establish upper and lower control limits (UCL and LCL, respectively) using historical data. Values that fall beyond the specified boundaries are considered point anomalies.

2.1.2. Contextual Anomaly

A contextual anomaly, akin to a point anomaly, pertains to a data point or sequence that is observed within a brief period but does not exhibit the same degree of deviation from the expected range as anomalies delimited by predefined UCLs and LCLs [60]. Nevertheless, considering the provided context, the data points deviate from the anticipated pattern or form. Consequently, detecting these anomalies may prove to be challenging.

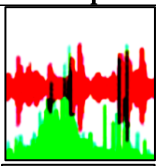
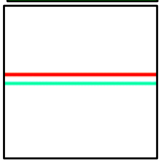
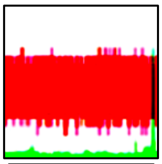
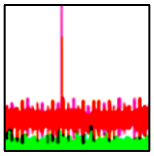
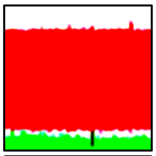
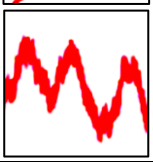
2.1.3. Collective Anomaly

This category of anomaly pertains to a collection of data points that warrant investigation as anomalous due to their progressive deviation from typical data patterns over a period [61], [62]. While individual values within this anomaly may appear to be inconsequential, when considered collectively, they generate skepticism. Given that they are not readily identifiable at first glance, long-term contexts are crucial for their detection.

2.1.4. Other Anomaly Types

The definition of abnormality is contingent upon the classification of an anomaly as a state that deviates from the norm. In broad terms, anomalies can be categorized into one of the three types; however, alternative viewpoints may further delineate anomalies into more particularized classifications. The taxonomy of anomaly patterns and illustrative instances from references [63], [64] are presented in Table 1, each data section was plotted in time domain and frequency domain with Fast Fourier Transform (FFT). Then a T- F image was generated by stacking time response image as channel 1 (red) and frequency response image as channel 2 (green).

Table 1. Description of the detailed classification of anomalies in time-series data

Anomaly Patterns	Description	Examples
Normal (assumption)	The amplitude and frequency are stable over time steps, and the time response is symmetrical.	
Missing	Most/all the data are missing, and the time/frequency response becomes 0.	
Minor	Compared to normal sensor data, the vibration amplitude is very small.	
Outlier	One or more outliers appear in the response.	
Square	The time response oscillates within a limiting range like a square wave.	
Trend	The data has an obvious non-stationary and monotonous trend.	
Drift	The vibration response is nonstationary, with random drift.	

A data point whose occurrence was previously either exceedingly rare or logically impossible constitutes an anomaly [65], [66]. On the other hand, anomaly classification may not apply to multivariate time series data as in the preceding instances. When dealing with multivariate time series data, it is crucial to consider the interrelation among variables in addition to the time axis. Diversified patterns manifest as the number of variables escalates. Atypical patterns may then exhibit irregularities, thereby creating ambiguity regarding the distinction between normal and abnormal states. The accuracy of detection results cannot be guaranteed by merely scanning and aggregating individual univariate time series data to identify anomalies [67]. This is because a small number of anomaly points may be obscured by the other normal variables, which could have a substantial impact on the entire target system. Such issues can be resolved by reducing the dimensions through the extraction of distinct variables or features, or by employing a model that is sufficiently complex to identify a multitude of patterns.

2.2. Anomaly Transformer

Anomaly Transformer is a modification of the Transformer architecture that integrates the Anomaly-Attention mechanism. The main purpose of this modification is to detect anomalous patterns in time series data more effectively [68]. The model consists of multiple layers consisting of feed-forward layers and Anomaly-Attention blocks arranged alternately. In light of the anomaly detection limitation of Transformers [69], [70], we modify the vanilla architecture to incorporate an Anomaly-Attention mechanism into the Anomaly Transformer Fig. 3. The overall architecture anomaly transformer is distinguished by the alternating stacking of feed-forward layers and Anomaly-Attention blocks. This stacking structure facilitates the discovery of latent associations underlying complex multi-level features. Suppose the model contains L layers with length-N. input time series $X \in R^{N\,x\,d}$. The overall equations of the l-th layer are formalized as:

$$\begin{gathered} Z^l = Layer - Norm\,(Anomaly - Attention\,(X^{l-1}) + X^{l-1}) \\ X^l = Layer - Norm\,(Feed - Forward\,(Z^l\ ) + Z^l\,) \end{gathered} \tag{1}$$

where $X^l \in R^{N\,x\,d}$model, $l \in \{1, \cdot\cdot\cdot, L\}$ denotes the output of the l-th layer with model channels. The initial input X 0 = Embedding (X) represents the embedded raw series. $Z^l \in R^{N\,x\,d}$model is the l-th layer's hidden representation. Anomaly-Attention (·) is to compute the association discrepancy.

A mechanism called Anomaly-Attention is designed to find anomalies in the data. This mechanism counts differences or discrepancies in the relationships between elements in the time data [71]. This helps the model find abnormal patterns by emphasizing unusual relationships in the data. The layer normalization method applied to the output of each layer of the model is known as the norm layer [72]. This method reduces dependence on the absolute values of inputs by standardizing those values based on layer statistics. This stabilizes and speeds up the model training process by ensuring that the distribution of data remains consistent across each layer. A simple Feed-Forward neural network consists of several linear layers with non-linear activation functions in between [73]. This maps a representation between the input and output, allowing the model to identify complex features in the data. Equation (1) can be explained as follows: Z^l is the hidden representation of the lth layer generated after going through the Anomaly-Attention mechanism and layer normalization. X^l is the output of the lth layer after applying feed-forward and layer normalization. $X^{(l-1)}$ is the input to the lth layer, which is the output of the previous layer. L is the total number of layers in the model, N is the length of the input time series, and d is the dimension of the model channel.

We include an illustrative example to clarify the working mechanism of the Anomaly Transformer model i.e. Time series data processed through linear transformation to extract important features. The Anomaly Attention mechanism uses a multiplication matrix to combine the outputs of prior association and serial association, which are then further processed in the reconstruction phase [74]. Normalization of the output is done to standardize the final layer so that the data is ready for analysis or decision-making.

