Article

Optimization of photovoltaic panels through machine learning algorithms linked to predictive maintenance

Optimización de paneles fotovoltaicos mediante algoritmos de aprendizaje automático vinculándolo con el área de mantenimiento predictivo

Ruiz-Garduño, Jhacer Kharen[a], Flores-Serrato, Leonel[b], González-Ramírez, Claudia Teresa*[c] and Viñas-Álvarez, Samuel Efrén[d]

[a] ROR National Technological Institute of Mexico • KSL-5132-2024 • 0000-0003-3353-7966 • 764417
[b] ROR National Technological Institute of Mexico • LGY-4995-2024 • 0009-0004-6555-0051
[c] ROR National Technological Institute of Mexico • G-6313- 2019 • 0000-0002-4106-4583 • 425737
[d] ROR National Technological Institute of Mexico • 5107-2024 • 0000-0001-5891-2801• 606583

CONAHCYT classification:

Area: Engineering
Field: Technological sciences
Discipline: Telecommunications technology
Subdiscipline: Photoelectric devices

https://doi.org/10.35429/EJRC.2024.10.18.5.13

History of the article:
Received: February 16, 2024
Accepted: December 14, 2024

* ✉ [claudia.gr@zitacuaro.tecnm.mx]

Abstract

Zitácuaro, Michoacán, faces the challenge of maximizing the performance of photovoltaic systems. This study addresses this challenge through the development and implementation of advanced predictive models based on machine learning and predictive maintenance strategies. Historical and real-time data from photovoltaic systems are collected, including meteorological variables, panel performance, and energy consumption. Advanced analysis techniques, such as linear regression, decision trees, and neural networks, are applied to forecast energy production and detect patterns of suboptimal performance. The methodology is divided into two stages: comprehensive data collection from sources like the National Meteorological Service and OpenWeatherMap, and detailed analysis through preprocessing techniques and predictive modeling. The results demonstrate significant improvements in prediction accuracy and failure reduction, optimizing energy efficiency and contributing to the sustainability of photovoltaic systems in the region.

Enhancing Photovoltaic System Performance in Zitácuaro, Michoacán

Objectives	Methodology	Contribution
* Develop Predictive Models * Implement predictive maintenance strategies *Improve Energy Efficiency *Enhance Data Utilization	*Data Collection *Predictive Modeling *Performance Optimization	*Enhanced Prediction Accuracy *Reduced System Failures *Increased Energy Efficiency *Methodological Framework

Predictive, Optimization, Neural

Resumen

La creciente adopción de tecnologías solares en Zitácuaro, Michoacán, presenta el desafío de maximizar el rendimiento de los sistemas fotovoltaicos. Este estudio aborda este reto mediante el desarrollo e implementación de modelos predictivos avanzados basados en aprendizaje automático y estrategias de mantenimiento predictivo. Se recopilan datos históricos y en tiempo real de sistemas fotovoltaicos, que incluyen variables meteorológicas, rendimiento de paneles y consumo energético. Se aplican técnicas avanzadas de análisis, como regresión lineal, árboles de decisión y redes neuronales, para prever la producción energética y detectar patrones de rendimiento sub óptimo. La metodología se divide en dos etapas: la recolección exhaustiva de datos de fuentes como el Servicio Meteorológico Nacional y OpenWeatherMap, y el análisis detallado mediante técnicas de pre procesamiento y modelado predictivo. Los resultados demuestran mejoras significativas en la precisión de predicciones y reducción de fallos, optimizando la eficiencia energética y contribuyendo a la sostenibilidad de los sistemas fotovoltaicos en la región.

Mejorando el Rendimiento de los Sistemas Fotovoltaicos en Zitácuaro, Michoacán

Objetivos	Metodología	Contribución
*Desarrollar Modelos Predictivos *Implementar Estrategias de Mantenimiento Predictivo *Mejorar la Eficiencia Energética *Potenciar la Utilización de Datos	*Recolección de Datos *Modelado Predictivo *Optimización del Rendimiento	*Precisión Mejorada en las Predicciones *Reducción de Fallos en el Sistema *Aumento de la Eficiencia Energética *Framework Metodológico

Predictivo, Optimización, Neural

Citation: Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.

Introduction

The growing demand for sustainable energy solutions has driven interest in optimizing photovoltaic systems, especially in regions like Zitácuaro, Michoacán, where the implementation of solar panels has become a priority to reduce dependency on non-renewable sources.

However, despite significant efforts in the installation of these systems, it has been observed that the performance of solar panels does not always reach the expected optimal levels. This underperformance can largely be attributed to the lack of effective optimization and predictive maintenance strategies.

In this context, the present research proposes an integrated approach that combines machine learning algorithms with predictive maintenance strategies to improve the efficiency of photovoltaic systems. The methodology employed in this study combines descriptive and applied methods, based on a thorough documentary review and experimental research. Historical and real-time data are collected from a photovoltaic system located in Zitácuaro, and predictive models are developed using advanced techniques such as linear regression, decision trees, and neural networks.

