Forecasting Solar Panel Power Outputs

Forecasting the Power Outputs of Photovoltaic Cells: a comparison of statistical, machine learning and deep learning models

This project demonstrates my skills using Python to build statistical models, machine learning models and deep learning models for multivariate time series forecasting and compare them using performance metrics and data visualisations.

Intro and Background

With the effects of climate change increasing (The Causes of Climate Change, 2023); diminishing fossil fuel resources (Shafiee and Topal, 2009); and rising energy costs (Guan et al., 2023), global energy production will become more reliant on renewable energy sources such as PV installations (Hu et al., 2016). Worldwide, countries are turning to renewable energy sources to combat these issues.

Forecasting the power outputs of PV systems is useful for several reasons:

* Aids in meeting government renewable energy goals.
* Helps solar panel companies predict lifespans and payback periods leading to economic benefits for the companies and prosumers.
* Allows grid companies to prepare for potential high periods of generations from these systems.

The technical process followed during this project is outlined in Figure 1.

Process

**Analyse how the residential PV systems power output changes with time and analyse weather features effects on power output.**

The distributions for each variable were examined by the density plots and box plots shown in Figure 2. The Shapiro-Wilk test for normality was used alongside these visualisations to test if the target variable of *Monthly Power Generation (kWh)* had a normal distribution. The result from this determined the target was not normally distributed.

Therefore, Spearman’s correlation was used to test the correlations between the variables and the target, Figure 3. These relationships are plotted in Figure 4 and the Spearman’s correlations and the corresponding p-values are displayed in Table 1.

**Incorporate target lags and weather data to build three long-term, multivariate forecasting models: statistical, machine learning and deep learning.**

**Use predictions for twelve months to compare results using plots and performance metrics.**