

# Aplicação do banco de dados de processos de mistura e reação química em microdispositivos a algoritmos de *machine learning*

Palavras-Chave: microfluídica, *machine learning*, *multilayer perceptron*

Wilson Legnaro Volpi Nakano [FEQ/Unicamp]

Prof. Dr. Osvaldir Taranto (orientador) [FEQ/Unicamp]

Dr. Harrson Silva Santana(coorientador) [FEQ/Unicamp]

## INTRODUÇÃO:

Recentemente o uso de microdispositivos vem crescendo na comunidade científica para diversas aplicações dentro da área da engenharia química. As vantagens derivadas da diminuição de escala de processos químicos, propiciou condições de transferência de massa, momento e calor, não compreendidas nas escalas industriais convencionais. Como grande atrativo desta área, verifica-se a utilização de microrreatores para processos reacionais, nos quais são obtidos rendimentos elevados.

Fazendo parte do corpo de estudo, as sínteses orgânicas por demandarem elevados valores de purezas do produto final e possuem rotas reacionais complexas, vêm sendo largamente aplicadas juntamente à microfluídica. Dentre as reações mais utilizadas, toma-se como exemplo a reação de transesterificação para a produção de biodiesel, partindo-se de etanol e óleo vegetal. Em microrreatores, a síntese de biodiesel apresentou rendimentos maiores que 95% (ARIAS *et al.*, 2012;

SANTACESARIA *et al.*, 2012; AGHEL *et al.*, 2014; BHOI *et al.*, 2014; SANTANA; SILVA, J. L. *et al.*, 2018).

Por mais promissora que a área se apresenta, existem ainda limitações acerca de dados de processos que utilizem microdispositivos, devido a seu uso crescente ser ainda recente.

Dentre as alternativas vigentes à obtenção de resultados e comportamentos de processos, verifica-se a possibilidade de uso de simulações computacionais, de forma que, seguindo a descrição fenomenológica e cinética correspondente à reação estudada, pode-se obter uma estimativa de um comportamento experimental. Porém, mesmo o uso das simulações demonstra limitações decorridas da dificuldade de se analisar reações com cinéticas complexas ou descrever corretamente a fenomenologia.

Comumente à microfluídica, mas em devidas proporções, as redes neurais artificiais, dentro do escopo da inteligência artificial (IA), representam uma solução alternativa para a problemática apresentada. Dentro da área chamada como *machine learning* (ML), a

IA possibilita o uso de modelos de regressão para descrever comportamentos de uma série de dados. Ao que tange a junção das duas áreas - microfluídica e IA -, verifica-se inúmeros estudos em acervos científicos que apresentam metodologias que utilizam ML para predição de rendimentos, controle e outros fatores em processos reacionais (GALAN *et al.*, 2020; MOAYEDI *et al.*, 2020; PIRDASHTI *et al.*, 2013).

Compilando as ideias centrais exemplificadas e baseando-se em aplicações reais dos temas abordados, o cerne da proposta do referido trabalho visa estabelecer, e analisar qualitativamente e quantitativamente o uso de redes neurais artificiais e o modelo de regressão - em específico, o modelo *multilayer perceptron* -, para a predição de resultados de síntese reacional através de microrreatores. A formulação da pesquisa foi discretizada em duas etapas: foi adicionado ao *software* novas funcionalidades como a automatização de uso de algoritmos de aprendizado de máquina (utilizando o banco de dados presente no *software* citado) e a validação dos resultados obtidos<sup>1</sup>.

## METODOLOGIA:

O modelo *multilayer perceptron* (MLP) está contido dentro da classe de aprendizado supervisionado, no qual representa uma aproximação de uma função não linear cabível de ser utilizado em casos de regressão. Detalhando-se a topologia de uma rede neural do

tipo MLP, considera-se as Figuras 1 e 2, onde se caracteriza um único neurônio e uma rede neural artificial, respectivamente.

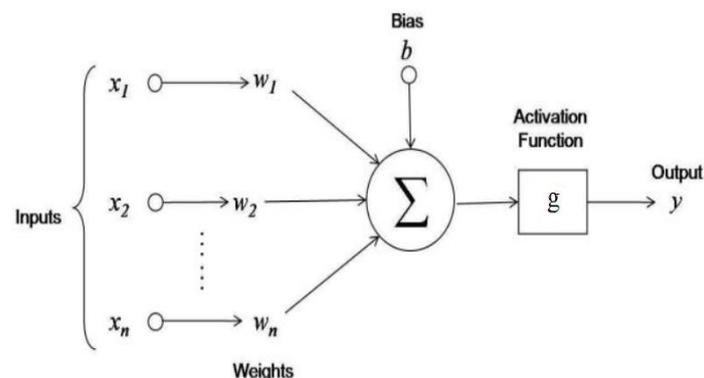


Figura 1 - Estrutura de um neurônio artificial (SUZUKI, 2017, p.147) Adaptado pelo autor.