Previous studies have demonstrated the importance of using historical and real-time data to improve the efficiency of photovoltaic systems. The Solar Comet Project – CS, led by Cecilia E. Sandoval-Ruiz, uses meteorological data analysis and advanced mathematical models to optimize solar energy production.

Similarly, Guillermo García Dávila and Miguel Ángel Hernández Flores have addressed the optimization of photovoltaic systems through simulations and computer tools to maximize efficiency and reduce operational costs. These approaches highlight the need to integrate environmental data and maintenance strategies to achieve optimal operation of solar panels.

The research also faces significant challenges in Zitácuaro, such as the variability of environmental conditions and the accumulation of dirt on the panels.

Traditional maintenance methods are often reactive and based on predefined time intervals, which can result in ineffective interventions. The use of machine learning algorithms allows for the analysis of large volumes of data to predict the behavior of solar panels and optimize maintenance routines, ensuring optimal performance for as long as possible.

This study aims to address these challenges by developing a predictive model and a maintenance system based on machine learning, with the goal of improving energy efficiency and reducing the operational costs of solar panels in Zitácuaro. The integration of meteorological data and predictive analysis is fundamental to achieving these objectives and contributing to the region's energy sustainability

Methodology

The research proposes an integrated approach to optimize photovoltaic systems through the use of machine learning algorithms and predictive maintenance strategies.

The methodology employed combines descriptive and applied methods, with a solid foundation in document review and experimental research. Historical and real-time data from a photovoltaic system located in Zitácuaro, Michoacán, are collected, and predictive models are developed using machine learning techniques such as linear regression, decision trees, and neural networks.

The review of previous studies highlighted the importance of using historical and real-time data to improve the efficiency of photovoltaic systems.

In Zitácuaro, Michoacán, the search for sustainable and efficient energy generation solutions, optimizing photovoltaic systems emerges as a critical necessity.

Various studies have explored innovative approaches to maximize solar panel efficiency and reduce operating costs through the use of machine learning algorithms and predictive maintenance strategies.

The optimization of photovoltaic systems is a field that has significantly evolved in recent years.

ISSN: 2414-4959.
RENIECYT-CONAHCYT: 1702902

Neama et al. (2023) emphasize the importance of renewable energy in sustainable development, highlighting how the optimization of these systems can maximize energy efficiency. David et al. (2022) point out that the implementation of optimized energy systems in rural areas not only improves electrification but also promotes economic and social development.

Artificial neural networks (ANN) and other machine learning algorithms have proven to be effective tools for predicting and optimizing the performance of photovoltaic systems. Cardil et al. (2015) implemented prediction models based on neural networks for the spread of forest fires, demonstrating their potential in renewable energy applications. Pereira et al. (2022) used neural networks for solar resource assessment, showing significant improvements in prediction accuracy. Predictive maintenance is a crucial strategy to ensure the efficiency and longevity of photovoltaic systems. Rosado et al. (2022) demonstrated how neural networks can predict the mass of peach fruits using non-destructive methods, a technique applicable to the monitoring and maintenance of solar panels.

Abbass (2023) proposed a framework based on Bayesian optimization and neural network connections to predict the energy performance of buildings, highlighting its application in photovoltaic systems.

Several studies have used simulation models to evaluate and improve the efficiency of photovoltaic systems. Osman et al. (2023) conducted sub-seasonal to decadal predictions to support climate services, providing a solid foundation for energy system planning and optimization. Beccali et al. (2017) developed a decision support tool based on ANN to assess energy performance and rehabilitation actions in non-residential buildings.

The practical application of these methodologies has been explored in various contexts. Báez Coronado (2024) designed an artificial intelligence model to predict solar energy production, applying it in the context of Antioquia, Colombia. Sandoval-Ruiz (2020) and Davila (2023) examined the optimization of photovoltaic systems in specific projects, highlighting the importance of integrating advanced technologies to improve efficiency and sustainability.

The accumulation of dirt on solar panels can significantly reduce their efficiency. González (2022) and Cobian Lufin (2023) investigated optimization models for cleaning planning in solar parks, proposing strategies that combine simulation techniques and machine learning models to keep panels in optimal condition.

The implementation of digital twins for the optimization of self-sufficient energy systems is an emerging trend. Rodríguez de Lope López et al. (2023) developed a digital twin for a residential self-sufficient energy system, demonstrating how these technologies can improve the management and efficiency of energy resources.

The state of the art in the optimization of photovoltaic panels using machine learning algorithms and predictive maintenance strategies reveals a dynamic and constantly evolving landscape. Advances in neural networks, simulation methods, and practical applications highlight the potential of these technologies to transform solar energy management, promoting greater efficiency and sustainability in the use of renewable resources.

Problem Description

In Zitácuaro, Michoacán, the implementation of photovoltaic solar panels has been a key initiative to promote the use of renewable energy and reduce dependence on non-sustainable energy sources. However, despite advances in the installation of these systems, their performance has not always reached the optimal levels expected.

This underperformance is largely due to the lack of effective optimization and predictive maintenance strategies. Solar panels are highly dependent on environmental conditions such as solar irradiation, temperature, humidity, and wind speed.

