

Rede Neural Artificial para previsão de propriedades de Nanoclusters de Pt_n ($n = 2 - 15$ e 55 átomos) em Python com Tensorflow + Keras

OCTÁVIO RODRIGUES FILHO¹; JOÃO PAULO CERQUEIRA FELIX²;
MAURÍCIO JEOMAR PIOTROWSKI³

¹Universidade Federal de Pelotas – octaviorodriguesf@gmail.com

²Universidade Federal de Pelotas – joao.felix@ufpel.edu.br

³Universidade Federal de Pelotas – mauricio.piotrowski@ufpel.edu.br

1. INTRODUÇÃO

Redes Neurais artificiais (RNAs) são modelos matemáticos (SVOZIL et al., 1997) baseados no funcionamento do cérebro, especificamente na atividade nervosa do neurônio (MCCULLOCH & PITTS, 1943) que, através da observação de seu comportamento, realizam aproximações matemáticas simplistas semelhantes ao seu processamento. RNAs podem ser utilizadas para predição de dados através do aprendizado com os erros via treinamento pelo algoritmo de retropropagação (RUMELHART et al., 1986), momento em que ocorre o reconhecimento de padrões que o cérebro humano não seria capaz de compreender. Nos últimos anos, observamos um aumento gradativo na utilização de *machine learning* (ML) nos campos da Física e Química (VON LILIENFELD & BURKE 2020), o qual se deve à capacidade de obtenção de propriedades preditivas e classificatórias de interesse científico das RNAs. Uma dessas possíveis aplicações seria na associação com a Teoria do Funcional da Densidade (DFT) (HOHENBERG & KOHN, 1964 e KOHN & SHAM, 1965), cujo foco reside na descrição atomística de sistemas físicos via Mecânica Quântica. Por sua vez, dentro do campo de estudo da DFT, encontramos os nanoclusters ou nanopartículas, que são aglomerados atômicos pequenos que compreendem a escala nanométrica ou sub-nanométrica (ou seja, o nanômetro é igual a um bilionésimo de metro) (ALONSO, 200). Sabendo que RNAs têm a propriedade fundamental de aproximação universal (CYBENKO, 1989 e HORNIK & STINCHCOMBE & WHITE, 1989), podemos aproximar qualquer função no R^n , gerando a possibilidade de existência de uma função simples que esteja em R^n retornar os mesmos ou semelhantes resultados da DFT. Dada a complexidade de tais cálculos, uma RNA simples não seria viável, logo, o emprego de múltiplas camadas se faz necessário para obtenção da não linearidade (SVOZIL et al., 1997), levando a uma predição que relacione tais dados. Entretanto, entradas simplistas não indicam que a rede é capaz de aprender DFT, mas sim, uma relação matemática que possivelmente possa existir entre esses dados. Portanto, fazer o treinamento e uma previsão além do aprendido, indicaria que a rede neural poderia saltar do limite estabelecido na previsão (por exemplo, clusters de 2 a 15 átomos), onde tal afirmação pode ser conferida, para uma predição de sistemas maiores (por exemplo, 55 átomos), tornando o mesmo um limiar para determinar a eficiência da RNA (em nosso caso, para metais de transição - MTs). Assim, o objetivo deste trabalho é investigar, estabelecer e criar uma relação entre as propriedades de nanoclusters de 2 a 15 átomos de MTs e para a energia de ligação (E_b) de Pt_{55} , com entradas simplistas, mas que tenham sentido devido à coerência física para classificar a espécie atômica. Com isso, espera-se prever

as propriedades dos nanoclusters através da criação de um modelo de RNA em Python com utilização de APIs (*Application Programming Interface*) auxiliares.

2. METODOLOGIA

A construção metodológica se deu através de APIs presentes no *Framework* da *Google Tensorflow* (versão 2.0), o qual é capaz de trabalhar com cálculo tensorial em conjunto com a biblioteca *Keras* (Tensorflow, 2023) para a construção de um modelo funcional com a capacidade de fazer camadas avançadas, aumentando a complexidade da RNA. Para obter um resultado de origem não linear, a função ativação escolhida foi a *SWISH* (RAMACHANDRAN et al., 2017) em conjunto com a *SIGMOID* (CYBENKO, 1989), onde as métricas foram o MSE (*Mean Square Error*), “*Validation Split*” de 20% e treinamento por 2000 vezes (*Epochs*) utilizando algoritmo ADAM (*Adaptive Moment Estimation*), todos originários do Tensorflow. Não existe uma topologia definida como o ideal para determinado objetivo em RNA (SVOZIL, 1997), logo, a rede foi baseada em camadas densamente conectadas com aplicações da técnica de *Dropout* (SRIVASTAVA et al., 2014). Foram utilizados para criação do banco de dados em CSV nanoclusters de 2 a 15 átomos de MTs (CHAVES et al., 2017). A rede não foi treinada com nenhum dado de Platina, pois são esses os dados que serão previstos, totalizando 11725 parâmetros de construção e 1218 dados de propriedades treinadas.

