

Análise Preditiva da Geração Fotovoltaica no Mercado de Energia do Complexo Solar Pirapora

Leonardo Fini¹, Samantha Lanzelotti², Marilaine Colnago³, Wallace Casaca⁴
FEC/UNESP, Rosana, SP

Atualmente, a implementação de fontes renováveis de energia tem sido uma pauta recorrente em conferências mundiais do clima e de sustentabilidade econômica. De fato, é notório os esforços envolta aos incentivos fiscais e fomento para a inserção das fontes renováveis em detrimento à utilização de combustíveis fósseis, uma vez que estes trazem resultados negativos seja em relação ao ambiente, sociedade e/ou economia [3]. Por exemplo, a partir do incentivo à projetos de minigeração e microgeração distribuída, é possível usuários obter compensação da energia elétrica fotovoltaica gerada em residências ou comércios com a empresa distribuidora, promovendo assim a diversificação energética das matrizes elétricas em diversos países do globo [2].

Sendo assim, prever a quantidade de energia que será gerada em um dado horizonte tem se tornado uma problemática relevante, sobretudo, na comercialização de venda e compra de energia elétrica no mercado livre de energia. Dessa forma, ferramentas de *Machine Learning* (ML) ajustadas para a tarefa de previsão da geração da energia elétrica tem sido cada vez mais empregadas no setor, já que viabilizam a construção de modelos preditivos mais assertivos e eficientes para tomadas de decisões e estratégias de planejamento energético.

Portanto, este trabalho propõe a implementação, ajuste e aprimoramento de métodos de ML para o problema de previsão (regressão) da geração elétrica do Complexo Solar de Pirapora (MG), a segunda maior usina fotovoltaica do Brasil. Para cumprir com esse propósito, foi utilizado a plataforma de programação *Anaconda® Navigator*, que viabilizou o uso da linguagem *Python* a fim de utilizar e testar algoritmos de ML. Em relação ao repositório de dados, foram adotados os dados climáticos diários coletados do Instituto Nacional de Meteorologia (INMET), bem como a série histórica da geração solar fotovoltaica (média diária) fornecida pelo Operador Nacional do Sistema Elétrico (ONS), ambos referente ao período de 01/01/2018 a 31/12/2021.

Antes da implementação e aplicação dos modelos preditivos, foi realizado a etapa de tratamento dos dados através da “limpeza” dos mesmos, de modo a identificar e tratar valores nulos ou inconsistentes (*data cleaning*), bem como a etapa de análise exploratória dos dados visando compreender o comportamento das variáveis em estudo, além de relacionar e exibir padrões que não seriam possíveis sem o auxílio de ferramentas gráficas e estatísticas.

Em relação aos modelos preditivos adotados, utilizou-se o *Random Forest*, *Gradient Boosting* e o *Support Vector Regressor* [1]. Também foi aplicado a ferramenta *Random Search* para a seleção dos melhores hiperparâmetros destes modelos, além da engenharia de recursos, que consiste na criação de novas variáveis a partir das já existentes, ambas técnicas possuem como intuito melhorar a assertividades desses algoritmos [1]. Finalmente, dividiu-se a base de dados de forma que os três primeiros anos foram utilizados para o treinamento dos modelos preditivos, e o último ano para a validação dos modelos pós treinados.

¹l.fini@unesp.br

²samantha.lanzelotti@unesp.br

³marilaine.colnago@unesp.br

⁴wallace.casaca@unesp.br

Para fins de validação quantitativa, foi computado o Erro Absoluto Médio Percentual (MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{|Y_i|} \times 100, \quad (1)$$

em que n representa o tamanho da amostra, Y_i os valores reais, e \hat{Y}_i os valores preditos. Após a aplicação dos modelos nos dados de teste, obteve-se valores de MAPE de 8,00%, 8,05% e 8,00%, para os modelos preditivos *Random Forest* (RF), *Gradient Boosting* (GB) e *Support Vector Regressor* (SVR), respectivamente. As previsões dos últimos 15 dias de 2021, dos três algoritmos avaliados, podem ser visualizadas e comparadas com os valores reais através da Figura 1.

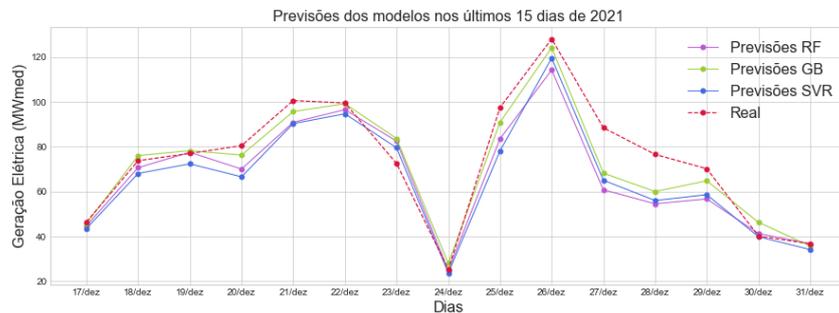


Figura 1: Previsões da geração elétrica (MW médio) dos modelos preditivos nos últimos 15 dias de 2021.

Por fim, a partir dos resultados alcançados com a referida pesquisa, foi possível confeccionar modelos com elevada acurácia, os quais podem servir de apoio para gestores e agentes do mercado de energia elétrica brasileiro. Além disso, pode-se utilizar a mesma metodologia empregada neste trabalho para a previsão do PLD (Preço de Liquidação das Diferenças), que é o preço atrelado a energia comercializada no mercado de curto prazo (mercado *spot*), por exemplo.

Agradecimentos

Os autores agradecem à FAPESP (#2021/08298-5) e ao CNPq (#316228/2021-4) pelo incentivo e fomento à pesquisa.

Referências

- [1] J. V. Leme et al. “Towards Assessing the Electricity Demand in Brazil: Data-Driven Analysis and Ensemble Learning Models”. Em: **Energies** 13.6 (2020), p. 1407. DOI: 10.3390/en13061407.
- [2] L. Ramos, M. Colnago e W. Casaca. “Data-driven Analysis and Machine Learning for Energy Prediction in Distributed Photovoltaic Generation Plants: A Case Study in Queensland, Australia”. Em: **Energy Reports** 8 (2022), pp. 745–751. DOI: 10.1016/j.egy.2021.11.123.
- [3] J. Yu et al. “Role of Solar-based Renewable Energy in Mitigating CO2 Emissions: Evidence from Quantile-on-Quantile Estimation”. Em: **Renewable Energy** 182 (2022), pp. 216–226. DOI: 10.1016/j.renene.2021.10.002.