

## Aplicação de redes neurais ART auto-expansíveis na detecção de pulsares

Vitor H. C. Luz<sup>1</sup>

DCC/ICEx/UNIFAL-MG, Alfenas, MG

Angela Leite Moreno<sup>2</sup>

DEMAT/ICEx/UNIFAL-MG, Alfenas, MG

Pulsares são estrelas de nêutrons com padrões únicos de eletromagnetismo, cujo campo magnético é da ordem de  $10^8$ . A medida que um pulsar gira seu campo magnético, induz um campo elétrico na sua superfície. Este campo elétrico é suficiente para arrancar partículas (elétrons) carregadas da superfície que, por sua vez, fluem para a magnetosfera onde são aceleradas, emitindo um feixe estreito ao longo das linhas do campo magnético. Quando o eixo do campo magnético coincide com a linha de visada da Terra é possível observar um pulso de radiação eletromagnética. Ao se estudar pulsares obtém-se uma maior compreensão de problemas físicos como, por exemplo, matéria escura, campos magnéticos galácticos e meio interestelar [1].

A captura dos dados de candidatos a pulsar é feita por radiotelescópios e, neste sentido, muitos ruídos são captados. O processo de validação de um candidato a pulsar é feito por um especialista que precisa observar o candidato por um longo período de tempo. Deste modo, o investimento em equipamentos tecnológicos e do tempo do especialista dedicado a uma única tarefa tornam esse processo de alto custo. Devido à necessidade de concentração por um longo período de observação e que o tempo necessário à validação de cada pulsar é variável, muitos candidatos que são pulsares acabam não sendo validados e, desta forma, retornam à base de dados. Com esse retorno à base de dados e descobertas de novos astros por radiotelescópios a quantidade de candidatos tornou inviável que especialistas verifiquem todos os candidatos, sendo necessário a utilização de modelos computacionais para a filtragem desses dados [3]. Nesta perspectiva, alguns modelos de classificação já vem sendo utilizados na abordagem desse problema utilizando a base de dados HTRU2 como, redes nebulosas KNN, Redes Kohonen, Random Forest. É neste contexto que este trabalho está inserido e tem como objetivo apresentar os resultados obtidos pelas Redes Neurais ART Auto-Expansíveis para o problema da detecção de pulsares.

A base de dados utilizada neste trabalho foi HTRU 2 [2] com 17898 exemplares de pulsares dos quais 1639 (9,157%) são classificados como pulsar e 16259 (90,843%) são classificados como não pulsar. Os resultados foram obtidos com aplicação de duas técnicas bastante difundidas na inteligência artificial: a busca exaustiva e a validação cruzada. A busca exaustiva consiste em explorar uma combinação de parâmetros de ajuste do modelo que maximizam sua performance. Neste trabalho, a busca ocorreu em relação aos parâmetros  $\alpha \in \{0, 1; 0, 13; 0, 2; 0, 25\}$  e  $\beta \in \{0, 1; 0, 13; 0, 2; 0, 25; 0, 30; 0, 35\}$ , sendo  $\rho$  variando entre 0,01 até 0,99 com passo de 0,001. A Rede Neural ART Fuzzy Auto-Expansível possui os três parâmetros, enquanto a Euclidiana possui apenas os parâmetros  $\beta$  e  $\rho$  [4]. A validação cruzada é quando dividimos um conjunto de dados em vários subconjuntos ( $k$ ), no caso 10, e fazemos o treinamento com  $k - 1$  e uma análise com 1, sem intersecção dos dados até que todos os  $k$  subconjuntos tenham sido usados para análise.

---

<sup>1</sup>vitorhugoclz@gmail.com

<sup>2</sup>angela.moreno@unifa-mg.edu.br

Essa técnica é bastante relevante para evitar *overfitting* do modelo e também com ela é possível verificar se o modelo está tendo resultados discrepantes dependendo dos dados apresentados.

Para interpretação do resultado podem ser utilizadas inúmeras métricas. Para estas redes duas métricas foram maximizadas a acurácia e a sensibilidade e podem ser observados na Tabela 1.

Tabela 1: Principais Resultados

	$\beta$	$\alpha$	$\rho$	ACC	Sens	MCC
ART Fuzzy Auto-Expansível	0,25	0,25	0,988	<b>97,45%</b>	77,85%	0,8387
ART Fuzzy Auto-Expansível	0,13	0,10	0,997	96,60%	<b>79,68%</b>	0,7927
ART Euclidiana Auto-Expansível	0,10	–	0,090	<b>97,23%</b>	78,70%	0,8257
ART Euclidiana Auto-Expansível	0,30	–	0,281	94,33%	<b>85,60%</b>	0,7177

Pode-se observar que ao se focar na sensibilidade da rede o modelo perde em acurácia, o mesmo acontecendo quando se foca na acurácia. Entretanto, o principal objetivo da rede é detectar candidatos a pulsar e, lembrando que dentro dos dados classificados como negativos, temos uma pequena porção de positivos rotulada erroneamente! Por outro lado, maximizar unicamente a sensibilidade sem levar em conta a acurácia e que pulsares são astros raros leva novamente a um problema de sobrecarga nos especialistas já que o modelo estaria indicando muitos candidatos.

Ao se focar na acurácia, os melhores resultados foram obtidos com a Rede ART *Fuzzy* Auto-Expansível, enquanto que ao se focar na sensibilidade a Rede ART Euclidiana Auto-Expansível obteve melhores resultados. Considerando a natureza do problema, a Rede Neural que mais se destaca para esse problema é a ART Euclidiana Auto-Expansível pois obteve uma sensibilidade de 85,6% sem perda significativa da acurácia. Por se tratar de dados desbalanceados, um fator que deve se levar em consideração é o Coeficiente de Correlação de Matthews (MCC) que em todas as previsões obteve um índice superior a 0,70, implicando que os modelos aqui discutidos são bons classificadores para o problema.

## Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, através do Programa Institucional de Bolsas de Iniciação Científica – PIBIC, e à Universidade Federal de Alfenas (UNIFAL-MG).

## Referências

- [1] Clark, H. A., Lewis, G. F. and Scott, P. Investigating dark matter substructure with pulsar timing - I. Constraints on ultracompact minihaloes, *Monthly Notices of the Royal Astronomical Society*, 456:1394-1401, 2016a. DOI: 10.1093/mnras/stv2743.
- [2] Dua, D. and Karra, T. E. *UCI Machine Learning Repository*: University of California, School of Information and Computer Science, 2017.
- [3] Lyon, R. J., Stappers, B. W., Cooper, S., Brooke, J. M. and Knowles, J. D. Fifty Years of Pulsar Candidate Selection: From simple filters to a new principled real-time classification approach, *Monthly Notices of the Royal Astronomical Society*, 459:1104-1123, 2016. DOI: 10.1093/mnras/stw656.
- [4] Moreno, A. L. *Redes Neurais ART e ARTMAP com Treinamento Continuado*. Saarbrücken: Novas Edições Acadêmicas, 2016.