Vol. 4, No. 3, 2024, pp. 1337-1352

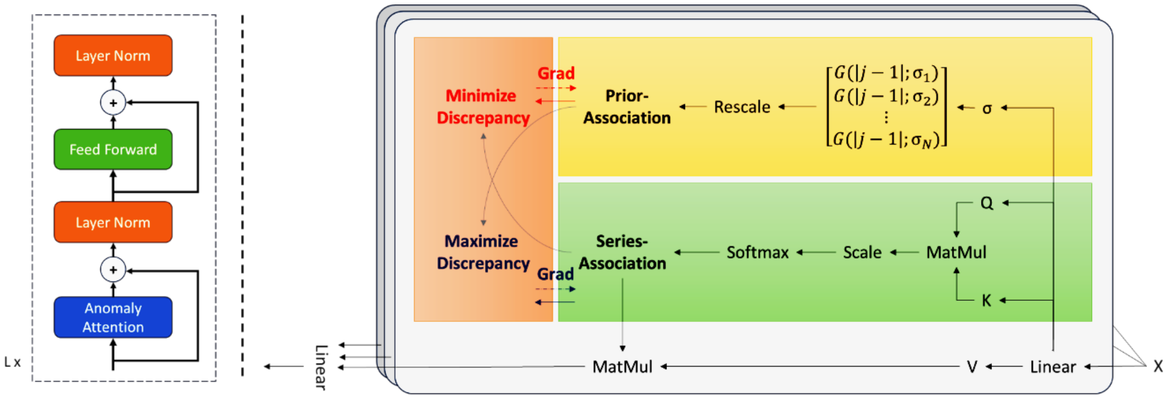


Fig. 3. Anomaly transformer architecture [75]

Fig. 3 represents an Anomaly Transformer, focusing on its components and how they contribute to the detection and processing of anomalies. Components of Anomaly Transformer: The input data, indicated by the letters X and Y, is processed through a linear transformation to extract the most important features. Key Conventions: The previous conventions are discussed at the top of the diagram. It includes calculating various Gaussian distributions, such as $G(j \mid j-1;;\sigma 1), G(j \mid j-2;\sigma), \ldots, G(j \mid j-N;\sigma N)$. These distributions are rescaled and used to stop the gradient flow, or Stop Grad, to reduce the difference. Serial Family: The Series Association is at the bottom. It starts by using a series of linear transformations and matrix multiplication (MatMul) to create queries (Q), keys (K), and values (V) from the input. To determine the relevance of the serial data, the questions and buttons are scaled and passed through the SoftMax function. The result is used to maximize the difference by allowing gradient flow (Grad). Anomaly Attention: Matrix multiplication (MatMul) is used to combine the output of the Prior Association and Serial Association. This combined result is further processed in the reconstruction phase. Standardized Layer Reconstruction Phase: To standardize the output, layer normalization is performed on the combined result. Feedback provisioning: The feed-forward network passes the normalized output to learn complex representations. Anomaly Attention: To find anomalies in the data, an anomaly attention mechanism is used. Last Layer Standardization: The output of the anomaly attention mechanism is normalized with the final layer. To extract features, the overall workflow of the input data X is linearly transformed. To calculate Prior-Association and Series-Association, these features concentrate on different elements of the data. The outputs of these groups are combined and standardized. Anomaly attention finds anomalies, and the feed-forward network refines the output. The final output is normalized again to make it consistent and ready for analysis or decision-making. Suggestions to Improve Understanding: Integrating more images or diagrams that show specific steps or components can be very helpful for improving understanding. For example: Prior-Association Gaussian distribution diagrams, step-by-step flowcharts showing the process in Series-Association, and visual representations of how anomaly attention works with real data examples.

3. Results and Discussion

Given the extensive collection of historical data on electrical energy, albeit with certain gaps, the decision was made to concentrate on the data from 2019. This is because the specified period has a consistent daily pattern and rarely has missing data, as illustrated in Fig. 4. Should any data become missing, they can be easily approximated, providing a comprehensive and valuable data source from which to develop this research approach.

Data from 2019 was chosen because this period shows consistent daily patterns and has a very low incidence of missing data. The consistency of the daily pattern is very important for this study as it allows us to analyze and detect anomalies more accurately. In anomaly detection research, data reliability is crucial to ensure the validity of the results obtained. In addition, the data from 2019 provides complete and continuous coverage, allowing us to build a more robust and reliable model.

Fig. 4 shows the daily pattern of electrical energy produced by the photovoltaic system at the study site. The data shows significant fluctuations in energy with peak production occurring during the day when sunlight intensity is maximized. This pattern helps identify anomalies such as a drop in production at a certain time of day which could be caused by environmental conditions such as extreme temperatures or heavy rainfall.

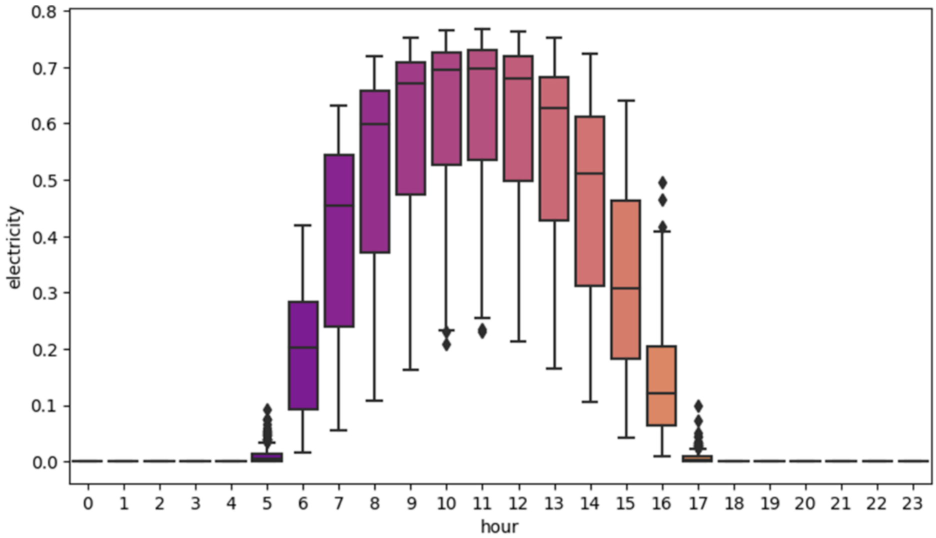


Fig. 4. Daily pattern of electrical energy from photovoltaic

We present the statistical outcomes and graphical representations of our three fundamental designs—anomaly criterion, learnable prior association, and optimization strategy—to intuitively illustrate how our model operates. The main purpose of presenting statistical results and graphical representations in this study is to demonstrate the effectiveness of the Anomaly Transformer model in detecting anomalies in photovoltaic energy production data in Malang, Indonesia. We highlight three basic designs that contribute to the overall model: anomaly criteria, learnable prior associations, and optimization strategies.