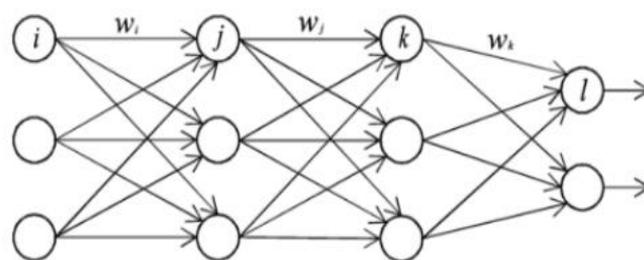


Figura 2 – Topologia da rede neural artificial (MENG *et al.*, 2020) Adaptado pelo autor.

Na Figura 1, os parâmetros ajustáveis (treináveis) do neurônio são expostos como:  $w_i$  (pesos) e  $b$  (bias). Em conjunto com os parâmetros de entrada ( $x$ ), *inputs*, e a função de ativação ( $g$ ), pode-se definir a função que descreve o valor de  $y$ , output, assim como a exposição na Equação 1.

$$y = g(x^T w + b) \quad (1)$$

A conexão de neurônios artificiais, conforme exposta na Figura 2, torna o valor do *output* dependente de um número maior de parâmetros treináveis (pesos e bias de cada neurônio presente na rede).

<sup>1</sup> O software em questão foi desenvolvido no programa de iniciação científica PIBIC 2019-2020 (Desenvolvimento de um banco de dados com informações de processos de mistura e reação

química em microdispositivos), inicialmente com a funcionalidade de banco de dados. A pesquisa em questão (2020-2021) adicionou novas funções ao programa.

O treinamento de uma rede neural ocorre na fase conhecida como “fase de treinamento”. Nesta, dados de entrada ( $x$ ) e saída ( $y$ ) são fornecidos e através de funções de otimização dos parâmetros ajustáveis, busca-se obter uma função que seja capaz de descrever o comportamento dos dados dentro dos limites inferior e superior inferidos. A aplicação e automação da aplicação (junto ao

software) do modelo MLP foi efetuado através da linguagem computacional da linguagem computacional *python* (versão 3.8). Em seu desenvolvimento, segregou-se em duas etapas distintas: [1] carregamento e tratamento dos dados (Figura 3), e [2] exposição dos resultados da rede MLP após treinamento (Figura 4).

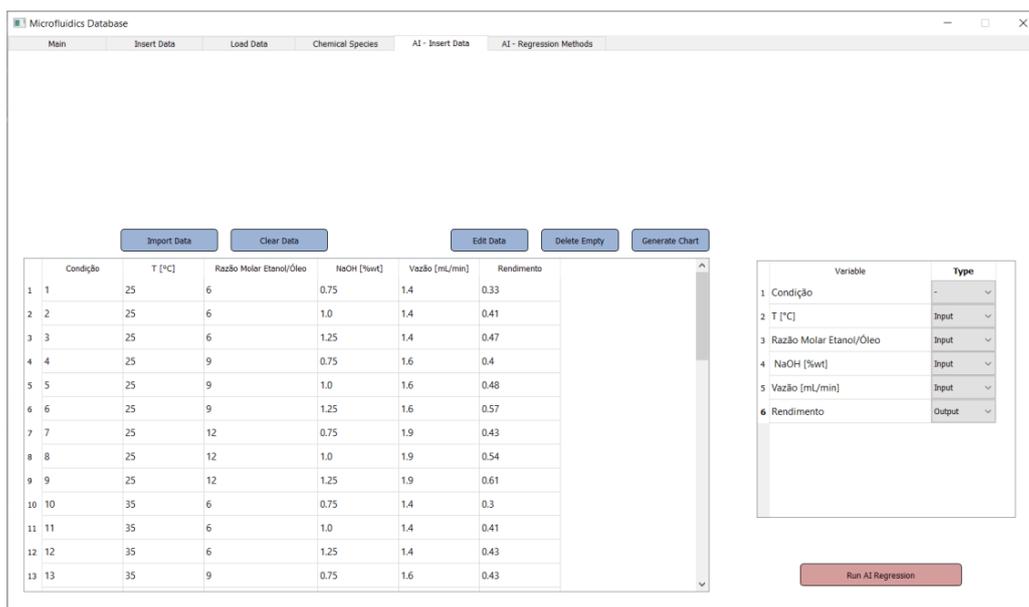


Figura 3 – Janela de carregamento e tratamento de dados do software desenvolvido.

