

PREDIÇÃO DO POTENCIAL DE GERAÇÃO DE ENERGIA SOLAR A PARTIR DE DADOS METEOROLÓGICOS

Lucas Almeida, Thiago Rizuti

Departamento de Ciência da Computação
Universidade Federal de Juiz de Fora

RESUMO

Uma das grandes preocupações ambientais atualmente é a utilização de recurso renováveis. No que se refere a geração de energia, fontes energéticas renováveis, como a solar, são uma alternativa mais sustentável que aquelas que utilizam a queima de combustíveis fósseis. A energia elétrica gerada através de células fotovoltaicas, por exemplo, pode ter a potência gerada limitada a fatores meteorológicos. Este trabalho objetiva prever qual o potencial de geração de energia solar com o uso de dados meteorológicos, como temperatura, umidade, cobertura de nuvens, pressão atmosférica e precipitação. Para tal foi proposto a utilização de um *Support Vector Regressor* e duas abordagens para otimizar seus parâmetros: um algoritmo genético e uma busca presa-predador.

PALAVRAS CHAVE: energia, solar, verde, predição, aprendizagem, otimização

1 Introdução

A questão energética é uma das mais importantes na atualidade. Gerando discussões que envolvem todas as camadas da sociedade, desde a política, setores econômicos e a comunidade científica.

As fontes de energia são geralmente classificadas em três categorias: fósseis, renováveis ou nucleares. As fontes fósseis são as mais exploradas pela humanidade e as mais danosas ao meio ambiente. A queima de combustíveis fósseis é responsável por altas taxas de emissão de CO_2 , que contribuem para o efeito estufa e aquecimento global. Desta forma, com o constante crescimento de demanda energética e a preocupação com o desenvolvimento do planeta, mostra-se urgente o uso de fontes alternativas de energia. A melhor alternativa é o uso de fontes vindas de recursos renováveis, tais como a energia solar, eólica, maremotriz e geotérmica (Azadeh *et al.*, 2013).

Uma das formas de se transformar irradiação solar em energia elétrica é através de células fotovoltaicas. Devido a novas tecnologias, a sua produção tem se tornado cada vez mais barato e o potencial de geração de energia cada vez maior. De forma que, em certos países, usinas de células fotovoltaicas estão no participando no mercado tanto quanto outras formas de geração de energia (Fernandez-Jimenez *et al.*, 2012). Um exemplo de aplicação são *data centers*, que estão empregando este tipo de energia a fim de reduzir custos e empregar uma estratégia sustentável (Fernandez-Jimenez *et al.*, 2012).

Ao utilizar a energia solar, é desejável que se saiba a previsão da energia disponível em uma determinada data ou período. Este trabalho foca na previsão do potencial de geração de energia elétrica através de células fotovoltaicas utilizando dados meteorológicos. Para tal foi proposto o uso de um *Support Vector Regressor* (SVR) para realizar a regressão dos dados. Um algoritmo genético e uma busca presa-predador foram usados para encontrar os melhores parâmetros para o SVR e otimizar a previsão.

Na seção 2 são apresentados os dados utilizados para este problema. As seções 3, 4 e 5 descrevem as abordagens propostas. A seção 6 apresenta os experimentos realizados e seus respectivos resultados. E na seção 7, são discutidas as conclusões e possíveis trabalhos futuros.

2 Dados Meteorológicos

Neste trabalho foram utilizados dados da base do *Measurement and Instrumentation Data Center* do *National Renewable Energy Laboratory*. Foram extraídos dados meteorológicos e a irradiação nas placas fotovoltaicas dos períodos de 10/11/2013 à 10/11/2014 e de 10/11/2015 à 10/11/2016. A tabela abaixo apresenta os dados utilizados e suas respectivas unidades de medida.

Tabela 1: Dados utilizados

Atributo (médias diárias)	Unidade
Irradiação	kWh/m^2
Temperatura	$^{\circ}C$
Umidade Relativa	%
Cobertura de Núvens	%
Pressão Atmosférica	$mbar$
Precipitação	mm

A irradiação diária média é a variável dependente a ser regredida, enquanto as medidas meteorológicas são as variáveis independentes usadas no processo.