In this study, we use several performance metrics to evaluate the effectiveness of the anomaly criteria applied by the Anomaly Transformer model. The main metric used includes discriminability, which is measured by the ability of the model to effectively distinguish between normal and anomalous data. Discriminability is measured by comparing the values generated for the normal and anomalous components. A good model will provide lower values for the normal component and higher values for the anomalous component, making it easier to identify anomalies. Success in anomaly detection is defined as the ability of the model to provide distinct values between normal and anomalous data. For example, in the context of anomaly detection in photovoltaic energy data, lower values for the normal component indicate that the model can identify data that is stable and conforms to expected patterns, while higher values indicate deviations or unusual conditions. We use association-based criteria to detect unusual patterns of relationships in the data. This method proves effective in identifying different types of anomalies, including point anomalies, contextual anomalies, and collective anomalies. The results obtained from the Anomaly Transformer model are compared with conventional anomaly detection methods. The graphs and statistical representations we present show that our model has a higher accuracy rate and better detection capabilities. Using these metrics, we can confirm that the Anomaly Transformer model is not only able to detect anomalies with high accuracy but can also provide deeper insights into the anomaly patterns that occur in photovoltaic energy production data. This is crucial for improving the stability and efficiency of energy systems in Malang and other regions with similar environmental conditions.

Observation of anomaly criteria to illustrate the operation of association-based criteria more understandably, we illustrate some examples in Fig. 5 and examine the performance of criteria in different types of anomalies, using Lai's taxonomy [76]. Lai's anomaly taxonomy is a classification used to categorize different types of anomalies in time series data. This taxonomy includes several types of anomalies, the first being Point Anomalies, A point anomaly is a sudden deviation from the norm by a single point or sequence of data. It is usually caused by a sensor error or unusual system operation. second Contextual Anomalies, Contextual anomalies are deviations that occur within a specific time context, although the data values themselves may not appear deviant when viewed in isolation. third Collective Anomalies, Collective anomalies are a series of data points that collectively exhibit unusual patterns, although the individual values may not appear anomalous.

Overall, we find that our proposed association-based criterion provides greater discriminability. Association-Based Criteria is a method used to detect anomalies by identifying patterns of relationships between data that appear together. This method aims to find unusual relationships or correlations between data elements that should not occur under normal conditions [77]. The main role of association-based criteria in detecting anomalies is Unusual Pattern Detection i.e. they can recognize unusual patterns that conventional methods cannot detect. For example, in the context of photovoltaic data, if there is an unusual correlation between high temperature and low energy production at any given time, this can be an indicator of an anomaly. Next Detection Error Reduction, by identifying unusual associations, this method helps reduce detection errors or false positives that may occur when using other anomaly detection methods. Lastly Deep Understanding, This method provides a deep understanding of how data elements interact with each other, thus enabling more accurate and efficient anomaly identification. The association-based criterion consistently produces lower values for the normal component, which stands in stark contrast to the situation in pattern-seasonal and point-contextual scenarios in Fig. 5. The jitter curves associated with the reconstruction criterion, on the other hand, disrupt the detection process and fail in the two situations mentioned above [78]. This validates the ability of our criterion to identify anomalies and assign unique values to normal and abnormal data points, thereby increasing the accuracy of the detection and decreasing the occurrence of false-positive results.

Association-based criteria are different from other anomaly detection criteria such as Threshold-Based Criteria, this method sets upper and lower limits for data values that are considered normal [79]. Data that exceeds these thresholds is considered anomalous. However, this method is less effective in detecting anomalies that occur due to the pattern of relationships between data. Secondly, Statistics-Based Criteria, this method uses descriptive statistics (e.g., mean, standard deviation) to determine the normality of the data. Anomalies are determined based on significant deviations from normal statistics. However, this method cannot detect anomalies that occur due to complex relationships between data. Third Machine Learning-based Criteria, this method uses machine learning algorithms to learn from historical data and detect anomalies. Association-based criteria can be combined with this method to improve the accuracy of anomaly detection.

Fig. 5 displays the different categories of anomalies in time series according to Lai's taxonomy, using association-based criteria and reconstruction criteria [80]. Anomalies are categorized into Point Anomalies and Pattern Anomalies [81]. Within Point Anomalies, there are Global Anomalies and contextual anomalies. Global Anomalies are data points that deviate far from other normal values. visualization of the red dots on the input time series graph shows the presence of global anomalies, Detection is successfully carried out by both criteria with high anomaly values. then Contextual Anomalies are deviations that occur in a specific time context. The visualization of the two red dots on the input time series graph shows a contextual anomaly, the reconstruction criterion produces detection errors while the association-based criterion successfully detects anomalies more precisely. Besides Pattern Anomaly, there are Shapelet Anomaly, Seasonal Anomaly, and Trend Anomaly. Shapelet Anomaly is a data pattern that differs significantly from other normal patterns in a short period. Visualization of the red area on the input time series graph shows a shapelet anomaly, the reconstruction criterion shows a detection error while the association-based criterion manages to detect the anomaly more precisely. next Seasonal Anomaly, seasonal anomaly is a pattern that differs in seasonal cycles. Visualization on the Red area on the input time series graph shows seasonal anomaly, Reconstruction criteria show detection error while association-based criteria successfully detect the anomaly more precisely. The next Trend Anomaly is an unusual trend change in the time

Vol. 4, No. 3, 2024, pp. 1337-1352

series data. Visualization on Red arrows on the input time series graph shows unusual trend changes, Reconstruction criteria show detection errors while association-based criteria successfully detect anomalies more precisely.