These factors can vary significantly over time, affecting the efficiency of converting solar energy into electricity. Additionally, the accumulation of dirt and other physical obstructions can further reduce the panels' ability to generate electricity.

Without adequate maintenance, these problems can lead to a considerable decrease in energy production, resulting in higher operating costs and a shorter lifespan for the panels. However, most traditional maintenance methods are reactive and based on predefined time intervals, which can result in unnecessary or late interventions.

On the other hand, optimizing panel configuration to maximize their energy efficiency is not always carried out systematically and data-driven, limiting the energy generation potential. The implementation of machine learning algorithms offers a promising solution to these problems.

These algorithms can analyze large volumes of historical and real-time data to identify patterns and predict the behavior of solar panels under various conditions. Valverde, L. et al (2023).

By applying machine learning techniques, it is possible to develop predictive models that anticipate maintenance needs before failures occur, thereby optimizing maintenance routines and ensuring that panels operate at their maximum capacity for as long as possible.

Despite the potential of these technologies, their practical application in Zitácuaro still faces several challenges (social, economic, geographical). The collection of accurate and real-time data, the selection and training of appropriate algorithms, and the implementation of an effective early warning system are critical aspects that require careful research and development.

This study aims to address these challenges by developing a predictive model and a maintenance system based on machine learning, with the objective of improving energy efficiency and reducing the operating costs of solar panels in Zitácuaro, Michoacán.

Data Collection from Local and National Sources:

- National Meteorological Service (SMN)
- National Forestry Commission (CONAFOR)
- National Water Commission (CONAGUA)

The reports provided by these organizations contain historical data related to meteorological conditions, such as radiation, temperature, and humidity, as these factors are relevant.

Determination of Energy Needs

First, the devices requiring electrical power were identified, including in our case study of a building: 25 LED lamps, 17 computers, and 6 printers. The energy consumption of each type of device was calculated based on its nominal power and daily operating time.

Real-time Data

- Solar Energy Information System (SIES): Provides real-time data on solar radiation and weather conditions in various regions of Mexico.
- Meteorological Data Platforms

- OpenWeatherMap: Provides current weather data and forecasts for any location.
- Weather Underground: Provides real-time weather data as well as forecasts for Zitácuaro.

Optimizing the Performance of a Photovoltaic System in Zitácuaro by Analyzing Historical and Real-time Data

Data Sources

- National Meteorological Service (SMN): Provides historical solar radiation and meteorological data in Mexico, can provide specific data for Zitácuaro.
- NASA's Surface Meteorology and Solar Energy (SSE): Offers historical data on solar irradiation and meteorological conditions. The SSE portal can filter data on solar irradiation and temperature for Zitácuaro.

Process

Collected Data

- Historical Solar Radiation: Monthly average solar irradiation in kWh/m² for the last 5 years.
- Historical Temperature: Monthly average temperature in °c during the same period.

Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.
https://doi.org/10.35429/EJRC.2024.10.18.5.13

- Humidity and Wind Speed: Monthly average data.

Real-time Data Collection

- OpenWeatherMap API: Provides current weather data and forecasts (solar radiation, temperature, humidity, and wind speed) for Zitácuaro.
- Solar Energy Information System (SIES): Offers real-time data on solar conditions in Mexico.

Collected Data

- Real-time Energy Generation: Data on energy generation from polycrystalline panels.
- Current Solar Radiation: Real-time measurements of solar irradiation in W/m².
- Current Environmental Conditions: Real-time temperature, humidity, and wind speed.

Data Integration and Analysis

Tools and Methods

- Database: SQL (like MySQL) or NoSQL (like MongoDB) to store historical and real-time data.
- Analysis Software: PVsyst to analyze and simulate the performance of polycrystalline solar panels.
- Visualization Tools: Tableau or Power BI to create dashboards and charts to visualize panel efficiency and correlate weather data with energy production.

Analysis Process

- Comparison of Historical vs. Real-time Data: Compare current energy production of polycrystalline panels with historical solar irradiation data to identify patterns or discrepancies.
- Environmental Conditions: Analyze how current environmental conditions affect panel performance compared to historical data.

Optimization

Operational Adjustments: Based on the analysis, make adjustments to the orientation or tilt of the panels or maintenance scheduling.

Predictive Models: Machine learning models to predict energy production and adjust the system according to forecasted conditions.

Database Creation for Research

Steps

1. Define Requirements: Identify the types of data needed, such as solar irradiation, temperature, humidity, wind speed, and panel performance. Organize data in the database (tables, relationships, etc.).

2. Select Database Technology: Use relational databases like MySQL or PostgreSQL for highly structured data and complex relationships. Use NoSQL databases like MongoDB for flexibility in the data schema or unstructured data.

3. Collect Data: Gather data from mentioned sources (Sistema Meteorológico Nacional (SMN), Administración Nacional de Aeronáutica y del Espacio(NASA), etc.) and store in the database. Consider automating real-time data collection using APIs or monitoring systems.