3. RESULTADOS E DISCUSSÃO

A estruturação do modelo não segue uma topologia pré-definida e sua montagem foi realizada de forma que os resultados das últimas camadas fossem específicos para cada dado de propriedade, ganhando o formato de uma RNA funcional. As propriedades escolhidas em ordem foram: energia de ligação (E_b), comprimento médio de ligação (d_{av}), número de coordenação efetivo (ECN), momento magnético total (m_t), gap Homo-Lumo (Gap). A topologia estrutural final é apresentada na Figura 1.

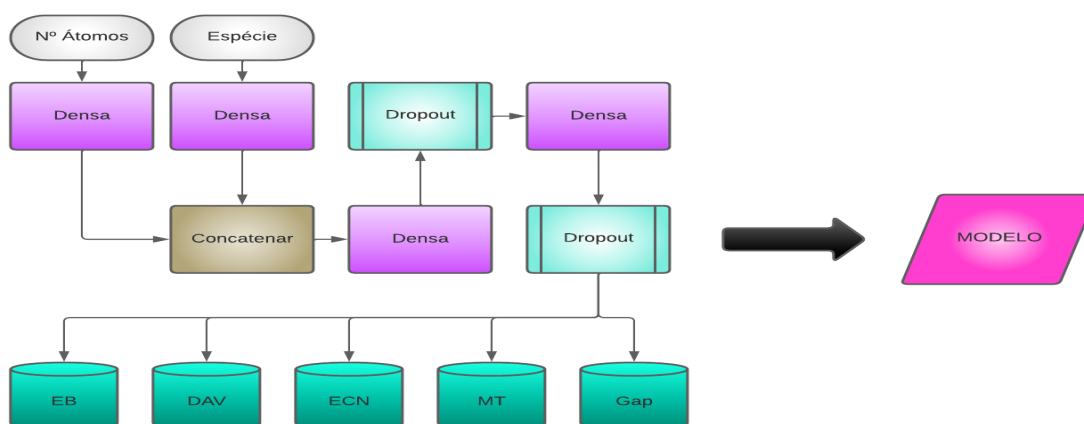


Figura 1: Estrutura da RNA para previsão de propriedades de nanocluster de metais de transição.

Após o treinamento, o modelo retorna o resultado de métricas impostas, como o erro MSE e a validação, onde foi configurado para que dos 1218 dados,

cerca de 20% fossem utilizados para validar o modelo, permitindo apontar sua capacidade preditiva. A Figura 2 apresenta os resultados para cada propriedade e o erro total.

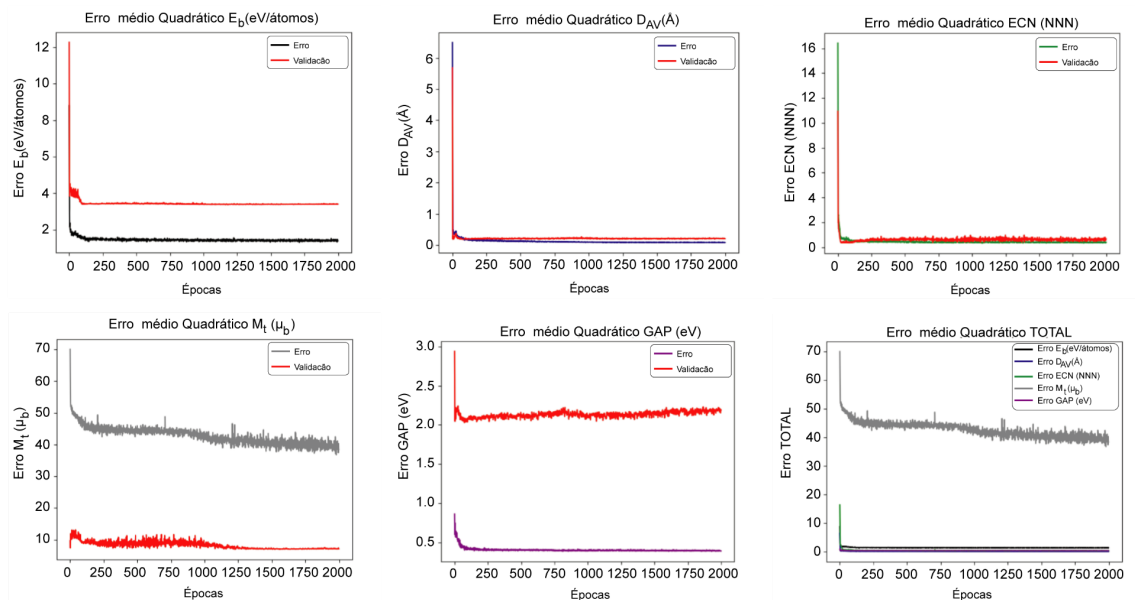


Figura 2: Resultados de treinamento: *Mean Square Error* e *Validation split* de 20%.

O modelo apresenta baixo *overfitting* (ou pode estar suavizado de maneira a não ser identificado com a quantidade de épocas impostas). O modelo foi capaz de realizar o treinamento, com destaque para m_t que, apesar do mais lento aprendizado, teve sua validação menor do que o erro. A E_b tem uma validação baixa, com um aprendizado com menos erros, indicando que, possivelmente, o gradiente pode estar oscilando em um mínimo local, tal fato é esperado quando tratamos de RNA com multicamadas (SVOZIL et al., 1997). Na Figura 3 mostra-se o comparativo das previsões com dados retirados da literatura.

