

Previsão de comportamento de manchas solares utilizando abordagens de redes neurais

Sunspot behavior forecast using neural networks approaches

RESUMO

Denise Pechebovicz
pechebovicz@alunos.utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil

Marcella Scoczynski Martins
marcella@utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil

Sthefanie Premebida
sthefanie@alunos.utfpr.edu.br
Universidade Tecnológica Federal do Paraná, Ponta Grossa, Paraná, Brasil

O estudo da atividade solar é muito importante para a predição de eventos solares e terrestres, permitindo termos uma melhor preparação para a ocorrência destes eventos. Em nosso estudo, utilizamos redes neurais, ELM e RBF, e técnicas de aprendizagem de máquina para prever número de manchas solares e uma previsão da atividade solar com base nos dados gravados do número de manchas solares diárias entre 1818 e 2019. Estes foram obtidos no site Solar Influences Data analysis Center (SIDC) e Sunspot Index and Long-term Solar Observations (SILSO). Os resultados mostram um elevado potencial de processamento, principalmente no modelo híbrido ELM + RBF que pode se tornar uma proposta competitiva para a previsão manchas solares.

PALAVRAS-CHAVE: Manchas Solares. Rede Neural. Astronomia e Astrofísica.

ABSTRACT

The study of solar activity is very important for the prediction of solar and earthly events, with the greater preparation for the occurrence of these events. In our study, we use neural networks, ELM and RBF, and machine learning techniques to predict the number of sunspots, thereby the solar activity based on the recorded data of the number of daily stains between 1818 and 2019. These were obtained on the website Solar Influences Data analysis Center (SIDC) and Sunspot Index and Long-term Solar Observations (SILSO). The results show a high processing potential, mainly in the hybrid model ELM + RBF, which becomes a competitive proposal for forecasting sunspots.

KEYWORDS: Sunspots. Neural Network. Astronomy and Astrophysics.

Recebido: 19 ago. 2020.

Aprovado: 01 out. 2020.

Direito autoral: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



INTRODUÇÃO

Atualmente, o principal parâmetro de referência para medir a atividade solar usa como base o número de manchas solares presentes no Sol a determinado momento. A atividade solar é um problema clássico que demonstra alta importância devido à exploração espacial. Satélites de baixa órbita requerem proteção contra o vento solar e partículas de alta energia, astronautas pelo espaço exigem monitoramento de campo magnético para a proteção da saúde contra radiação solar. Além disso, radiocomunicação, cabos de redes e energia elétrica na Terra também sofrem interferência da atividade solar (N. SAFIULLIN, S. PORSHNEV, E N. KLEEORIN, 2018).

No intuito de fazer uma nova predição do número de manchas solares, facilitar o conhecimento de futuros eventos solares e prever máximos e mínimos da atividade solar, foi utilizado métodos computacionais e de aprendizado de máquina, neste estudo foram utilizados Função de Base Radial (RBF) e Máquina de Aprendizado Extremo (ELM).

Aprendizado de máquina, do inglês Machine Learning, é uma inteligência artificial que com algoritmos reconhece padrões e se torna capaz de fazer previsões, um sistema de aprendizagem de máquina, podendo ser empregado em diversas áreas envolvendo banco de dados grandes e complexos sendo capaz de entregar um resultado rápido e preciso. Neste trabalho utilizamos no banco de dados 73565 dias registrados pelo Solar Influences Data analysis Center (SIDC).

O nosso objetivo com esse artigo foi fazer um método de previsão de manchas solares, com isso, uma previsão da atividade solar permitindo uma maior preparação para eventos solares, como as ejeções de massa coronal, que são liberações de grandes quantidades de matéria da atmosfera solar, resultado de instabilidades magnéticas em larga escala associada a erupções solares. Também o vento solar, que é um fluxo contínuo de partículas energizadas por altas temperaturas, constituídas principalmente por prótons e elétrons que escapam da atração gravitacional do Sol. Esses eventos solares resultam em várias interações com a Terra, e essas interações resultam em algumas complicações, como comprometimento da saúde dos astronautas, interferência nas telecomunicações e em equipamentos elétricos e satélites (D. H. HATHAWAY, 2015).