4. Design and Configure Database: Design a data model including tables for solar irradiation, temperature, humidity, wind speed, and panel performance. Configure indices to improve query performance and create SQL queries or scripts for data extraction.

5. Implement and Maintain Database: Develop the database and enter historical and real-time Conduct Periodic Maintenance to Ensure Data Integrity and Accuracy

Data Analysis

- Utilize data analysis tools such as Excel, Tableau, Power BI, or programming languages like Python to extract and analyze stored data.
- Implement a predictive model and machine learning algorithms to optimize the performance of solar panels.

ISSN: 2414-4959.
RENIECYT-CONAHCYT: 1702902

Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.
https://doi.org/10.35429/EJRC.2024.10.18.5.13

For this database process, a variety of key data must be collected:

Meteorological Data

Solar Radiation:

- Daily/Monthly/Annual Average: Measured in kWh/m² or W/m².
- Specific Data: Direct, diffuse, and global irradiation.

Ambient Temperature:

- Daily/Monthly/Annual Average: Measured in °C.
- Range: Maximum and minimum daily temperature.

Relative Humidity:

- Daily/Monthly/Annual Average: Measured in %.
- Additional Data: Minimum and maximum humidity.

Wind Speed:

- Daily/Monthly/Annual Average: Measured in m/s or km/h.
- Additional Data: Maximum and minimum speed.

Atmospheric Pressure:

- Daily/Monthly/Annual Average: Measured in hPa or mmHg.
- Additional Data: Pressure variations that may affect production.

Weather Conditions:

- Precipitation: Measured in mm or liters per square meter.
- Cloud Cover: Percentage of sky covered or type of cloudiness.

Solar Panel Performance Data

Energy Generation:

- Daily/Monthly/Annual Production: Measured in kWh.
- Additional Data: Total and per panel production.

Conversion Efficiency:

- Instantaneous and Average Efficiency: Measured in %.

Panel Condition:

- Panel Temperature: Measured in °C.
- Shade and Obstruction: Data on shaded or dirty areas.

Tilt and Orientation Angle:

- Specific Data: Tilt angle and orientation with respect to north.

Maintenance and Failures:

- Maintenance History: Dates and types of maintenance performed.
- Incidents and Failures: Record of failures and downtime.

Installation System Data

Installation Characteristics:

- Panel Type: Polycrystalline, monocrystalline, etc.
- Inverter Specifications: Type and capacity.
- System Configuration: Series/parallel and panel arrangement.

Installation Conditions:

- Location: Exact geographic coordinates.
- Surrounding Areas: Description of possible obstructions or reflections.

Operational Data

Energy Consumption Data:

- Daily/Monthly/Annual Consumption: Measured in kWh.
- Consumption Profile: Hours of highest and lowest demand.

Economic Data:

- Maintenance Cost: Costs associated with maintenance and repairs.
- Energy Cost: Cost per kWh produced and savings generated.

ISSN: 2414-4959.
RENIECYT-CONAHCYT: 1702902

Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.
https://doi.org/10.35429/EJRC.2024.10.18.5.13

Comparison and Validation Data

Benchmarking Data:

- Comparison with Similar Systems: Performance data of photovoltaic systems in similar conditions.

Historical Project Data:

- Previous Projects: Historical data from similar projects for model and technique validation.

Selection of Machine Learning Algorithms

The following machine learning algorithms are evaluated and selected for modeling:

- Linear Regression: Used to predict energy production based on environmental variables. Suitable for establishing simple linear relationships between variables.
- Decision Trees: Applied to model decisions based on multiple variables. Ideal for capturing non-linear relationships and providing clear interpretations of factors affecting energy production.
- Neural Networks: Employed to capture complex, non-linear relationships between variables. Useful when relationships between variables are highly complex.

Model Training

For each selected algorithm, training is performed using the training data set:

Linear Regression:

- Fit the model to historical energy production and environmental conditions.
- Evaluate the model's ability to predict energy production based on environmental data.

Decision Trees:

- Construct the decision tree using the training data set.
- Adjust tree parameters to improve accuracy and avoid overfitting.

Neural Networks:

- Design the neural network architecture, including the number of layers and neurons.
- Train the model using optimization algorithms like gradient descent.
- Adjust hyperparameters to improve model performance.

Model Validation and Evaluation

After training, each model is validated and evaluated using the test data set. The evaluation metrics considered are:

- Mean Squared Error (MSE): Measures the average squared difference between the model's predictions and actual values.
- Coefficient of Determination (R^2): Evaluates the proportion of variability in the dependent variable explained by the model.
- Precision and Recall: Especially relevant for classification if a classification approach is used instead of regression.

Selection of the Best Model

The performance metrics of all evaluated models are compared. The model with the best combination of accuracy, lowest error, and generalization ability is selected for final implementation.

Model Implementation

The selected model is implemented in the photovoltaic system for real-time prediction and optimization. Monitoring mechanisms are established to evaluate model performance and adjust the model as necessary based on new data.
Optimization of a Photovoltaic System's Performance

- Objective: Predict energy production and optimize the configuration of solar panels using machine learning.
- Data: Energy production (kWh/day), solar radiation (kWh/m²), temperature (°C), humidity (%), wind speed (m/s).