The visualization of prior associations. To approach the series association, the prior association is learned during minimax optimization. Therefore, the acquired σ can represent the degree of adjacent concentration of the time series. Fig. 6 illustrates that σ undergoes modifications to accommodate diverse data patterns within time series. In particular, the prior association of anomalies typically exhibits a smaller σ value than typical time points, which corresponds to the inductive bias of anomalies towards adjacent concentrations.

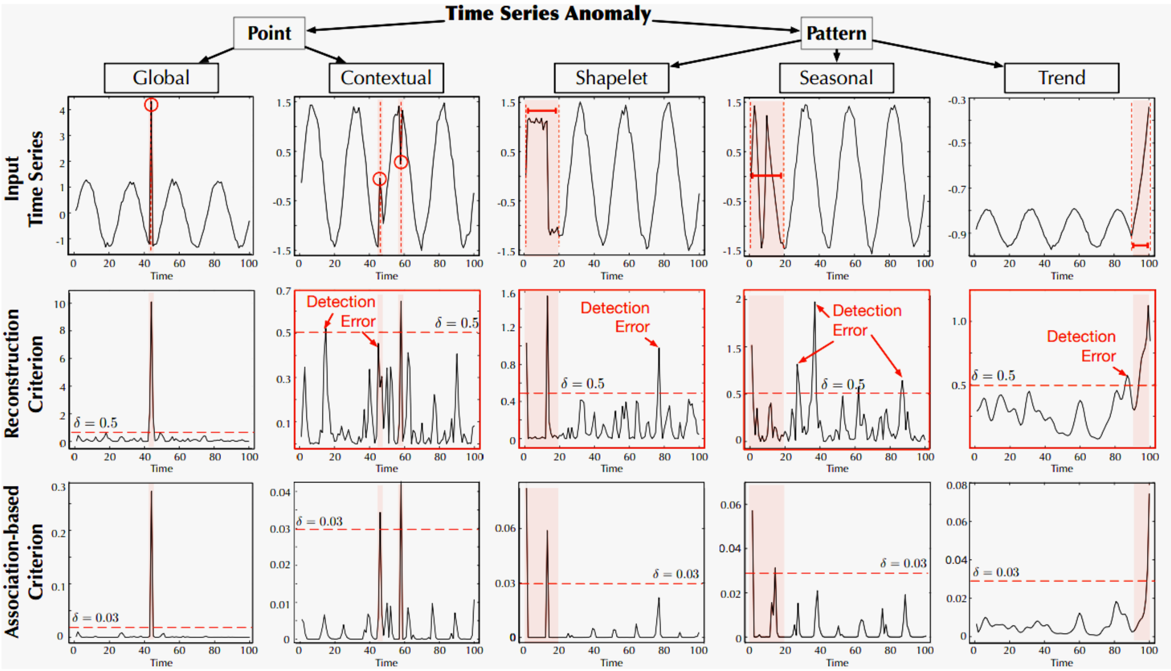


Fig. 5. Visualization of different anomaly categories [9]

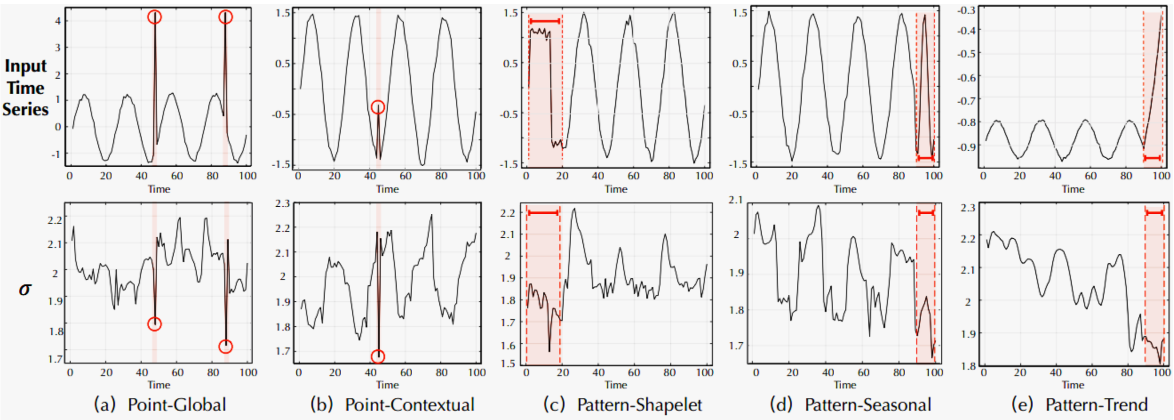


Fig. 6. Learned scale parameter σ for different types of anomalies (highlighted in red) [9]

Fig. 6 illustrates the different types of anomalies in the time series and how the learned scale parameter, σ, is modified to detect these anomalies. There are five types of anomalies, namely Point-Global Fig. 6 (a): Input Time Series i.e. The input time series is shown at the top of Fig. 6 (a) with two global point anomalies highlighted with red circles. σ: The graph below shows the increasing σ values at these anomalous points. The red highlights in this section indicate the location of the global point anomalies in the time series. Point-Contextual Fig. 6 (b): Input Time Series i.e. The input time series is shown with one contextual point anomaly highlighted with a red circle. σ: The graph below shows the increasing value of σ at the point contextual anomaly. The red highlight shows the contextual point anomaly in the local context of the time series. Pattern-Shapelet Fig. 6 (c): Input Time Series I.e. Input time series with shapelet pattern anomalies highlighted with red areas. σ: The graph below shows a significant increase in the σ value in the area containing the anomalous shapelet pattern. The red highlights reflect the duration and location of the shapelet pattern anomalies. Pattern-Seasonal Fig. 6 (d): Input Time Series I.e. Input time series with seasonal pattern anomalies highlighted with red areas. σ: The graph below shows the increase in σ value in the area containing the seasonal anomaly. The red highlights indicate the period and location of the seasonal anomaly in the time series. Pattern-Trend Fig. 6 (e): Input Time Series I.e. Input time series with trend pattern anomalies highlighted with red areas. σ: The graph below shows the increase in the σ value in the area with the trend anomaly. The red highlights indicate the period and location of the trend pattern anomaly in the time series.