A principal contribuição deste trabalho é a aplicação de um modelo de rede neural recentemente desenvolvido que engloba a Função de Base Radial (RBF) e a Máquina de Aprendizagem Extremo (ELM) juntas na mesma arquitetura.

MATERIAIS E MÉTODOS

A emissão da energia Solar é gerada a partir de reações nucleares no núcleo do Sol onde átomos de hidrogênio se fundem e essa emissão varia na posição do Sol e no tempo, recebendo influência do campo magnético terrestre. Observações de manchas solares, mostraram que a área que elas cobrem cresce rapidamente de um mínimo (próximo de zero) ao máximo em 3 a 4 anos, e a redução do máximo ao mínimo é mais lenta (D. H. HATHAWAY, 1994).

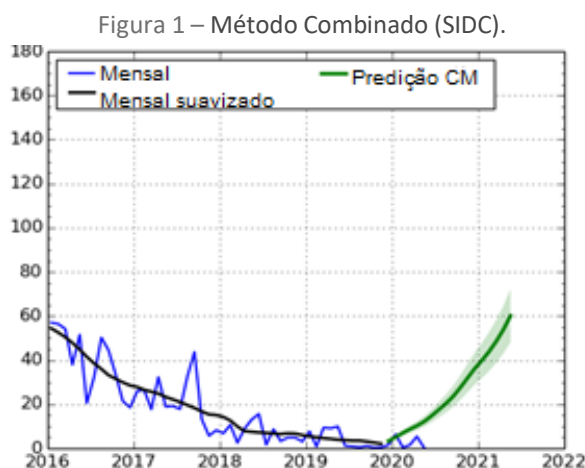
Um ciclo solar completo dura em média 22 anos, sendo conhecido como ciclo de Hale (SAFIULLIN, PORSHNEV, KLEEORIN, 2018), após um ciclo solar de

aproximadamente 11 anos, conhecido como ciclo solar de Schwabe, o Sol inverte sua polaridade magnética e a cada duas inversões é contabilizado como um ciclo solar completo (GIOVANNELI, 1939). Porém para a contabilização de ciclos solares, é baseado no ciclo solar de Schwabe, o qual atualmente estamos no ciclo 25, iniciado em 2020.

Atualmente, a maior referência para a medição da atividade solar é o número de manchas solares presentes no sol em determinado momento. Desde 1981 a análise de imagens fornecidas por satélites e observatórios solares, para monitoramento automatizado, está sendo feita pelo Centro de Análise de Dados de Influências Solares (SIDC) que produz mensalmente o número internacional relativo de manchas solares, do inglês Relative Sunspot Number, R_i . Esse é estatisticamente calculado a partir de contribuintes e observações disponíveis entre observadores profissionais e astrônomos amadores (F. CLETTE et al, 2007).

O principal observatório para previsão de manchas solares é o Observatório de índice de manchas solares e observações solares de longo prazo (SILSO), que desempenha o papel de disseminar e fazer estudos mais aprofundados sobre o número de manchas solares e ciclos solares. Este observatório vem produzindo previsões da atividade solar diariamente, mensalmente e de ciclos solares.

Para isso são utilizados três métodos de previsão de médio prazo, isto é, valores médios semestrais dos números de manchas solares que estão atualmente em operação em tempo real: o método McNish-Lincoln (NGDC) (A. MC NISH and J. LINCOLN, 1949), o método Curvas Padrão (HATHAWAY, WILSON, REICHMANN, 1994) e o método combinado (K. DENKMAYR and P. CUGNON, 1997). A precisão da previsão para o método McNish-Lincoln é aumentada em 17 a 30%, no método padrão em 5 a 21% e no método combinado em 6 a 57%. O método combinado está representado na figura 1.



Fonte: <http://www.sidc.be/silso/predicm> (30 de junho de 2020) - Traduzido.