ISSN: 2414-4959.
RENIECYT-CONAHCYT: 1702902

Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.
https://doi.org/10.35429/EJRC.2024.10.18.5.13

Algorithm Selection:

- Linear Regression to establish a basic relationship between solar radiation and energy production.
- Decision Trees to model complex interactions between multiple environmental variables.
- Neural Networks to capture complex patterns in energy production and optimize system configuration.

Training and Evaluation:

- Models are trained with historical data over a 5-year period.
- Model performance is validated with a recent data set.

Implementation:

- The best-performing model is used to adjust panel angles and improve real-time efficiency.

Results

Development and Evaluation of Predictive Models

1. Data Collection:
- Dependent Variables: Energy production (kWh/day).
- Independent Variables: Solar radiation, temperature, humidity, wind speed.

2. Data Preprocessing:
- Data Cleaning: Remove or impute missing values and handle outliers.
- Normalization/Standardization: Scale variables to improve model efficiency, especially for neural networks.

3. Data Splitting:
- Training Set: Use approximately 70-80% of the data.
- Test Set: Use the remaining 20-30% to evaluate model accuracy.

4. Development of Predictive Models:
- Linear Regression:

```
from sklearn.linear_model import
LinearRegression
from sklearn.metrics import
mean_squared_error, r2_score
# Definir el modelo
modelo_lineal = LinearRegression()
 # Entrenar el modelo
modelo_lineal.fit(X_train, y_train)
# Hacer predicciones
y_pred_lineal =
modelo_lineal.predict(X_test)
# Evaluar el modelo
mse_lineal =
mean_squared_error(y_test,
y_pred_lineal) r2_lineal =
r2_score(y_test, y_pred_lineal)
print(f"Regresión Lineal - MSE:
{mse_lineal}, R²: {r2_lineal}”)
```

Decision Trees

```
from sklearn.tree import
DecisionTreeRegressor
# Definir el modelo
modelo_decision =
DecisionTreeRegressor()
# Entrenar el modelo
modelo_decision.fit(X_train, y_train)
# Hacer predicciones
y_pred_decision =
modelo_decision.predict(X_test)
# Evaluar el modelo
mse_decision =
mean_squared_error(y_test,
y_pred_decision)
r2_decision = r2_score(y_test,
y_pred_decision)
print(f"Árboles de Decisión - MSE:
{mse_decision}, R²: {r2_decision}")
```

Neural Networks

```
from sklearn.neural_network import
MLPRegressor
# Definir el modelo
modelo_nn =
MLPRegressor(hidden_layer_sizes=(100,),
max_iter=1000)
# Entrenar el modelo
modelo_nn.fit(X_train, y_train)
# Hacer predicciones
y_pred_nn = modelo_nn.predict(X_test)
# Evaluar el modelo
mse_nn = mean_squared_error(y_test,
y_pred_nn)
r2_nn = r2_score(y_test, y_pred_nn)
print(f"Redes Neuronales - MSE: {mse_nn}, R²:
{r2_nn}")
```


Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.
https://doi.org/10.35429/EJRC.2024.10.18.5.13

Mean Squared Error (MSE):

Description

Measures the average of the squares of the errors, i.e., the average difference between actual values and predictions.

Formula: $MSE=1 \sum {}^{n}{}_{i=1} n(yi-yi\hat{})2$

A lower MSE value indicates better model performance.

Coefficient of Determination (R^2):

Measures the proportion of variance in the dependent variable that is explained by the independent variables in the model.

Formula: $R2=1-\frac{\sum_{i=1} n\ (yi-\bar{y})2}{\sum_{i=1} n\ (yi-yi\hat{})2}$

An R^2 value close to 1 indicates that the model explains the variability of the data well. A value close to 0 indicates that the model does not explain the variability well.d

Predictive Maintenance

Prevent failures in photovoltaic panels and reduce operational costs by implementing an early warning system that detects and predicts maintenance needs.

Data Collection

Data to Collect:

- Historical and Real-Time Data of Photovoltaic Panels:
- Energy Production: kWh/day.
- Solar Radiation: kWh/m²/day.
- Temperature: °C.
- Humidity: %.
- Wind Speed: m/s.
- Previous Maintenance Data:
• Failure History: Type of failure, date, and time.
• Maintenance Interventions: Type of intervention, date, and time.

Additional Sensor Data:

- Light Tracking Sensors: Data on panel orientation and alignment.
- Panel Status Sensors: Temperature and voltage of each panel.

Development of the Predictive Model

Data Preprocessing:

- Data Cleaning: Impute missing values, remove outliers.
- Normalization/Standardization: Scale variables to improve model performance.

Algorithm Selection:

- Logistic Regression: To predict the probability of failure based on panel characteristics.
- Decision Trees: To identify patterns and conditions that precede failures.
- Neural Networks: To capture complex and nonlinear relationships between variables.