3 Support Vector Regressor

O algoritmo *Support Vector Regressor* (SVR) (Drucker *et al.*, 1997) é um algoritmo de regressão linear criado a partir do algoritmo *Support Vector Machine*. O objetivo desse algoritmo é

aproximar parâmetros $\mathbf{w} \in \mathbb{R}^n$ e $b \in \mathbb{R}$ tais que para todos os pontos do conjunto de treinamento representados como $\{(\mathbf{x}_i, y_i) : i = 1, \dots, m\}$ a Equação 1 seja satisfeita:

$$|y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle + b| < \varepsilon \quad (1)$$

na qual ε é um valor pré-estipulado e representa o erro tolerável.

Para encontrar os parâmetros deve-se resolver o problema de otimização representado na Equação (2):

$$\begin{aligned} &\text{minimizar} \quad \frac{1}{2} \|\mathbf{w}\|^2 + C \sum \xi_i \\ &\text{sujeito a} \quad y_i - \langle \mathbf{w}, \mathbf{x}_i \rangle + b \leq \epsilon + \xi_i \\ &\quad \quad \quad \langle \mathbf{w}, \mathbf{x}_i \rangle + b - y_i \leq \epsilon + \xi_i \\ &\quad \quad \quad \xi_i > 0 \end{aligned} \quad (2)$$

Neste trabalho foi utilizada a SVR implementada na biblioteca *scikit-learn* do *Python*.

4 Algoritmo Genético

4.1 Representação do Indivíduo

Foi utilizada uma abordagem de algoritmo genético com valores reais (Wu *et al.*, 2007), no qual valores são codificados diretamente no cromossomo. Desta forma, o cromossomo é representado pela dupla de parâmetros da máquina de vetor suporte, como mostrado na Equação 3.

$$X = \{C, \gamma, \epsilon\} \quad (3)$$

4.2 Função de aptidão

A função de aptidão avalia a qualidade de cada um dos indivíduos da população. Neste caso, foi usado o inverso da raiz do erro médio quadrático da SVR treinada com os valores do indivíduo. Assim, quanto menor o erro da SVR, mais apto é o indivíduo.

$$aptidao = \frac{1}{erro} \quad (4)$$

4.3 Seleção

Foi proposta o uso da seleção por roleta, aleatório proporcional à aptidão. Ou seja, quanto maior a aptidão do indivíduo, maior sua chance de ser selecionado para reprodução.

4.4 Crossover

O cruzamento é feito de forma que dois indivíduos X_1 e X_2 geram dois novos indivíduos X_3 e X_4 . Um dos dois operadores mostrados a seguir é aplicado, escolhido de forma aleatória de forma aleatória. Onde σ é um micro número aleatório.

$$X_3 = X_1 + \sigma(X_1 - X_2) \quad X_4 = X_2 - \sigma(X_1 - X_2) \quad (5)$$

$$X_3 = X_1 + \sigma(X_2 - X_1) \quad X_4 = X_2 - \sigma(X_2 - X_1) \quad (6)$$

4.5 Mutação

O operador de mutação é feito sobre um dos valores do indivíduo escolhido aleatoriamente. O operador é mostrado pela equação abaixo, onde x_n é o valor escolhido, LB_n é seu limite inferior, UB_n seu limite superior e r é um valor aleatório no intervalo (0,1).

$$x_n = LB_n + r(UB_n - LB_n) \quad (7)$$

4.6 Substituição

A cada geração é feita a substituição pelos novos indivíduos na população. É usada uma taxa de elitismo de 0.15, isto é, a nova população é formada pelos melhores 15% dos indivíduos da população atual. O restante são os novos indivíduos gerados à partir do cruzamento.

5 Busca Presa-Predador

A Busca Presa-Predador é uma (Tilahun e Ong, 2015) é uma meta-heurística bio inspirada. O algoritmo se baseia nas interações entre presas e predadores nas quais ocorrem perseguição das presas pelos predadores, e.g., a predação de zebras por leões.