O método combinado consiste no ajuste do quadrado mínimo e na interpolação em todos os ciclos registrados desde 1749. Onde a linha azul representa os dados mensais, linha preta os dados mensais suavizados e linha verde representa a previsão pelo Método Combinado. O ajuste é feito no número de manchas solares mensais suavizadas, observadas em 13 meses com todos os valores disponíveis desde o último mínimo ou o último máximo do ciclo. A previsão

começa 5 meses antes do último mês decorrido e dura mais 18 meses até 12 meses à frente (K. DENKMAYR and P. CUGNON, 1997).

Uma maneira de analisar e tentar encontrar possíveis padrões de comportamentos das manchas solares seria utilizando um grande banco de dados para análise, mas essa alternativa utiliza de um longo tempo de treinamento de rede e para resultados mais rápidos utilizamos métodos computacionais e de aprendizado de máquina, que neste caso são utilizados Função de Base Radial (RBF) e Máquina de Aprendizado Extremo (ELM).

A função de Base Radial (RBF) é uma rede neural linear e tem a sua aprendizagem supervisionada, ou seja, quando a rede aprende com o conjunto de dados já definido e as entradas e saídas são conhecidas (S. S. HAYKIN et al, 2009). Já Máquina de Aprendizado Extremo (ELM) pode ser dita como uma versão simplificada das redes neurais e é uma rede randômica que pode ser utilizada para classificação, regressão, aproximação (HUANG, ZHU e SIEW, 2006), dentre outras aplicações (H. SIQUEIRA et al, 2014), (H. SIQUEIRA et al, 2018).

O primeiro passo foi organizar os dados a serem usados pela inteligência artificial. Foram utilizados os dados do SIDC e SILSO, onde 73.565 dias, de 1º de janeiro de 1818 à 31 de maio de 2019, foram analisados. Os dados primários foram divididos em entradas e saídas e divididos em três partes: treino, validação e teste. As proporções escolhidas foram as seguintes: Treinamento 70%; conjunto de validação 15%; conjunto de testes 15%.

O processo de aprendizagem ocorre através de um número definido de épocas, onde o conjunto de treinamento é submetido a um loop de adivinhar os resultados, sugerindo os mesmos e minimizando os erros conforme os acertos. Após inúmeras épocas o conjunto de validação é usado para verificar se os resultados calculados são satisfatórios e também analisar o comportamento da rede quando submetida a um novo conjunto de dados. Se esta etapa não for executada a rede responderá perfeitamente ao analisar os dados treinados, porém perde desempenho ao tentar prever com base em novos dados, nosso principal objetivo.

RESULTADOS E DISCUSSÃO

Dependendo dos resultados, a rede pode ser declarada treinada e pronta para executar a previsão sobre o conjunto de dados de teste. Após o conjunto de testes ter sido enviado para as redes treinadas, os dados e resultados esperados foram comparados e usamos o erro quadrático médio (MSE) como um indicativo de desempenho, conforme mostra a equação (1).

$$MSE = \frac{1}{2nd} \sum_{i=0}^{nd} (Rsv - Psv)^2 \quad (1)$$

Onde nd representa o número de dias, Rsv o valor diário real da quantidade de manchas e Psv o valor diário real previsto de quantidade de manchas. O processo de treinamento executa três redes neurais diferentes com os mesmos recursos. Os melhores resultados foram obtidos ao juntar as duas redes.

A tabela 1 apresenta os melhores e piores MSE para o processo de teste com diferentes quantidades de neurônios para o modelo de junção ELM e RBF, modelo híbrido que obtivemos melhores resultados.

Tabela 1 – ELM + RBF MSE

Quantidade de neurônios	Melhor	Pior
500	0.0213	0.0245
1000	0.0109	0.0127
1500	0.0096	0.0103
2000	0.0078	0.0084

Fonte: Autoria Própria (2020).

CONCLUSÃO

Nós estendemos nossa análise comparando a rede ELM + RBF com um trabalho relativo, comparando a acurácia da predição. Esses resultados mostram que nosso método pode ter uma performance competitiva com demais métodos de previsão da atividade solar. Porém, conforme o SIDC, métodos baseados puramente na atividade solar passada, torna-se não confiável no final de cada ciclo e durante os mínimos, mas possui boa performance no meio de cada ciclo.