Example

```
from sklearn.model_selection import
train_test_split
from sklearn.ensemble import
RandomForestClassifier
from sklearn.metrics import
classification_report, confusion_matrix

# Preparar los datos
X = data[['radiacion', 'temperatura', 'humedad',
'velocidad_viento', 'estado_panel']]
y = data['fallo']

# Dividir datos
X_train, X_test, y_train, y_test =
train_test_split(X, y, test_size=0.3,
random_state=42)

# Definir y entrenar el modelo
modelo_rf =
RandomForestClassifier(n_estimators=100,
max_depth=10, random_state=42)
modelo_rf.fit(X_train, y_train)

# Hacer predicciones
y_pred = modelo_rf.predict(X_test)

# Evaluar el modelo
print(confusion_matrix(y_test, y_pred))
print(classification_report(y_test, y_pred))
```


Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.
https://doi.org/10.35429/EJRC.2024.10.18.5.13

Predicción de producción de energía con lineal

```
# Importar las librerías
import numpy as np
import pandas as pd
from sklearn.model_selection import
train_test_split
from sklearn.linear_model import
LinearRegression
from sklearn.metrics import
mean_squared_error, r2_score
import matplotlib.pyplot as plt

# Generar datos
np.random.seed(42)
n_samples = 100
X = np.random.rand(n_samples, 4) * 100
# Datos de entrada: irradiación solar,
temperatura, humedad y velocidad del viento
estado_panel = np.random.randint(0, 2,
size=(n_samples, 1))  # Estado del panel: 0
(bueno) o 1 (necesita mantenimiento)
y = 3.5 * X[:, 0] + 1.5 * X[:, 1] - 2 * X[:, 2] +
X[:, 3] + estado_panel[:, 0] * 5 +
np.random.randn(n_samples) * 10
 # Datos de salida: producción de energía en
kWh

# Combinar X y estado_panel en una sola matriz
de entrada
X = np.hstack((X, estado_panel))

# Dividir el conjunto de datos en entrenamiento
y prueba
X_train, X_test, y_train, y_test =
train_test_split(X, y, test_size=0.2,
random_state=42)

# Entrenar el modelo de regresión lineal
model = LinearRegression()
model.fit(X_train, y_train)

# Predecir los valores de producción de energía
para los datos de prueba
y_pred = model.predict(X_test)

# Evaluar el modelo
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f"Mean Squared Error: {mse}")
print(f"R² Score: {r2}")
# Verificar formas
print(f"Shape of X_test[:, 0]: {X_test[:,
0].shape}")
print(f"Shape of y_test: {y_test.shape}")
print(f"Shape of y_pred: {y_pred.shape}")

# Visualizar los resultados
plt.figure(figsize=(10, 6))
plt.scatter(X_test[:, 0], y_test, color='blue',
label='Datos reales')  # Usar irradiación solar
para el eje x
plt.scatter(X_test[:, 0], y_pred, color='red',
label='Predicciones', alpha=0.5)  # Usar
irradiación solar para el eje x
plt.xlabel('Irradiación Solar (kWh/m²)')
plt.ylabel('Producción de Energía (kWh)')
plt.title('Predicción de Producción de Energía
con Regresión Lineal')
plt.legend()
plt.show()

# Datos en tiempo real
radiacion_actual = 80  # kWh/m²
temperatura_actual = 25  # °C
humedad_actual = 60  # %
velocidad_viento_actual = 5  # m/s
estado_panel_actual = 0  # 0 (bueno) o 1
(necesita mantenimiento)

# Crear un array con los datos en tiempo real

datos_en_tiempo_real =
np.array([radiacion_actual, temperatura_actual,
humedad_actual, velocidad_viento_actual,
estado_panel_actual]).reshape(1, -1)

# Predecir la producción de energía con los
datos en tiempo real
prediccion_energia =
model.predict(datos_en_tiempo_real)
print(f"Predicción de Producción de Energía
(kWh): {prediccion_energia[0]}")
```

Box 1

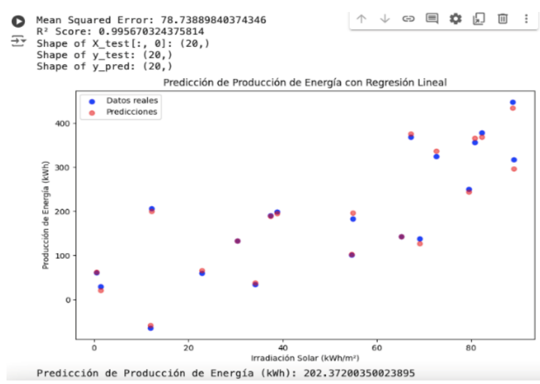

Figure 1

Production Prediction with Linear Regression
Source: Data analysis and visualization performed using *Google Colab (https://colab.research.google.com/)*