Assim como no algoritmo genético o indivíduo é dado pela Equação 3.

O algoritmo pode ser definido pelo seguinte conjunto de passos:

1. Gerar um conjunto de soluções iniciais.
2. Calcular medida de sobrevivência e determinar predador e presas.
3. Mover o predador e as presas no espaço de soluções.
4. Verificar o critério de parada, caso não tenha ocorrido voltar ao passo 2.

5.1 Medida de Sobrevivência

A medida de sobrevivência é um valor que representa a capacidade da presa escapar do predador. Esse valor é inversamente proporcional a função objetivo em problemas de minimização. Neste trabalho a medida de sobrevivência é inversamente proporcional à raiz do erro médio quadrático.

Após o cálculo da medida de sobrevivência a solução com menor medida de sobrevivência será o predador.

5.2 Movimentos

A distância movida pelas presas e pelo predador são calculadas a partir de variáveis de entradas λ_{max} e λ_{min} que definem, respectivamente, o passo grande e o passo pequeno e de variáveis aleatórias no intervalo (0,1).

O predador se movimenta de modo a combinar um movimento aleatório e a perseguição à presa com menor medida de sobrevivência.

As presas podem ser mover de três maneiras diferentes: fugindo aleatoriamente do predador, seguindo as outras presas ou – para a presa na dianteira, i.e., com maior medida de sobrevivência – buscando melhores caminhos. A presa na dianteira sempre buscara melhores caminhos e as outras seguiram o grupo com uma dada probabilidade p_{follow_up} e com probabilidade $1 - p_{follow_up}$ fugiram aleatoriamente.

Para emular a direção de uma presa fugindo aleatoriamente é sorteado um vetor unitário y_r e a presa segue pela direção apontada por esse vetor ou na direção contrária $-y_r$ dependendo de qual desses movimentos se distancia do predador.

Quando a presa i resolve seguir as outras é calculado um vetor unitário y_i que é a direção média ponderada pela medida de sobrevivência das outras presas em relação a i . A presa i segue nessa direção com um passo inversamente proporcional a sua medida de sobrevivência e com um pequeno passo na direção aleatória y_r .

Por fim, a busca por melhores caminhos é uma busca local realizada sorteando direções e escolhendo a posição com maior medida de sobrevivência.

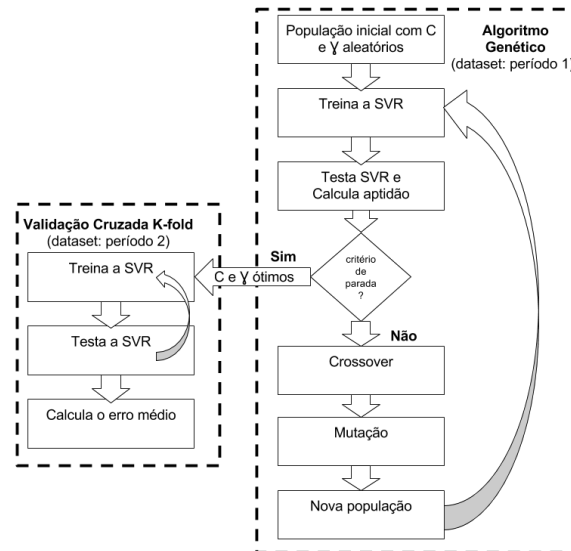
6 Experimentos

Foram realizados testes para cada uma das abordagens propostas: algoritmo genético e busca presa-predador. A seguir são apresentados os experimentos e os resultados para cada uma delas.

6.1 Algoritmo Genético

Para o algoritmo genético, os testes feitos com o 500 gerações e tamanho da população de 100 indivíduos. O treinamento é feito com os dados do período de 2015 à 2016. Após encontrar os parâmetros a avaliação final é feita com os dados de 2013 à 2014 utilizando validação cruzada. Como apresentado na Figura 6.1.