Acreditamos que com essa pesquisa, a aplicação de aprendizado de máquina no treinamento das redes neurais, pode trazer grande avanço no estudo da atividade solar, a partir da otimização da predição de manchas solares. A predição da atividade solar permite uma maior preparação para eventos solares que trazem complicações na Terra e também a satélites, astronautas, entre outros.

AGRADECIMENTOS

Denise Pechebovicz agradece a Marcella Martins, Virginia Baroncini, Hugo Siqueira pela ajuda no desenvolvimento dessa pesquisa como também ajuda financeira e logística para apresentações em congressos e aos membros da equipe de iniciação científica, Sthefanie Premevida, Thiago Camargo e Vinícios Soares, pelo apoio e trabalho realizado. Agradece também pela participação no Programa de Voluntariado em Iniciação Científica e Tecnológica (Edital PROPPG – PIVICT 2019/2020) concedida ao acadêmico.

REFERÊNCIAS

R. G. Giovanelli, “**The relations between eruptions and sunspots.**” The Astrophysical Journal, vol. 89, p. 555, 1939.

N. Safiullin, S. Porshnev, e N. Kleeorin, “**Monthly sunspot numbers forecast with artificial neural network combined with dynamo model: comparison with modern methods**”, em 2018 Ural Symposium on Biomedical Engineering, Radio electronics and Information Technology (USBREIT). IEEE, 2018, pp. 199–202.

D. H. Hathaway, “**The solar cycle**”, Living reviews in solar physics, vol. 12, no. 1, p. 4, 2015

R. G. Giovanelli, “**The relations between eruptions and sunspots**”. The Astrophysical Journal, vol. 89, p. 555, 1939.

D. H. Hathaway, “**The solar dynamo**”, NASA Technical Report NASA-TM-111102, NAS 1.15:111102, 1994.

F. Clette, D. Berghmans, P. Vanlommel, R. A. Van der Linden, A. Koeckelenbergh, and L. Wauters, “**From the wolf number to the international sunspot index: 25 years of Sids**”, Advances in Space Research, vol. 40, no. 7, pp.919–928, 2007.

D. H. Hathaway, “**The solar cycle**”, Living Reviews in Solar Physics, vol. 7, no.1, p. 1, Mar 2010. [Online]. Página: <https://doi.org/10.12942/lrsp-2010-1>.

D. H. Hathaway, R. M. Wilson, e E. J. Reichmann, “**The shape of the sunspot cycle**”, Solar Physics, vol.151, no. 1, pp. 177–190, 1994.

A. Mc Nish e J. Lincoln, “**Prediction of sunspot numbers**”, Eos, Transactions American Geophysical Union, vol. 30, no. 5, pp. 673–685, 1949.

K. Denkmayr and P. Cugnon, “**About sunspot number medium-term predictions**”, in Solar-Terrestrial Prediction Workshop V, Japan, vol. 103, 1997.

S. S. Haykin et al. “**Neural networks and learning machines**”. New York: Prentice Hall, 2009.

G.-B. Huang, Q.-Y. Zhu, e C.-K. Siew, “**Extreme learning machine: theory and applications**”, Neuro computing, vol. 70, no. 1-3, pp. 489–501, 2006.

G.-B. Huang, D. H. Wang, and Y. Lan, “**Extreme learning machines: a survey**”, International journal of machine learning and cybernetics, vol. 2, no. 2, pp. 107–122,2011.

H. Siqueira, L. Boccato, R. Attux, e C. Lyra, “**Unorganized machines for seasonal stream flow series forecasting**”, International journal of neural systems, vol. 24, no. 03, p. 1430009, 2014.

H. Siqueira, L. Boccato, I. Luna, R. Attux, e C. Lyra, “**Performance analysis of unorganized machines in stream flow forecasting of brazilian plants**”, Applied Soft Computing, vol. 68, pp. 494–506, 2018.