Box 2

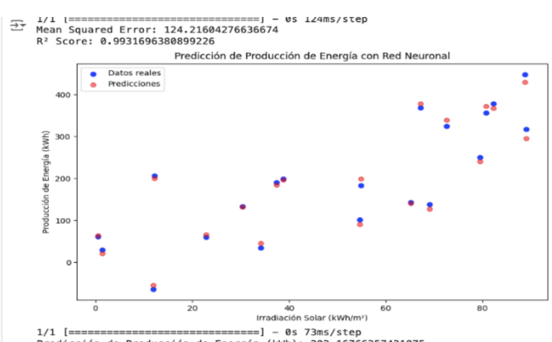

Figure 2

Production Prediction with Neural Network

ISSN: 2414-4959.
RENIECYT-CONAHCYT: 1702902

Ruiz-Garduño, Jhacer Kharen, Flores-Serrato, Leonel, González-Ramírez, Claudia Teresa and Viñas-Álvarez, Samuel Efrén. [2024]. Optimization of Photovoltaic Panels through Machine Learning Algorithms Linked to Predictive Maintenance. ECORFAN-Journal Republic of Cameroon. 10[18]1-13: e51018113.
https://doi.org/10.35429/EJRC.2024.10.18.5.13

Evaluation Metrics

The MSE and R^2 were calculated for both models (Neural Network and Linear Regression). We printed the evaluation metrics to compare the performance of both models.

Neural Network: Tends to be more powerful and capable of capturing complex patterns in the data, but may require more data and training time.

Linear Regression: Is simpler and faster to train but may not adequately capture nonlinear patterns in the data.

Conclusions

This study developed and compared predictive models to optimize photovoltaic panels and perform predictive maintenance using linear regression and neural networks. Both approaches significantly improved the efficiency and sustainability of photovoltaic systems.

Linear regression demonstrated adequate capability for predicting energy production based on environmental variables such as solar radiation, temperature, humidity, and wind speed. It is a simpler and faster model to train but is limited in scenarios with complex nonlinear relationships.

Neural networks, on the other hand, showed greater accuracy in predicting energy production by better capturing complex relationships between variables, although they require more data and training time.

For predictive maintenance, an effective system was developed to predict maintenance needs, detect anomalies, and prevent failures before they occur, reducing operational costs and extending the lifespan of the panels.

Implementation in Zitácuaro, Michoacán, Mexico, showed that it is possible to maximize solar energy production and optimize panel maintenance, contributing to greater sustainability and energy efficiency in the region. The use of machine learning in optimizing photovoltaic systems and predictive maintenance increases efficiency, reduces costs, and promotes sustainable practices, with potential for adaptation to different scales and environments.

Conflict of Interest

The authors declare that they have no conflicts of interest. There are no financial interests or personal relationships that could have influenced the preparation or results of this article.

Author Contributions

Ruiz-Garduño, Jhacer Kharen: Designed the study, performed data analysis, and wrote the final manuscript.

Flores-Serrato, Leonel Alejandro: Contributed to the study design.

Viñas-Álvarez, Samuel Efrén: Contributed to the design and development of the study.

González-Ramírez, Claudia Teresa: Participated in the analysis, design, and development of the study.

Funding

We thank the National Technological Institute of Mexico (TECNM) for funding this project, in the 2024 PI call, enabling its realization and success. We also express our gratitude to the Institute of Science, Technology, and Innovation (ICTI) of Michoacán for their support.

Their contributions were crucial to the development and execution of this research, and their commitment to sustainability and scientific advancement was vital in achieving the proposed objectives. Without their collaboration, this work would not have been possible.

Acknowledgements

We thank the Biocenosis Association Zone Monarca for their invaluable collaboration and support in the development of this project.

Their dedication and technical assistance have been fundamental to the success of our research. Their contribution has been essential in achieving the proposed objectives.

Data Availability

The data supporting the findings of this study are available upon request from the corresponding author. For access, please contact Jhacer Kharen Ruiz Garduño at jhacer.rg@zitacuaro.tecnm.mx

References

Neama, N. , Abbood, R. and Daham, K. (2023) Renewable Energy Is an Essential Variable in Achieving Sustainable Development. Open Journal of Business and Management, 11, 1116-1124. doi: 10.4236/ojbm.2023.113062.

David, M. , Uzorka, A. and Makeri, Y. (2022) Optimisation of a Renewable Energy System for Rural Electrification. Journal of Power and Energy Engineering, 10, 1-15.

Cardil, T., Cortés, A.,Margalef, T., Molina, D., Pelegrín,L. & Ramírez, J. (2015)Forest Fire Propagation Prediction Based on Overlapping DDDAS Forecasts, Procedia Computer Science, Volume 51,2015,Pages 1623-1632, ISSN 1877-0509

Pereira, S., Abreu,E., Iakunin, M., Cavaco, A., Salgado, R. & Canhoto,P. (2022) Method for solar resource assessment using numerical weather prediction and artificial neural network models based on typical meteorological data: Application to the south of Portugal, Solar Energy,Volume 236,2022,Pages 225-238, ISSN 0038-092X,

Osman,M., Domeisen, D., Robertson, A. & Weisheimer, A.(2023) Sub-seasonal to decadal predictions in support of climate services, Climate Services,Volume 30,2023, 100397, ISSN 2405-8807,

Beccali, M., Ciulla,G., Lo Brano, V., Galatioto, A. & Bonomolo, M. (2017) Artificial neural network decision support tool for assessment of the energy performance and the refurbishment actions for the non-residential building stock in Southern Italy, Energy, Volume 137, 2017, Pages 1201-1218, ISSN 0360-5442,

Báez, C. (2024). Diseño de un modelo de inteligencia artificial para predecir la energía solar producida. Trabajo de grado profesional Mechanical Engineering. University of Antioquia, Medellín, Colombia,(2024).