From the explanation, all types of anomalies can be run in this study. However, it is important to pay attention to specific values such as accuracy, as has been measured in previous studies [37]. This will help in evaluating the effectiveness of the methods used and ensure that the results obtained have high validity. As such, this research not only demonstrates the ability to detect anomalies, but also provides quantitative evidence supporting the accuracy of such detection.

4. Conclusion

Using the Anomaly Transformer model, this study identifies anomalous patterns in photovoltaic energy in Malang City, Indonesia. The research focused on specific times between 05:00 and 16:00 WIB, when environmental mismatches, such as high temperatures in the morning and heavy rainfall in the afternoon, impact the efficiency and stability of the photovoltaic energy supply. A thorough analysis of these environmental parameters shows that. This conclusion also includes an explanation of the methodology and algorithms used to detect the anomalies, as well as the limitations and assumptions made during the analysis. In addition, it is important to talk about the impact of the detected anomalies on the performance and stability of photovoltaic energy systems, and how they can be applied elsewhere with comparable environmental conditions. By employing the Anomaly Transformer and using quantitative analysis to support the results, this research makes new theoretical contributions. One recommendation for future research is the development of mitigation methods and the development of additional research lines to mitigate anomalies. Therefore, this research not only improves energy efficiency in Malang City but also provides useful information for applications elsewhere. It also makes the connection between research objectives and results stronger and opens the door for further exploration and engagement in the industry.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: This research received no external funding.

Acknowledgment: Universitas Negeri Malang.

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

References

[1] R. Gozalo-Brizuela, E. C. Garrido-Merchán, “A survey of Generative AI Applications,” *arXiv*, 2023, https://doi.org/10.48550/arXiv.2306.02781.

[2] K. M. Adavala *et al*., “Deep Generative Models for Data Synthesis and Augmentation in Machine Learning,” *Journal of Electrical Systems*, vol. 20, no. 3s, pp. 1242–1249, 2024, https://doi.org/10.52783/jes.1435.

[3] H. Sun, T. Zhu, Z. Zhang, D. Jin, P. Xiong and W. Zhou, "Adversarial Attacks Against Deep Generative Models on Data: A Survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 4, pp.

3367-3388, 2023, https://doi.org/10.1109/TKDE.2021.3130903.

[4] D. Reshetova, Wei-Ning Chen, A. Özgür, "Training generative models from privatized data," *arXiv*, 2023, https://doi.org/10.48550/arXiv.2306.09547.

[5] A. Grover, "Generative Decision Making Under Uncertainty," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, no. 13, pp. 15440-15440, 2023, https://doi.org/10.1609/aaai.v37i13.26807.

[6] B. Yilmaz, "Generative adversarial network for load data generation: Türkiye energy market case," *Mathematical Modelling and Numerical Simulation with Applications*, vol. 3, no. 2, pp. 141-158, 2023, https://doi.org/10.53391/mmnsa.1320914.

[7] L. Tausani, A. Testolin, and M. Zorzi, "Investigating the Generative Dynamics of Energy-Based Neural Networks," *Brain Informatics*, vol. 13974, pp. 96–108, 2023, https://doi.org/10.1007/978-3-031-43075-6_9.

[8] B. Irawan, W. Wirawan, B. A. Ikawanty, and A. Takwim, "Analysis of the season effect on energy generated from hybrid PV/WT in Malang Indonesia," *Eastern-European Journal of Enterprise Technologies*, vol. 5, no. 8(119), pp. 70-78, 2022, https://doi.org/10.15587/1729-4061.2022.266082.

[9] S. Pfenninger and I. Staffell, "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data," *Energy*, vol. 114, pp. 1251-1265, 2016, https://doi.org/10.1016/j.energy.2016.08.060.

[10] C. I. Cahyadi, S. Suwarno, A. A. Dewi, M. Kona, M. Arif, and M. C. Akbar, "Solar Prediction Strategy for Managing Virtual Power Stations," *International Journal of Energy Economics and Policy*, vol. 13, no. 4, pp. 503–512, 2023, https://doi.org/10.32479/ijeep.14124.

[11] M. N. Hidayat, D. N. Akbar, I. N. Syamsiana, and I. Ridzky, "Analysis of potential development of the hybrid power plants in coastal areas of Malang Regency-Indonesia," *AIP Conference Proceedings*, vol. 2255, no. 1, 2020, https://doi.org/10.1063/5.0014621.

[12] A. Khan, "Enhanced Evolutionary Sizing Algorithms for Optimal Sizing of a Stand-Alone PV-WT-Battery Hybrid System," *Applied Sciences*, vol. 9, no. 23, p. 5197, 2019, https://doi.org/10.3390/app9235197.

[13] S. M. Shahandashti, K. Mostaan, and B. Ashuri, "Methods for Assessing Longitudinal Variations of Energy Production by PV Systems," *AEI 2013: Building Solutions for Architectural Engineering*, pp. 113–122, 2013, https://doi.org/10.1061/9780784412909.012.

[14] Y. Kwon, S. Kim, Y.-S. Choi, and S. Kang, "Generative Modeling to Predict Multiple Suitable Conditions for Chemical Reactions," *Journal of Chemical Information and Modeling*, vol. 62, no. 23, pp. 5952-5960, 2022, https://doi.org/10.1021/acs.jcim.2c01085.

[15] H. Young, M. Du, O. Bastani, "Neurosymbolic Deep Generative Models for Sequence Data with Relational Constraints," *36th Conference on Neural Information Processing Systems*, 2021, https://proceedings.neurips.cc/paper_files/paper/2022/file/f13ceb1b94145aad0e54186373cc86d7-Paper-Conference.pdf.

[16] Y. Alnumay, A. J. Alrasheed, H. Trigui, A. Halawani, M. Alshiekh, and S. Patel, "Synthetic Data Generation for Machine Learning Applications in the Energy Industry," *Abu Dhabi International Petroleum Exhibition and Conference*, 2022, https://doi.org/10.2118/211821-MS.

[17] T. Vieijra, J. Haegeman, F. Verstraete, and L. Vanderstraeten, "Generative modeling with projected entangled-pair states," *Physical Review B*, vol. 104, no. 23, p. 235141, 2022, https://doi.org/10.1103/PhysRevB.104.235141.

