



## PREVISÃO DE VELOCIDADE DO VENTO POR APRENDIZADO DE MÁQUINAS

MOTA, R.A.F.<sup>1.</sup>; SANTOS L.I.<sup>2</sup>

<sup>1</sup>Discente do curso de pós graduação em Modelagem Computacional da UNIMONTES – *Campus Montes Claros*; <sup>2</sup>Docente do curso superior em Ciência da Computação IFNMG – *Campus Montes Claros*.

### Introdução

Em razão do acelerado crescimento do uso de energias renováveis limpas para a preservação do ambiente e diversificação da matriz energética, a inserção dessas energias nas redes elétricas também precisa acompanhar essa evolução. O uso seguro e eficiente dos sistemas de transmissão depende continuamente de previsões precisas das gerações de energia, assim como para a otimização da inserção desse recurso. Diante disso, são necessárias previsões meteorológicas direcionadas à geração de energia para o melhor proveito dos recursos renováveis naturais, como radiação solar e energia eólica. Logo, este trabalho implementa métodos variados de aprendizado de máquina para a previsão de variáveis meteorológicas, com foco na velocidade do vento, de principal importância para a energia eólica, processo o qual tem o potencial de permitir que os sistemas de geração renováveis sejam capazes de obter maior eficiência nos níveis de energia produzidos em sistemas de alta complexidade.

### Material e Métodos

Este trabalho realiza, primeiramente, o tratamento de dados meteorológicos para a previsão de recursos naturais a implementação de diferentes modelos de aprendizado de máquinas e verificação dos resultados obtidos, sendo esse processo realizado pelas linguagens Python e R. Para isso, o conjunto de dados meteorológicos para a cidade de Macau, no Rio Grande do Norte, para o ano de 2019 — obtido em INMET (2022) — passa por uma filtragem de dados faltantes e pontos fora da curva, de modo que sejam devidamente corrigidos, usando a biblioteca Pandas. Essa cidade possui alto potencial eólico e boa qualidade de dados no site do INMET, sendo então a base do trabalho.

Os modelos de aprendizado de máquina são implementados por meio das bibliotecas Scikit-Learn, ELM e XGBoost, enquanto que os modelos autorregressivos tracionais são implementados em R, por ser um ambiente capaz de proporcionar a divisão entre treino e teste para essa classe de modelos de maneira mais clarificada. Dessa forma, foram este trabalho demonstra a aplicação dos seguintes métodos: Persistência — o qual considera que o próximo valor a ser previsto será igual ao último registrado, o qual é utilizado como *baseline* para os resultados dos demais modelos —, SARIMAX, Regressão linear (RL), *Multilayer Perceptron* (MLP), *Extreme Learning Machine* (ELM), *Support Vector Regression* (SVR), *Random Forest* (RF), *Gradient Boosting Machine* (GBM), *Extreme Gradient Boosting* (XGB) e *Light Gradient Boosting Machine* (LGBM). Com exceção dos dois primeiros, que não possuem hiperparâmetros variados, os modelos são testados tanto sem a otimização de hiperparâmetros, ou seja, utilizando a configuração padrão das bibliotecas. Já na busca pelos hiperparâmetros otimizados, é utilizada a busca por árvores de decisão, da biblioteca Sci-Opt.

Por fim, são realizadas as avaliações das performances de cada modelo por meio do Erro Médio Absoluto (MAE), Erro Médio Absoluto Percentual (MAPE), Erro Quadrático Médio (MSE), Raiz do Erro Quadrático Médio (RMSE) e o  $R^2$ , além do registro do tempo total envolvido no processo de implementação de cada método, com a ressalva de que os modelos em Python foram implementados no Google Colab (em máquinas virtuais de mesmos atributos) e o SARIMAX foi implementado em



R como explicado anteriormente, também no Google Colab.

## Resultados e Discussão

Com a metodologia definida, foram utilizados os dados de Macau em 2019 para, enfim, se obter a previsão para cada modelo de regressão aplicado no contexto de séries temporais de velocidade do vento. Dessa forma, são mostrados os resultados conforme a Tabela 1 e Figura 1. Dessa forma, as principais análises obtidas, levando em consideração o modelo da persistência como base e o MAE como métrica principal de interesse para a geração eólica — por estar na mesma unidade de medida da velocidade do vento e ter a mesma proporção que o vento a passar nas turbinas em si —, indicam uma boa precisão do SARIMAX, em troca de um alto custo computacional, a melhoria de precisão do ELM ao ser utilizado com uma otimização de hiperparâmetros e a pouca mudança dos outros modelos de aprendizado de máquinas quando comparados entre modelos com as configurações padrão das suas respectivas bibliotecas e quando utilizados com a otimização de hiperparâmetros.

Em termos de balanceamento entre custo computacional e precisão, pode-se destacar as seguintes observações: o tempo extremamente baixo da RL com uma precisão próxima das melhores obtidas, as performances rápidas com boa precisão dos modelos MLP e XGB sem otimização de suas configurações, e a baixa relação entre uso de otimização para busca de hiperparâmetros nos modelos de aprendizado de máquina e a melhoria expressiva dos resultados obtidos nos modelos, com exceção do ELM. Já para a métrica MAPE, foram obtidos resultados melhores para MLP e modelos baseados em Árvores de Decisão do que no SARIMAX, porém, por conta da velocidade do vento se aproximar de zero em alguns casos, o MAE continua sendo mais concreto para as comparações.

