

Previsão de Geração de Energia em uma Usina Eólica a partir de Séries Temporais

Manoel Victor Florencio de Souza
Escola Superior de Tecnologia
Universidade do Estado do Amazonas
Manaus, AM, Brasil
mvfds.snf@uea.edu.br

Luis Cuevas Rodriguez
Laboratório LUDUS
Universidade do Estado do Amazonas
Manaus, AM, Brasil
lrodriguez@uea.edu.br

Carlos Trinchet Varella
Centro de Estudios CAD/CAM
Universidad de Holguín
Holguín, Holguín, Cuba
carlos.trinchet@uho.edu.cu

ABSTRACT

Forecasting the generation of renewable energy is necessary mainly for the operators of these systems, in the field of wind energy it is no different, especially considering the inconstancy and unpredictability of the wind behavior. Therefore, this paper proposes and evaluates different models based on machine learning aiming at this problem, with emphasis on a wind farm located on the coast of Cuba. The results of the work were used to develop an operational monitoring system for the plant capable of emitting alarms when the actual energy generation is different from the generation made by the model.

Keywords

Energia Eólica, Aprendizado de Máquina, Previsão de geração de energia

1. INTRODUÇÃO

A economia mundial na atualidade e a própria humanidade está em busca de fontes de energia alternativas e renováveis que possam substituir os combustíveis fósseis como principal fonte de energia.

Entre as energias renováveis que nos últimos anos tem um maior impulso pela comunidade científica está a energia eólica, na qual se pode transformar a energia do vento em energia útil. Este processo ocorre em usinas eólicas compostas geralmente por várias turbinas eólicas, capazes de gerar energia.

Algumas das dificuldades fundamentais enfrentadas pelos operadores destes sistemas de energia são a imprevisibilidade e a variabilidade da geração de energia eólica. A geração vai depender, entre outros elementos: da velocidade do vento e sua variação, o ângulo das pás da turbina em torno de seu eixo, do controle de passo, das revoluções por minutos da turbina, a densidade do ar e a área varrida pelo rotor; por outro lado ela é influenciada pela presença de condições ambientais complexas, ventos de amplitude e mudança de direção, a existência da camada atmosférica e as deformações estáticas. Tendo estes elementos em conta, é evidente a urgência do desenvolvimento ou a seleção de boas ferramentas de previsão [1].

Este artigo propõe o uso de séries temporais para previsão da geração da energia, ou seja, é considerado que a geração de energia num momento t_n depende do conjunto $T = \{t_0, t_1, \dots, t_n\}$, de registros anteriores. Este trabalho está organizado da seguinte

maneira: na Seção 2 está descrita a Solução Proposta e as ferramentas utilizadas para resolver o problema, por fim na Seção 3 são discutidas as Considerações Finais.

2. SOLUÇÃO PROPOSTA

Nesta seção serão apresentados os modelos inteligentes avaliados, a base de dados utilizada, o experimento realizado, o método de pré-processamento, o treinamento dos modelos inteligentes e os resultados dos testes.

2.1 Modelos

Entre os algoritmos de Aprendizado de Máquina (AM) tem aqueles que dado um conjunto de exemplos rotulados, constrói um estimador. Se esse domínio for um conjunto de valores discretos, tem-se um problema de classificação, quando se tem um aprendizado supervisionado, ou de agrupamento, quando se tem um problema de aprendizado não-supervisionado, se o domínio for contínuo e ordenado de valores, tem-se um problema de regressão. Diferentes tipos de algoritmos de AM podem encontrar diferentes fronteiras de decisão [2].

2.2.1 Árvore de Regressão. A árvore de decisão, ou árvore de regressão, utiliza de uma estratégia de divisão e conquista para separar um problema complexo em problemas mais simples []. A Árvore Regressora utilizada neste trabalho usa o Erro Médio Quadrado para avaliar a qualidade de uma separação [2].

2.2.2 Regressão Linear. É um método estatístico para estimar a condicional (valor esperado) de uma variável alvo y , dados os valores de variáveis predictoras X [3]. Este método é muito empregado em problemas de aprendizado de máquina por se tratar de uma forma simples de criação de funções regressoras.

2.2.3 Lasso. Em estatística e aprendizado de máquina, Lasso (*Least Absolute Shrinkage and Selection Operator*) é um método de regressão que realiza tanto a seleção de variáveis quanto a regularização com o objetivo de diminuir, ou aumentar, a função objetivo.

A fórmula do cálculo do Lasso é bem similar a da Regressão Linear com a adição de alguns parâmetros adicionais, o algoritmo se dá pela minimização da função objetivo. O parâmetro que difere da Regressão Linear para o Lasso é o λ que trata do fator de ajuste. Nos experimentos deste trabalho o valor utilizado para o fator de ajuste foi de $\lambda = 1$.