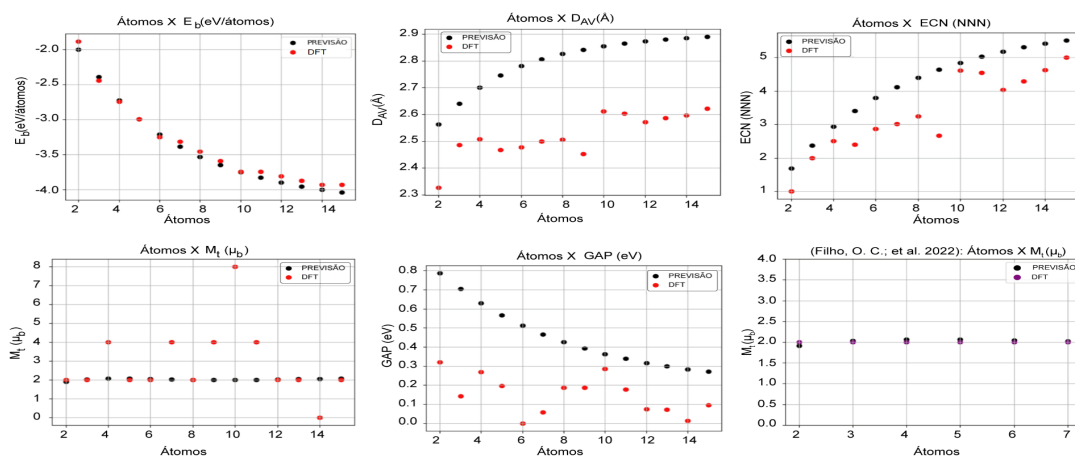


Figura 3: Propriedades em relação ao número de átomos: relação entre DFT e RNA.

4. CONCLUSÕES

O modelo proposto conseguiu prever E_b em acordo com a literatura, com uma média de erros para clusters de 2 a 15 átomos de Pt de 0,056, atingindo o objetivo principal. Para m_t , o modelo não foi satisfatório, porém, em comparação com outra referência (FILHO et al., 2022) (Figura 3) o modelo conseguiu atingir com êxito essa previsão. Tal fato indica que, se o conjunto de dados validados for diferente, outros resultados podem ser encontrados de acordo com o padrão reconhecido. Para as demais propriedades identificou-se desvios baixos no treinamento e médios na previsão. Para a última previsão, relativa à predição de E_b de Pt_{55} , o resultado foi de -4,19 eV, sendo que o valor obtido via DFT é de -4,47 eV. O que indica que o modelo poderia saltar cerca de 40 passos à frente do que foi treinado para tal propriedade, mas obtendo um aumento gradativo de erro médio, dado que a previsão de propriedades para 55 átomos apresentou uma diferença com a DFT de 0,28. Tal modelo promissor ainda passará por melhorias.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- SVOZIL, D.; et al. Introduction to multi-layer feed-forward neural networks. **Chemometrics and Intelligent Laboratory Systems**, v. 39, n. 1, p. 43-62, 1997.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, v. 5, p. 115-133, 1943.
- RUMELHART, D. E.; et al. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, 1986.
- VON LILIENFELD, O. A.; BURKE, K. Retrospective on a decade of machine learning for chemical discovery. **Nature Communications**, v. 11, n. 1, p. 4895, 2020.
- HOHENBERG, P.; KOHN, W. Inhomogeneous Electron Gas. **Physical Review B**, v.864, p. 136, 1964.
- KOHN, W.; SHAM, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. **Physical Review A**, v. 40, p. 11331, 1965.
- ALONSO, J. A. Electronic and Atomic Structure, and Magnetism of Transition Metal Clusters. **Chemical Reviews**, v. 100, p. 637-677, 2000.
- HORNIK, K.; et al. Multilayer feedforward networks are universal approximators. **Neural Networks**, v. 2, n. 5, p.359-366, 1989.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of control, Signals and Systems**, v. 2, n. 4, p. 303-314, 1989.
- RAMACHANDRAN, P.; et al. Searching for activation functions. **arXiv: preprint arXiv:1710.05941**, 2017.
- SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. **The Journal of Machine Learning Research**, v. 15, n. 1, p. 1929-1958, 2014.
- CHAVES, A. S.; et al. Evolution of the structural, energetic, and electronic properties of the 3d, 4d, and 5d transition-metal clusters (30 TM_n systems for $n = 2-15$): a density functional theory investigation. **Physical Chemistry Chemical Physics**, 2017, 19, 15484.
- Tensorflow.**Tensorflow[software]**. Disponível em: <https://www.tensorflow.com.org>. acesso em: 3 agosto 2023.
- Filho, O. C.; et al. Nanoclusters subnanométricos de Pt_n ($n = 2-7$ ÁTOMOS): Um estudo via DFT-PBE. **XXXI Congresso de Iniciação Científica**. 2022.