[18] Y. Wang, M. L. Kamari, S. Haghighat, and P. T. T. Ngo, "Electrical and thermal analyses of solar PV module by considering realistic working conditions," *Journal of Thermal Analysis and Calorimetry*, vol. 144, pp. 1925-1934, 2021, https://doi.org/10.1007/s10973-020-09752-2.

[19] E. Sauter, M. Mughal, and Z. Zhang, "Evaluation of Machine Learning Methods on Large-Scale Spatiotemporal Data for Photovoltaic Power Prediction," *Energies*, vol. 16, no. 13, p. 4908, 2023, https://doi.org/10.3390/en16134908.

[20] V. Belov, "Optical Communication on Scattered or Reflected Laser Radiation," *Light & Engineering*, vol. 27, no. 1, pp. 15-24, 2019, https://doi.org/10.33383/2018-105.

[21] M. Lietzow and S. Wolf, “Scattered polarized radiation of extrasolar circumplanetary rings,” *Astronomy & Astrophysics*, vol. 671, no. A113, p. 12, 2023, https://doi.org/10.1051/0004-6361/202245474.

[22] A. Pradhan, B. Panda, L. Nanda, C. Jena and S. Sahoo, "Analysis of Various Types of Reflectors on The Performance of PV Panel," *2022 International Conference for Advancement in Technology (ICONAT)*, pp. 1-6, 2022, https://doi.org/10.1109/ICONAT53423.2022.9725825.

[23] J. R. Kumar, S. P. Mishra, P. P. Padhi, V. R. Krishna and P. K. Dash, "Real Time Comparative Analysis between Solar Fixed Plate and Tracking Methodology," *2022 2nd Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON)*, pp. 1-5, 2022, https://doi.org/10.1109/ODICON54453.2022.10010223.

[24] E. G. Pignaton, D. De Paula e Silva, F. B. B. D. Silva, J. L. F. Salles, and J. F. Fardin, “Validation of Photovoltaic Model for Application in A Distributed Energy Source Using Weather Data,” *Revista Contemporanea*, vol. 3, no. 6, pp. 6483-6496, 2023, https://doi.org/10.56083/RCV3N6-096.

[25] E. Olvera-Gonzalez, M. M. Rivera, N. Escalante-Garcia, and E. Flores-Gallegos, “Modeling Energy LED Light Consumption Based on an Artificial Intelligent Method Applied to Closed Plant Production System,” *Applied Sciences*, vol. 11, no. 6, p. 2735, 2021, https://doi.org/10.3390/app11062735.

[26] Y. Zhi, T. Sun, and X. Yang, “A physical model with meteorological forecasting for hourly rooftop photovoltaic power prediction,” *Journal of Building Engineering*, vol. 75, p. 106997, 2023, https://doi.org/10.1016/j.jobe.2023.106997.

[27] D. Kasprowicz and G. Kasprowicz, “Generative Modeling of Semiconductor Devices for Statistical Circuit Simulation,” *Electronics*, vol. 13, no. 11, p. 2003, 2024, https://doi.org/10.3390/electronics13112003.

[28] A. Patel and O. V. G. Swathika, “Day-Ahead Solar Power Forecasting Using Statistical and Machine Learning Methods,” *Integrated Green Energy Solutions Volume 2*, pp. 71–101, 2023, https://doi.org/10.1002/9781394193738.ch26.

[29] H. C. Bloomfield, D. J. Brayshaw, M. Deakin, and D. Greenwood, “Hourly historical and near-future weather and climate variables for energy system modelling,” *Earth System Science Data*, vol. 14, no. 6, pp. 2749–2766, 2022, https://doi.org/10.5194/essd-14-2749-2022.

[30] F. A. Toro-Cardona, J. L. Parra, and O. R. Rojas-Soto, “Predicting daily activity time through ecological niche modelling and microclimatic data,” *Journal of Animal Ecology*, vol. 92, no. 4, pp. 925-935, 2023, https://doi.org/10.1111/1365-2656.13895.

[31] T. Klinsuwan, W. Ratiphaphongthon, R. Wangkeeree, R. Wangkeeree, C. Sirisamphanwong, “Evaluation of Machine Learning Algorithms for Supervised Anomaly Detection and Comparison between Static and Dynamic Thresholds in Photovoltaic Systems,” *Energies*, vol. 16, no. 4, p. 1947, 2023, https://doi.org/10.3390/en16041947.

[32] H. Li, K. Liu, C. Chen and L. Chen, "An Anomalous Data Identification Approach for PV Generation Based on Quadratic and Intra-group Variance Method," *2023 IEEE 6th International Electrical and Energy Conference (CIEEC)*, pp. 3294-3299, 2023, https://doi.org/10.1109/CIEEC58067.2023.10166874.

[33] T. Peng and F. Wen, "Photovoltaic Panel Fault Detection Based on Improved Mask R-CNN," *2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT)*, pp. 1187-1191, 2023, https://doi.org/10.1109/ICCECT57938.2023.10140742.

[34] H. D. Tafti *et al.*, "Extended Functionalities of Photovoltaic Systems With Flexible Power Point Tracking: Recent Advances," *IEEE Transactions on Power Electronics*, vol. 35, no. 9, pp. 9342-9356, 2020, https://doi.org/10.1109/TPEL.2020.2970447.

[35] L. Abduh, L. Omar, and I. Ivrissimtzis, “Anomaly Detection with Transformer in Face Anti-spoofing,” *Journal of WSCG*, vol. 31, no. 1-2, pp. 91-98, 2023, https://doi.org/10.24132/JWSCG.2023.10.

[36] A. P. Wibawa *et al.*, “Mean-Median Smoothing Backpropagation Neural Network to Forecast Unique Visitors Time Series of Electronic Journal,” *Journal of Applied Data Sciences*, vol. 4, no. 3, pp. 163-174, 2023, https://doi.org/10.47738/jads.v4i3.97.

[37] T. Park, K. Song, J. Jeong, and H. Kim, “Convolutional Autoencoder-Based Anomaly Detection for Photovoltaic Power Forecasting of Virtual Power Plants,” *Energies*, vol. 16, no. 14, p. 5293, 2023,

https://doi.org/10.3390/en16145293.