## Considerações finais

Em conclusão, pode-se inferir que os modelos de aprendizado de máquina com as configurações padrão são capazes de obter melhores indicadores de custo-benefício para essa tarefa específica empregada neste trabalho. Como a cidade de Macau possui bom potencial eólico, os ventos possuem maior constância em seu comportamento, mostrando que para a previsão de séries temporais de baixa variabilidade, há um grande ganho no custo computacional ao se utilizar os modelos como já são padronizados em cada biblioteca. Além disso, apesar da alta precisão do SARIMAX, ao se considerar que foram obtidas precisões muito próximas com modelos mais rápidos, verifica-se a possibilidade de melhorias futuras na execução desses modelos, de forma a alcançar ou superar a precisão do SARIMAX com velocidades de execução maiores.

## Agradecimentos

Agradecimentos ao IFNMG pela atribuição da bolsa de iniciação científica, na qual foram implementadas as etapas desse trabalho e outras análises relativas a energias renováveis.

## Referências

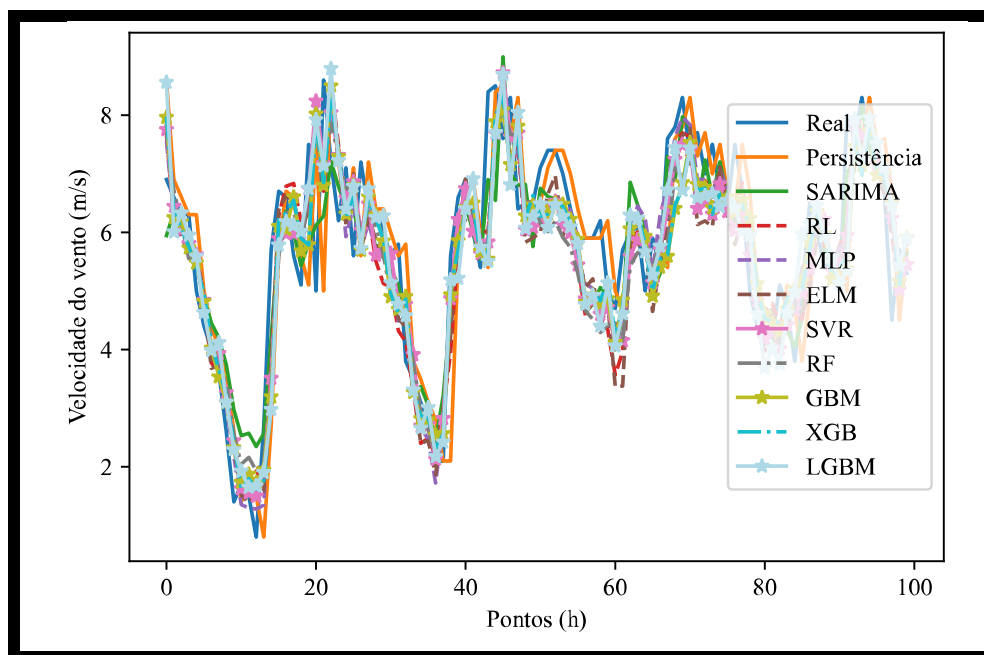
- ABEEOLICA. Associação Brasileira de Energia Eólica. 2022. Disponível em: <<https://abeeolica.org.br/>>. Acesso em: 08 ago. 2023.
- ANEEL. Agência Nacional de Energia Elétrica. 2022. Disponível em: <<https://www.gov.br/aneel/pt-br>>. Acesso em: 08 ago. 2023.
- BURKOV, A. **The hundred-page machine learning book**. [S.l.]: Andriy Burkov Quebec City, QC, Canada, 2019. v. 1.



INMET. Instituto Nacional de Meteorologia. 2022. Disponível em: <<https://portal.inmet.gov.br/dadoshistoricos>>. Acesso em: 08 ago. 2023.

ONS. Operador Nacional do Sistema Elétrico. 2022. Disponível em: <<http://www.ons.org.br/>>. Acesso em: 08 ago. 2023.

WWEA. World Wind Energy Association. 2021. Disponível em: <<https://wwindea.org/>>. Acesso em: 08 ago. 2023.



**Figura 1.** Previsões para a velocidade do vento em Macau, com modelos sem otimização de hiperparâmetros. Elaboração própria (2022).

**Tabela 1.** Resultados gerais, onde os métodos com índice 0 são possuem busca de hiperparâmetros.

Modelo	MAE (m/s)	MAPE (%)	MSE (m <sup>2</sup> /s <sup>2</sup> )	RMSE (m/s)	R2	Tempo (s)
Persistência	0,81	16,71	1,20	1,08	0,61	0
SARIMAX	0,64	14,54	0,71	0,81	0,79	121,17
RL	0,69	14,84	0,79	0,89	0,73	0,03
MLP <sub>0</sub>	0,65	13,78	0,75	0,87	0,75	4,51
ELM <sub>0</sub>	0,73	14,94	0,90	0,95	0,70	0,39
SVR <sub>0</sub>	0,68	13,84	0,80	0,90	0,73	0,67
RF <sub>0</sub>	0,67	13,77	0,80	0,89	0,73	4,07
GBM <sub>0</sub>	0,68	13,97	0,83	0,91	0,72	1,42
XGB <sub>0</sub>	0,67	13,78	0,82	0,91	0,72	0,37
LGBM <sub>0</sub>	0,69	14,08	0,83	0,91	0,72	1,49
MLP	0,65	13,78	0,75	0,87	0,75	84,90
ELM	0,69	14,76	0,82	0,91	0,72	24,36
SVR	0,68	13,84	0,80	0,90	0,73	65,26
RF	0,67	13,77	0,80	0,89	0,73	215,25
GBM	0,68	13,97	0,83	0,91	0,72	54,10
XGB	0,67	13,78	0,82	0,91	0,72	27,17
LGBM	0,69	14,08	0,83	0,91	0,72	64,60

Fonte: Elaboração própria (2022).