

Relatório # 2

Engenharia de recursos para modelagem de previsão de geração de energia em usinas fotovoltaicas

Resumo

Este relatório descreve o processo de seleção e criação dos recursos mais relevantes para alimentar um modelo de previsão de geração de energia em usinas fotovoltaicas, utilizando uma abordagem dirigida por dados.

Engenharia de Recursos

Os dados selecionados para a modelagem dirigida por dados são provenientes de duas fontes principais: medições coletadas da estação solarimétrica e medições coletadas dos inversores de frequência.

A escolha das variáveis foi fundamentada em análises de correlação, conhecimentos de engenharia elétrica e experimentação para identificar os recursos mais relevantes. Mais especificamente, as seguintes variáveis foram selecionadas para modelar a geração de energia em usinas fotovoltaicas:

variáveis da estação solarimétrica:

- `poa_irradiance`: irradiância [W/m^2] no plano dos painéis.
- `module_temperature`: temperatura [$^{\circ}\text{C}$] do módulo solarimétrico.
- `air_temperature`: temperatura [$^{\circ}\text{C}$] ambiente.
- `timestamp`: data [ano-mês-dia] e hora [hora:minuto:segundo] de cada observação/medição.

variáveis dos inversores de frequência:

- `dc_power`: potência DC [W].
- `ac_active_power`: potência AC (ativa) [kW].

- `daily_yield`: geração diária [kWh]. A geração diária é a soma cumulativa da energia gerada naquele dia, até aquele instante de tempo.
- `total_yield`: geração total [kWh] (soma cumulativa da energia gerada) até aquele instante de tempo.
- `timestamp`: data [ano-mês-dia] e hora [hora:minuto:segundo] de cada observação/medição.

Para aprimorar a qualidade dos resultados de um modelo dirigido por dados, podem ser criadas diversas outras variáveis artificiais (recursos extras) a partir de transformações matemáticas das variáveis, como:

- **Transformações trigonométricas:** Cosseno e seno das variáveis para capturar padrões cíclicos.
- **Variáveis de atraso:** Atrasar as variáveis em diferentes intervalos de tempo para capturar efeitos de dependência temporal.
- **Médias móveis:** Calcular a média móvel das variáveis para suavizar ruídos e tendências de curto prazo.
- **Decomposição de séries temporais:** Extrair componentes de tendência, sazonalidade e residuais das variáveis para capturar padrões subjacentes.

Com o objetivo de simplificar o modelo, melhorar sua capacidade de generalização e reduzir a complexidade computacional, as variáveis individuais dos inversores foram agregadas globalmente pela soma, resultando nas seguintes variáveis:

variáveis globalmente agregadas (pela soma):

- `global_dc_power`: soma das potências DC de todos os inversores.
- `global_ac_active_power`: soma das potências ativas de todos os inversores.
- `global_total_yield`: soma das gerações totais de todos os inversores.
- `global_daily_yield`: soma das gerações diárias de todos os inversores.

A variável dependente (variável alvo), a ser prevista, será a variável *Global Daily Yield* [kWh], a partir da qual podem ser calculados outros indicadores.

Os preditores da variável alvo, isto é, as variáveis independentes (entradas ou recursos) são as demais variáveis (*POA Irradiance* [W/m²], *Module Temperature* [°C], *Air Temperature* [°C], *Global DC Power* [W], *Global AC Active Power* [kW] e *Global Total Yield* [kWh]), as quais, como mencionado, são dadas em função do tempo.

Conclusão

O relatório detalha a abordagem utilizada na engenharia de recursos para modelagem de geração de energia em usinas fotovoltaicas.

A combinação dos recursos selecionados (via análises de correlação, conhecimentos do domínio e experimentação empírica), transformações matemáticas (incluindo agregações globais) e a definição precisa das variáveis preditoras e da variável alvo contribuem para a construção de modelos robustos, eficazes e precisos.