[38] X. Xiao, Z. Yang, and X. Gao, "Anomaly Detection of Power Time-Series Data Based on Multi-Dimensional Transformer Network," *Computer-Aided Design & Applications*, vol. 21, no. S7, pp. 15-27, 2023, https://doi.org/10.14733/cadaps.2024.S7.15-27.

[39] H. Tang, Q. Wang, G. Jiang, "Retracted: Time Series Anomaly Detection Model Based on Multi-Features," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, 2023, https://doi.org/10.1155/2023/9820841.

[40] A. W. Saputra, A. P. Wibawa, U. Pujianto, A. B. Putra Utama, and A. Nafalski, "LSTM-based Multivariate Time-Series Analysis: A Case of Journal Visitors Forecasting," *ILKOM Jurnal Ilmiah*, vol. 14, no. 1, pp. 57-62, 2022, https://doi.org/10.33096/ilkom.v14i1.1106.57-62.

[41] M. B. Rahmoune, A. Iratni, A. S. Amari, A. Hafaifa, and I. Colak, "Fault detection and diagnosis of photovoltaic system based on neural networks approach," *Diagnostyka*, vol. 24, no. 3, pp. 1–10, 2023, https://doi.org/10.29354/diag/166428.

[42] F. Harrou, Y. Sun, A. Dorbane and B. Bouyeddou, "Sensor fault detection in photovoltaic systems using ensemble learning-based statistical monitoring chart," *2023 11th International Conference on Smart Grid (icSmartGrid)*, pp. 1-6, 2023, https://doi.org/10.1109/icSmartGrid58556.2023.10170985.

[43] A. Chaudhary and R. Agarwal, "Machine Learning Techniques for Anomaly Detection Application Domains," *Paradigms of Smart and Intelligent Communication, 5G and Beyond*, pp. 129-147, 2023, https://doi.org/10.1007/978-981-99-0109-8_8.

[44] U. Patole, N. Gorhe, K. Chavan, P. Kulkarni, J. Dale, "Anomaly Detection System," *International journal of Scientific Research in Engineering and Management*, vol. 08, no. 05, pp. 1-4, 2024, https://doi.org/10.55041/IJSREM32871.

[45] R. J. Nayak, J. P. Chaudhari, "Anomaly detection using deep learning based model with feature attention," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 1, pp. 383-390, 2024, http://doi.org/10.11591/ijai.v13.i1.pp383-390.

[46] E. Gilabert, E. Jantunen, C. Emmanouilidis, A. Starr, and A. Arnaiz, "Optimizing E-Maintenance Through Intelligent Data Processing Systems," *Engineering Asset Management 2011*, pp. 1-9, 2014, https://doi.org/10.1007/978-1-4471-4993-4_1.

[47] S. Sucipto, D. Dwi Prasetya, and T. Widiyaningtyas, "Educational Data Mining: Multiple Choice Question Classification in Vocational School," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 23, no. 2, pp. 379-388, 2024, https://doi.org/10.30812/matrik.v23i2.3499.

[48] P. K. Malone, "Developing Realistic Schedule Risk Impacts," *2023 IEEE Aerospace Conference*, pp. 1-16, 2023, https://doi.org/10.1109/AERO55745.2023.10115766.

[49] D. Andriani, M. F. Fakhrullah and M. Y. Syafei, "Proposed Maintenance Scheduling Using Failure Mode and Effect Analysis and Monte Carlo Methods on Critical Components," *2023 9th International Conference on Signal Processing and Intelligent Systems (ICSPIS)*, pp. 1-6, 2023, https://doi.org/10.1109/ICSPIS59665.2023.10402655.

[50] P. Chiranjeevi, Yadavalli Ramya, Chinthala Balaji, Bathini Shashank, and Abbdi Sainath Reddy, "Uncovering time series anomaly using deep learning technique," *World Journal of Advanced Research and Reviews*, vol. 22, no. 1, pp. 879-887, 2024, https://doi.org/10.30574/wjarr.2024.22.1.1129.

[51] S. K. Dani, C. Thakur, N. Nagvanshi and G. Singh, "Anomaly Detection using PCA in Time Series Data," *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, pp. 1-6, 2024, https://doi.org/10.1109/IATMSI60426.2024.10502929.

[52] S. K. Perepu and V. S. Pinnamaraju, "A novel unsupervised method for root cause analysis of anomalies using sparse optimization techniques," *2022 10th International Conference on Systems and Control (ICSC)*, pp. 416-422, 2022, https://doi.org/10.1109/ICSC57768.2022.9993819.

[53] M. Fahim and A. Sillitti, "Anomaly Detection, Analysis and Prediction Techniques in IoT Environment: A Systematic Literature Review," *IEEE Access*, vol. 7, pp. 81664-81681, 2019, https://doi.org/10.1109/ACCESS.2019.2921912.

[54] S. Schmidl, P. Wenig, and T. Paperbrock, "Anomaly detection in time series: a comprehensive evaluation," *Proceedings of the VLDB Endowment*, vol. 15, no. 9, pp. 1779-1797, 2022, https://doi.org/10.14778/3538598.3538602.

[55] A. A. Cook, G. Mısırlı and Z. Fan, "Anomaly Detection for IoT Time-Series Data: A Survey," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6481-6494, 2020, https://doi.org/10.1109/JIOT.2019.2958185.

[56] A. Pranolo *et al.*, "Exploring LSTM-based Attention Mechanisms with PSO and Grid Search under Different Normalization Techniques for Energy demands Time Series Forecasting," *Knowledge Engineering and Data Science*, vol. 7, no. 1, pp. 1-12, 2024, http://dx.doi.org/10.17977/um018v7i12024p1-12.

[57] D. M. Hawkins, "Identification of Outliers," *Springer Dordrecht*, 1980, https://doi.org/10.1007/978-94-015-3994-4.

[58] L. Donghua, Yang., Kaiqi, Zhang., Hong-wei, Gao., Jianzhong, "Anomaly and change point detection for time series with concept drift," *World Wide Web*, vol. 26, pp. 3229-3252, 2023, https://doi.org/10.1007/s11280-023-01181-z.