XI Computer on the Beach

2 a 4 de Setembro de 2020, Baln. Camboriú, SC, Brasil

2.2.4 Regressão por Vetores de Suporte. As Máquinas de Vetores de Suporte são embasadas pela teoria de aprendizado estatístico, essa teoria estabelece uma série de princípios que devem ser seguidos na obtenção de classificadores com boa capacidade de generalização. O algoritmo ϵ -SVR (*Support Vector Regression ou Regressão por Vetores de Suporte*) é um tipo de Máquina Vetores de Suporte especializado para problemas de regressão e tem como objetivo encontrar um função que produza saídas contínuas para os dados de treinamento que desviem no máximo de ϵ de seu rótulo desejado [2].

2.2.5 Redes Neurais Artificiais. Este tipo de modelo inteligente se encaixa na classificação de método baseado em otimização, pois o treinamento deste modelo consiste em minimizar (ou maximizar) uma função objetivo. No caso de problemas supervisionados, o valor dos elementos que se quer descobrir é considerado na formulação realizada [2].

2.2.6 Rede Long Short-term Memory. As Redes Long Short-term Memory (LSTM) são um tipo especial de Redes Neurais Artificiais que foram criadas com o objetivo de combater o problema do desaparecimento do gradiente. O princípio básico por trás desta arquitetura é transmitir e armazenar informações importantes para decisões futuras [4].

2.2 Base de Dados

A base de dados em questão foi cedida pela usina de Gibara, localizada no litoral norte de Holguín em Cuba. Nela contém as medições de geração de energia por cada turbina eólica no período referente à 01 de Fevereiro de 2018 até 31 de Março de 2018, estas medições são referentes a geração acumulada de cada turbina eólica de 10 em 10 minutos, caracterizando a tarefa em uma previsão de curto prazo. A usina está formada por 12 turbinas eólicas, delas foi selecionada uma, a número 7, para seu estudo. Por problemas decorridos durante o registro destes dados somente foi possível aproveitar 2416 registros, destes foram usados 2160, os valores destes registros estão entre 0 e 733 KW, seguindo a distribuição normal com média 307.49W e desvio padrão de 191.82W.

2.3 Experimento

O experimento realizado neste trabalho se iniciou com duas fases de pré-processamento dos dados de geração de energia. Primeiramente, o tratamento dos dados para possibilitar o treinamento do modelo, transformando o vetor em uma matriz e então feita a padronização dos valores da matriz. Depois, foi realizado o particionamento dos dados para treino e teste, seguido pelo treinamento e avaliação dos modelos.

A arquitetura da RNA utilizada neste artigo teve seus parâmetros escolhidos a partir de uma técnica chamada Busca em Grade e os que apresentaram os melhores resultados foi a seguinte configuração:

- Função de ativação: Identidade
- Número de camadas ocultas: 2
- Número de neurônios nas camadas ocultas: (11, 1)
- Otimizar: Adam

SOUZA, Manoel V. F. et al.

- Número de épocas: 600

A configuração da Rede LSTM usada neste trabalho foi baseada na arquitetura da Rede Neural Artificial apresentada anteriormente que apresentou melhores resultados, no entanto ainda possui algumas diferenças, sua configuração completa está descrita a seguir:

- Número de camadas LSTM: 1
- Número de unidades na camada LSTM: 50
- Número de épocas: 100
- Tamanho do batch: 32
- Função de ativação nas camadas ocultas: ReLU
- Função de ativação na camada de saída: Softmax

2.4 Pré-processamento

O pré-processamento da série foi feito em 2 momentos, descritos a seguir: (1) foi aplicado uma técnica de janela deslizante, onde foram gerados 2016 vetores contendo 144 registros, este número de registros foi decidido a partir da quantidade máxima de registros de 10 minutos contidos em 1 dia, para prever o registro seguinte; e (2) foi aplicado uma padronização dos valores dos registros seguindo a Equação 3, onde Z_i representa um registro da série padronizado, X_i um registro da série, μ a média dos valores dos registros da série e d o desvio padrão do conjunto de valores.

2.5 Treinamento

A separação dos elementos da matriz resultante do pré-processamento visto na seção anterior a fim de realizar o treinamento e avaliação dos modelos foi feita a partir da função *Time Series Split* da API (*Application Programming Interface*) da biblioteca *Scikit-learn*. Esta função realiza a partição de uma série temporal em $n + 1$ partes, sendo n o número de partições informado, com o objetivo de encontrar um número ótimo de elementos a serem usados para treinamento do modelo inteligente. Tal partição é realizada de acordo com a Figura 2 [10].

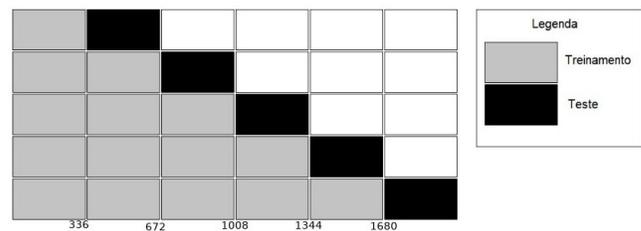


Figura 1: Particionamento de uma Série considerando $n = 5$

Assim, o treinamento dos modelos inteligentes foi realizado com $n = 5$, separando em 5 conjuntos de treinamento com os 336, 672, 1008, 1344 e 1680 primeiros elementos da matriz de 2016 posições e a avaliação foi feita a partir da previsão dos seguintes 336 elementos.

