Predictive Maintenance for Solar Panels - AWS Ascend Zalestra ML Challenge

As solar energy systems become increasingly popular in sustainable energy infrastructures, maintaining high performance and reducing downtime are essential. Traditional maintenance methods for photovoltaic (PV) systems are often reactive, leading to energy loss and increased costs. The goal of this project is to develop a Machine Learning model that predicts performance degradation and potential failures in solar panels.

Task

You are tasked to develop a Machine Learning model that predicts performance degradation and potential failures in solar panels using historical and real-time sensor data, enabling predictive maintenance and optimal energy output.

Dataset Description

The dataset folder contains:

- train.csv: 20,000 rows × 17 columns

- test.csv: 12,000 rows × 16 columns

- sample_submission.csv: 5 rows × 2 columns

Columns provided:

id: Unique row identifier

temperature: Ambient air temperature (°C), affects panel output efficiency

irradiance: Solar energy received per unit area (W/m²)

humidity: Air moisture content, may impact panel performance over time

panel_age: Age of the solar panel in years

maintenance_count: Number of previous maintenance activities

soiling_ratio: Efficiency reduction due to dust/debris on panel (0.0-1.0)

voltage: Voltage output from the panel (V)

current: Current output from the panel (A)

module_temperature: Temperature of the panel surface

cloud_coverage: Sky coverage by clouds (percentage)

wind_speed: Wind speed (m/s), aids cooling and affects temperature

pressure: Atmospheric pressure at the panel location (hPa)

string_id: Identifier for panel group (e.g., A1, B2)

error_code: Error codes from panel diagnostics (e.g., E00, E01)

installation_type: Mounting type: fixed, tracking, or dual-axis

efficiency: Target variable representing panel energy output efficiency

Evaluation Metric

The model's performance will be evaluated using the following metric:

Score = 100 * (1 - sqrt(MSE(actual, predicted)))

Submission Guidelines

- The index is "id" and the target is "efficiency".

- Submission must be a .csv file with shape 12,000 × 2.

- Ensure correct index values and column names as per sample_submission.csv.