Sandoval, C.(2020). Solar Comet Project - Cs For Photovoltaic System Optimization. Universidad Ciencia Y Tecnología, 24(100), 74-87. Retrieved from

Davila, G. G. (2023). Optimization and Improvement of Energy Efficiency in a Building through the Installation of Photovoltaic Panels.

González, M. (2022). Optimization Models for Cleaning Planning in Solar Photovoltaic Parks Master's thesis, Pontifical Catholic University of Chile.

Cobian, E. (2023). Optimización de limpieza de paneles solares en plantas fotovoltaicas.

Rodríguez, L., Maestre, V., Díez, L., Ortiz, A., Agüero, R., & Ortiz, I. (2023). Modeling a Digital Twin for the Optimization of a Residential Self-Supply Energy System.

Bady, M. (2015) Towards Sustainable Power Generation Using Solar Chimney. Open Access Library Journal, 2, 1-9.

Valverde, L., López, C. and Santiago, J. (2023) Design and development of a comprehensive renewable energy system. ECORFAN Journal-Mexico. 2023. 14-31:39-45

Bashery, M. (2023) A comprehensive framework based on Bayesian optimization and skip connections artificial neural networks to predict buildings energy performance, Journal of Building Engineering, Volume 77, 2023, 107523, ISSN 2352-7102,

Di Giovanni, G., Rotilio, M., Giusti, L., & Ehtsham, M. (2024). Exploiting building information modeling and machine learning for optimizing rooftop photovoltaic systems. Energy and Buildings, 313, 114250.

Buturache, A. and Stancu, S. (2021) Wind Energy Prediction Using Machine Learning. Low Carbon Economy, 12, 1-21. DOI: 10.4236/lce.2021.121001.

Costa, P. , Simões, T. and Estanqueiro, A. (2019) A GIS Methodology for Planning Sustainable Renewable Energy Deployment in Portugal. Energy and Power Engineering, 11, 379-391.

David, M. , Uzorka, A. and Makeri, Y. (2022) Optimisation of a Renewable Energy System for Rural Electrification. Journal of Power and Energy Engineering, 10, 1-15.

Emodi, V. , Yusuf, S. and Boo, K. (2014) The Necessity of the Development of Standards for Renewable Energy Technologies in Nigeria. Smart Grid and Renewable Energy, 5, 259-274.

Ruggeri, A., Calzolari, M., Scarpa, M., Gabrielli, L., & Davoli, P. (2020). Planning Energy Retrofit on Historic Building Stocks: A Score-Driven Decision Support System. Energy and Buildings, Volume 224, 2020, 110066, ISSN 0378-7788,

Abdulla, H., Sleptchenko, A., & Nayfeh, A. (2024). Photovoltaic systems operation and maintenance: A review and future directions. Renewable and Sustainable Energy Reviews, 195, 114342.

Fard, Z., Zomorodian, Z.,Tahsildoost, M.(2023) Development of a Machine Learning Framework Based on Occupant-Related Parameters to Predict Residential Electricity Consumption in the Hot and Humid Climate, Energy and Buildings, Volume 301, 2023, 113678, ISSN 0378-7788,

González, P., Zamarreño, J. (2005). Prediction of Hourly Energy Consumption in Buildings Based on a Feedback Artificial Neural Network. Energy and Buildings, Volume 37, Issue 6, 2005, Pages 595-601, ISSN 0378-7788,

Hernández Flores, M. Á. (2021). Design of an Automatic Optimization Tool for Installing Photovoltaic Panels on Building Roofss.

Mendoza, R., Ruiz, J, Jiménez, A & González, C. (2023) Prototype of a sustainable electricity generation system in an indigenous Mazahua community. Journal of Technological Prototypes. 2023. 9-23:7.

Rodríguez, L., Maestre, V., Díez, L., Ortiz, A., Agüero, R., & Ortiz, I. (2023). Modeling a Digital Twin for the Optimization of a Residential Self-Supply Energy System.

Silva, R., Penso, R., Dalapicula, G., Magalhães, C., Toledo, E., Cruz, C., Valiati, C., Nascimento, M., & Cecon, P. (2022). Artificial Neural Network as an Alternative for Peach Fruit Mass Prediction by Non-Destructive Method. Scientia Horticulturae, Volume 299, 2022, 111014, ISSN 0304-4238,

Biswas, C., Chakraborti, A., & Majumder, S. (2024). Recent advancements in artificial intelligence and machine learning in sustainable energy management. In Sustainable Energy Solutions with Artificial Intelligence, Blockchain Technology, and Internet of Things (pp. 35- 46). CRC Press.