2.6 Resultado dos Testes

A partir do treinamento descrito na seção anterior foi possível avaliar os modelos inteligentes com as métricas RMSE e MAE no conjunto de teste, os melhores resultados alcançados podem ser observados na Tabela 1, nesta constam o modelo, o melhor número de partições usadas (MNPU) no treinamento de tal modelo e as métricas RMSE e MAE.

Tabela 1: Resultados obtidos

Modelo	MNPU	RMSE	MAE
Árvore de Regressão	5	69,32	49,01
Máquinas de Vetores de Suporte	5	51,60	32,07
Rede neural Artificial	5	51,08	32,46
Regressão Linear	3	41,67	30,77
Lasso	3	40,28	29,97
Rede Long Short-term Memory	3	35,68	23,84

3. CONSIDERAÇÕES FINAIS

É possível verificar na Tabela 1 que a Rede LSTM obteve a melhor performance entre os modelos testados e suas métricas foram baixas quando comparadas com a escala dos dados.

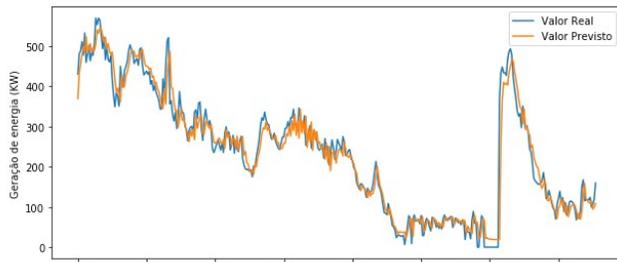


Figura 2: Previsão da Rede LSTM nos exemplos de teste

Na Figura 3, que foi criada a partir da previsão e do modelo sobre os exemplos de teste, é visível que o modelo conseguiu se adequar bem aos dados de treino e generalizar para os dados de teste, tendo em vista que conseguiu acompanhar o comportamento da série temporal.

A partir destes resultados, o modelo gerado por rede LSTM foi empregado no desenvolvimento de um sistema para monitoramento operacional da usina. O sistema foi desenvolvido a partir de que os operadores destas instalações enfrentam várias dificuldades, entre elas os altos custos dos processos de manutenção do equipamento.

SOUZA, Manoel V. F. et al.

A manutenção desse equipamento requer o desligamento da turbina do sistema e a utilização de tempo e recursos importantes para a desmontagem do aerogerador para manutenção. Por outro lado, o fabricante dá uma referência da energia que deve ser gerada pelo equipamento (conjunto de tabelas), os aerogeradores instalados em Gibara tem um comportamento diferente aos valores apresentados pelo fabricante. Esta diferença está relacionada com as características próprias do território onde estão instalados: o comportamento do clima, as habilidades dos operários, as adaptações feitas ao equipamento para trabalhar como 60 Hz, o tempo de uso das usinas, a localização geográfica delas, entre outros fatores que ainda estão em estudos por um grupo de pesquisadores da Universidade de Holguín.

Nesta planta eólica são necessárias ferramentas que possam prever com maior certeza a energia que será gerada pelas usinas. Da mesma forma, obter uma ferramenta que possa servir como possível diagnóstico do trabalho geral das usinas. A partir da hipóteses que se os aerogeradores mantém um comportamento estável na geração com respeito à predição feita por um modelo pré-treinado, significa que o aerogerador está trabalhando corretamente, se o comportamento apresentado for muito diferente do modelo matemático, pode indicar um mal funcionamento do aerogerador e um possível dano posterior, chamando a atenção sobre a necessidade de manutenção ou diagnóstico mais detalhado. O modelo criado, testado e avaliado neste trabalho é a parte responsável pela verificação do comportamento da usina.

4. AGRADECIMENTOS

A presente pesquisa foi apoiada pelo Programa de Apoio à Iniciação Científica (PAIC/FAPEAM/UEA) 2018 à 2019, com o código de projeto 15.077.

REFERÊNCIAS

- [1] Ma Lei, Luan Shiyang, Jiang Chuanwen, Liu Hongling, and Zhang Yan. 2009. A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews* 13, 4 (2009), 915–920.
- [2] Katti Faceli, Ana Carolina Lorena, João Gama, André Carlos Ponce de Leon Carvalho, et al. 2011. *Inteligência Artificial: Uma abordagem de aprendizado de máquina*. (2011).
- [3] H Seltman. 2012. *Experimental Design for Behavioral and Social Sciences*. Online Text–Book(2012).
- [4] Nikhil Buduma and Nicholas Locascio. 2017. *Fundamentals of deep learning: Designing next-generation machine intelligence algorithms*. "O'Reilly Media, Inc."