Figura 1: Experimento para o Algoritmo Genético



Foram feitos testes variando as taxas de mutação e *crossover*, realizando dez execuções independentes para cada configuração. A Tabela 2, a seguir, apresenta os resultados obtidos, para cada configuração é mostrada a margem de erro média.

Tabela 2: Resultados obtidos para o AG.

Taxa de Mutação	Taxa de <i>Crossover</i>	Margem de erro média
1	1	16,9%
0.05	0.95	16,9%
0.1	0.9	17,5%
0.15	0.85	17,3%
0.2	0.8	17,1%
0.5	0.5	17,0%
0.3	0.7	17,1%
0.5	0.8	16,9%

6.2 Busca Presa-Predador

Assim como os testes com o algoritmo genético, para a busca presa-predador a primeira etapa do algoritmo é a busca por parâmetros efetuada usando o conjunto de dados referente ao período de 2015 à 2016. Quando chega-se à condição de parada é realizado o aprendizado do SVR utilizando os parâmetros encontrados para realizar validação cruzada no conjunto de dados referentes ao período de 2013 e 2014. A condição de parada para este o algoritmo presa-predador é a realização de 500 movimentos por indivíduo e no total existem 100 indivíduos na população a todo o momento.

Em (Tilahun e Ong, 2015) são citados os vários parâmetros do algoritmo presa-predador. O número de parâmetros torna inviável testes utilizando variações de todos os parâmetros. Por

isso decidiu-se manter alguns parâmetros fixos e variar outros parâmetros que influenciam mais o algoritmo, a saber, λ_{max} , λ_{min} e p_{follow_up} . Além disso, para fins comparativos usou-se

Tabela 3: Resultados obtidos para o Algoritmo Presa-Predador.

λ_{max}	λ_{min}	p_{follow_up}	Margem de erro média
0.1	0.01	0.8	23,4%
0.1	0.001	0.8	22,7%
0.01	0.0001	0.8	21,6%
0.1	0.0001	0.8	23,1%
0.1	0.01	0.2	23,4%
0.1	0.001	0.2	23,3%
0.01	0.0001	0.2	22,3%
0.1	0.0001	0.2	23,2%

7 Conclusão

Os resultados obtidos estão aquém do esperado pois existe pouca melhora após a aplicação de ambos os métodos no conjunto de dados iniciais. Entretanto, ao analisar as Tabelas 2 e 3 é perceptível que o algoritmo genético alcançou um desempenho superior ao algoritmo presa-predador o que alude a um possível equívoco ao não variar todos os parâmetros deste.

Por fim, conclui-se que embora os resultados deste trabalho possam não ser uteis para algumas tarefas como previsão de energia solar em *data centers*, porém ainda é possível utilizá-los em outros contextos como a previsão de energia solar média anual em uma localidade.

Referências

- National Renewable Energy Laboratory measurement and instrumentation data center. <http://www.nrel.gov/midc/>. Acessado em 22/01/2017.
- Azadeh, A., Babazadeh, R. e Asadzadeh, S. M.** (2013), Optimum estimation and forecasting of renewable energy consumption by artificial neural networks. *Renewable and Sustainable Energy Reviews*, v. 27, p. 605–612.
- Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. e Vapnik, V.** Support vector regression machines. *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS* 9, p. 155–161. MIT Press, 1997.
- Fernandez-Jimenez, L. A., Muñoz-Jimenez, A., Falces, A., Mendoza-Villena, M., Garcia-Garrido, E., Lara-Santillan, P. M., Zorzano-Alba, E. e Zorzano-Santamaria, P. J.** (2012), Short-term power forecasting system for photovoltaic plants. *Renewable Energy*, v. 44, p. 311–317.
- Tilahun, S. L. e Ong, H. C.** (2015), Prey-predator algorithm: A new metaheuristic algorithm for optimization problems. *International Journal of Information Technology & Decision Making*, v. 14, n. 06, p. 1331–1352.
- Wu, C. H., Tzeng, G. H., Goo, Y. J. e Fang, W. C.** (2007), A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. *Expert Systems with Applications*, v. 32, n. 2, p. 397–408.