SEP 769 – CYBER PHYSICAL SYSTEMS
DEEP LEARNING – PROJECT

Purva Singh | Chris Xavier Mathias | Vipin Chandran Muthirikkaparambil | Suraj Ramesh

Project Title:

Anomaly Detection and Forecasting of
Solar Power Plant Performance across Deep Learning Models

Project Abstract:

This project introduces an integrated deep learning framework for short-term solar power forecasting and inverter-level anomaly detection. Utilizing high-resolution sensor and generation data from two solar plants in India, the study evaluates four forecasting models, Feedforward Neural Network (FFNN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and a hybrid CNN-LSTM on their ability to predict AC power output. Among these, the LSTM model, optimized using the Nadam optimizer, demonstrated the best overall performance. For anomaly detection, a deep autoencoder is deployed to identify deviations in inverter behavior, with Dynamic Time Warping (DTW) applied to localize anomalous intervals. The unified system supports intelligent monitoring, fault localization, and predictive maintenance in photovoltaic (PV) systems.

Project Introduction:

The increasing demand for renewable energy solutions has amplified the need for precise forecasting and real-time monitoring in solar power systems. Solar photovoltaic (PV) performance is inherently variable, affected by irradiation, temperature, and panel health. Traditional statistical methods often fall short in capturing the non-linear and temporal nature of solar generation. This project addresses those gaps by employing deep learning architectures that can better model complex dependencies over time. In addition to forecasting, inverter faults and inefficiencies can result in generation losses if not detected early. Hence, a parallel unsupervised anomaly detection system is included to enhance operational resilience. By combining forecasting with localized fault detection, the proposed framework enables efficient energy management, increased reliability, and faster response times in solar PV systems.

Methodology Used:

Forecasting Approach

The forecasting module involves preprocessing plant-level data at 15-minute intervals, handling missing values using interpolation and forward-filling techniques. Key weather features, irradiation, module temperature, and ambient temperature are engineered alongside temporal encodings (hour_sin, hour_cos). Four models are built and evaluated:

- FFNN: Uses flattened 96-time step windows. Despite capturing overall trends, it performs poorly at predicting sharp peaks.
- CNN: Applies 1D convolutions over time-steps to capture local dependencies. Performs well on smooth patterns but underestimates high peaks.
- LSTM: Captures long-term dependencies using stacked LSTM layers. When paired with the Nadam optimizer, it achieves the lowest RMSE and MAE.
- CNN-LSTM Hybrid: Combines CNN's short-term pattern recognition with LSTM's sequence learning. Provides smooth and stable forecasts but slightly underperforms on extreme values.

All models use early stopping, Min-Max normalization, and are evaluated using RMSE and MAE.

Anomaly Detection Approach

For anomaly detection, inverter-level AC power is reshaped into daily profiles (96 time slots). After data cleaning and imputation, profiles are normalized and fed into a deep autoencoder. The encoder compresses each 96-dimensional profile, and the decoder reconstructs it. Anomalies are flagged if reconstruction error (MSE) exceeds a threshold defined as $\mu + 3\sigma$. To pinpoint when the anomaly occurs during the day, Dynamic Time Warping (DTW) is used. Each inverter-day is divided into 24 one-hour segments, and segment-wise DTW distances between original and reconstructed profiles are computed. Segments exceeding the 95th percentile of the DTW error distribution are identified as locally anomalous, supporting granular diagnostics.

Results:

Forecasting Results

- FFNN: RMSE = 9154.23 kW, MAE = 1370.95 kW. Accurately models daily trend but fails during peaks.
- CNN: RMSE = 2905.11 kW, MAE = 1446.18 kW. Stable predictions but underperforms during output spikes.
- LSTM: RMSE = 2206.16 kW, MAE = 1182.46 kW. Best-performing model, effectively capturing temporal dependencies.
- CNN-LSTM: RMSE = 2905.11 kW, MAE = 1446.18 kW. Balanced performance but smooths out spikes.

LSTM emerges as the most suitable model due to its superior handling of time-based variability and spiky power outputs.

Anomaly Detection Results

The autoencoder model successfully flagged 18 out of 1012 inverter-day profiles as anomalous. DTW further localized these anomalies to specific one-hour segments, with common deviations occurring during midday. Visual overlays of original and reconstructed profiles confirm abnormal behavior such as power drops and sudden spikes. This combination of autoencoder and DTW enables accurate and explainable diagnostics at both profile and temporal segment levels.