[59] J. Van Zyl *et al.*, "Subspace based Anomaly Detection Framework for Point Clouds," *2022 IEEE 18th International Conference on e-Science (e-Science)*, pp. 316-325, 2022, https://doi.org/10.1109/eScience55777.2022.00045.

[60] K. Chen, M. Feng, T. S. Wirjanto, "Time-series Anomaly Detection via Contextual Discriminative Contrastive Learning," *arXiv*, 2023, https://doi.org/10.48550/arXiv.2304.07898.

[61] F. Xu, N. Wang, X. Zhao, "Exploring Global and Local Information for Anomaly Detection with Normal Samples," *arXiv*, 2023, https://doi.org/10.48550/arXiv.2306.02025.

[62] G. S. Fuhnwi, J. O. Agbaje, K. Oshinubi, O. J. Peter, "An Empirical Study on Anomaly Detection Using Density-based and Representative-based Clustering Algorithms," *Journal of the Nigerian Society of Physical Sciences*, vol. 5, no. 2, pp. 1-13, 2023, https://doi.org/10.46481/jnsps.2023.1364.

[63] Y. Bao, Z. Tang, H. Li, and Y. Zhang, "Computer vision and deep learning–based data anomaly detection method for structural health monitoring," *Structural Health Monitoring*, vol. 18, no. 2, pp. 401–421, 2019, https://doi.org/10.1177/1475921718757405.

[64] Z. Tang, Z. Chen, Y. Bao, and H. Li, "Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring," *Structural Control and Health Monitoring*, vol. 26, no. 1, p. e2296, 2019, https://doi.org/10.1002/stc.2296.

[65] A. C. Jackson, S. Lacey, "The discrete Fourier transformation for seasonality and anomaly detection of an application to rare data," *Data Technologies and Applications*, vol. 54, no. 2, pp. 121-132, 2020, https://doi.org/10.1108/DTA-12-2019-0243.

[66] S. Cai, R. Sun, H. Mu, X. Shi & G. Yuan, "A Minimum Rare-Itemset-Based Anomaly Detection Method and Its Application on Sensor Data Stream," *Computer Supported Cooperative Work and Social Computing*, vol. 1042, pp. 116-130, 2019, https://doi.org/10.1007/978-981-15-1377-0_9.

[67] Z. Tian, M. Zhuo, L. Liu, J. Chen and S. Zhou, "Anomaly detection using spatial and temporal information in multivariate time series," *Scientific* Reports, vol. 13, no. 4400, 2023, https://doi.org/10.1038/s41598-023-31193-8.

[68] G. Wang *et al.*, "Anomaly Detection for Data from Unmanned Systems via Improved Graph Neural Networks with Attention Mechanism," *Drones*, vol. 7, no. 5, p. 326, 2023, https://doi.org/10.3390/drones7050326.

[69] A. Vaswani *et al.*, "Attention is all you need," *arXiv*, 2017, https://doi.org/10.48550/arXiv.1706.03762.

[70] A. Patel *et al.*, "Cross Attention Transformers for Multi-modal Unsupervised Whole-Body PET Anomaly Detection," *Deep Generative Models*, vol. 13609, pp. 14-23, 2023, https://doi.org/10.1007/978-3-031-18576-2_2.

[71] P. Guo, C. Wei and Z. Yin, "Anomaly Detection of Spacecraft Reconstructed Signals Based on Attention Mechanism," *2023 6th International Symposium on Autonomous Systems (ISAS)*, pp. 1-5, 2023, https://doi.org/10.1109/ISAS59543.2023.10164514.

[72] S. Shleifer, J. Weston, M. Ott, "NormFormer: Improved Transformer Pretraining with Extra Normalization," *arXiv*, 2021, https://doi.org/10.48550/arXiv.2110.09456.

[73] G. Yagawa, A. Oishi, "Feedforward Neural Networks," *Computational Mechanics with Neural Networks*, pp. 11-23, 2021, https://doi.org/10.1007/978-3-030-66111-3_2.

[74] C. Yin, S. Zhang, J. Wang and N. N. Xiong, "Anomaly Detection Based on Convolutional Recurrent Autoencoder for IoT Time Series," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 1, pp. 112-122, 2022, https://doi.org/10.1109/TSMC.2020.2968516.

[75] J. Xu, H. Wu, J. Wang, and M. Long, "Anomaly Transformer: Time Series Anomaly Detection with Association Discrepancy," *arXiv*, 2022, https://doi.org/10.48550/arXiv.2110.02642.

[76] K.-H. Lai, D. Zha, J. Xu, Y. Zhao, G. Wang, X. Hu, "Revisiting time series outlier detection: Definitions and benchmarks," *35th Conference on Neural Information Processing* Systems, pp. 1-13, 2021, https://datasets-benchmarks-proceedings.neurips.cc/paper_files/paper/2021.

[77] D. Rout, A. Kotangale, S. Nath and B. Roy, "An Association Based Approach to Elicit and Measure Impact of Features on Sales of a Garment Retail," *2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1)*, pp. 1-6, 2023, https://doi.org/10.1109/ICAIA57370.2023.10169499.

[78] A. Tong, G. Wolf and S. Krishnaswamy, "Fixing Bias in Reconstruction-based Anomaly Detection with Lipschitz Discriminators," *Journal of Signal Processing Systems*, vol. 94, pp. 229-243, 2022, https://doi.org/10.1007/s11265-021-01715-6.

[79] S. Chauhan and S. Lee, "Machine Learning-Based Anomaly Detection for Multivariate Time Series With Correlation Dependency," *IEEE Access*, vol. 10, pp. 132062-132070, 2022, https://doi.org/10.1109/ACCESS.2022.3230352.

[80] K. Shaukat *et al.*, "A Review of Time-Series Anomaly Detection Techniques: A Step to Future Perspectives," *Advances in Information and Communication*, pp. 865-877, 2021, https://doi.org/10.1007/978-3-030-73100-7_60.

[81] M. H. Krishna, N. K, G. Charmitha, T. Vignesh, V. Ch and S. Kuchibhotla, "Studies on Anomaly Detection Techniques," *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 813-817, 2023, https://doi.org/10.1109/ICCMC56507.2023.10083885.