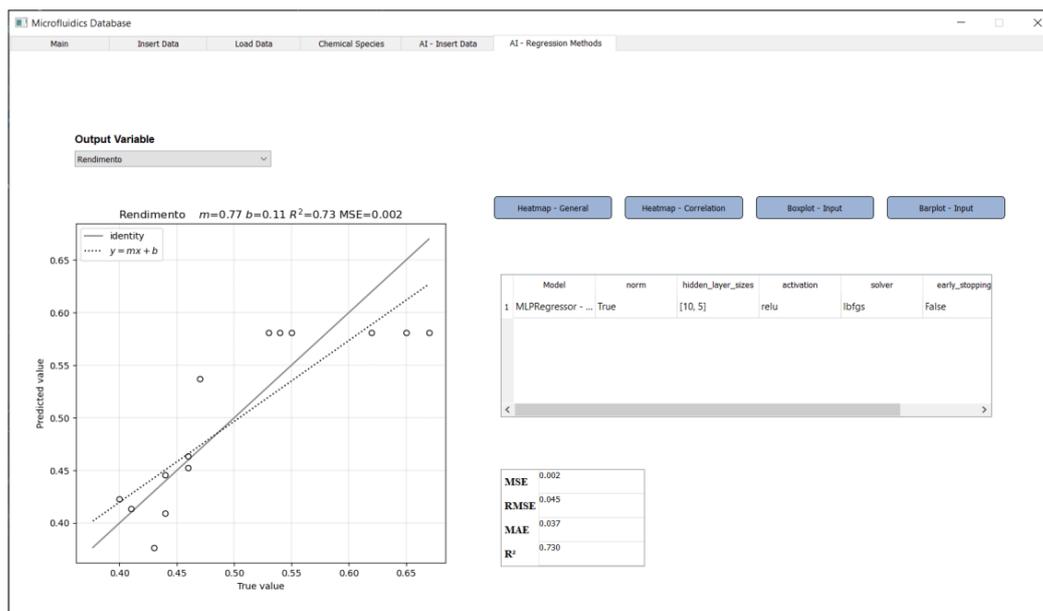


Figura 4 – Janela de exposição dos resultados obtidos pela MLP através do software desenvolvido.

[1]: Na janela exposta na Figura 3, as seguintes funcionalidades estão presentes: tabela de inserção de dados (manual ou exportação); exclusão de linhas com dados faltantes; gráfico de plotagem de dispersão entre duas variáveis; tabela para aferição de *inputs* e *outputs* para o modelo de *machine learning*, e execução do modelo ML.

[2]: Na janela exposta na Figura 4, as seguintes funcionalidades estão presentes: para casos com mais de um *output* na rede neural, há a opção de se escolher qual variável terá os resultados expostos; gráfico para comparação dos valores obtidos *versus* esperados, gráfico *heatmap* com ou sem

correlação de Pearson, exposição da topologia da rede neural, e valores dos erros estatísticos MSE, RMSE, MAE e R<sup>2</sup>.

## RESULTADOS E DISCUSSÃO:

Utilizando os dados presentes no banco de dados (Zhai, *et al*, 2010; Moayedi *et al*, 2020; Dai *et al*, 2019; Deng *et al*, 2019), modelos MLP foram testados a fim de se validar as novas funções do *software*. A Tabela 1 expõe os melhores resultados encontrados para cada conjunto de dados.

Referência	Processo	Topologia	Função de Ativação	Otimizador	R <sup>2</sup>	RMSE
Zhai <i>et al</i> , 2010	Reforma à vapor de metano	[10, 20]	ReLU	lbfgs	0,92	4,32E-02
Moayedi <i>et al</i> , 2020	Produção de biodiesel	[17]	ReLU	lbfgs	0,94	2,85E-02
Dai <i>et al</i> , 2019	Produção de fenol	[11]	ReLU	lbfgs	0,91	4,55E-02
Deng <i>et al</i> , 2019	Reforma à vapor de DME	[15]	ReLU	lbfgs	0,92	4,28E-02

Tabela 1: Topologia das redes neurais MLP e erros estatísticos obtidos para diferente quatro conjunto de dados presentes no banco de dados referido.

A descrição da topologia caracteriza a quantidade de camadas ocultas e o número de neurônios em cada camada da rede MLP. Ao caso do processo “Produção de biodiesel”, a rede com maior capacidade preditiva apresentou apenas uma camada oculta e nesta, 17 neurônios. Dentre os resultados

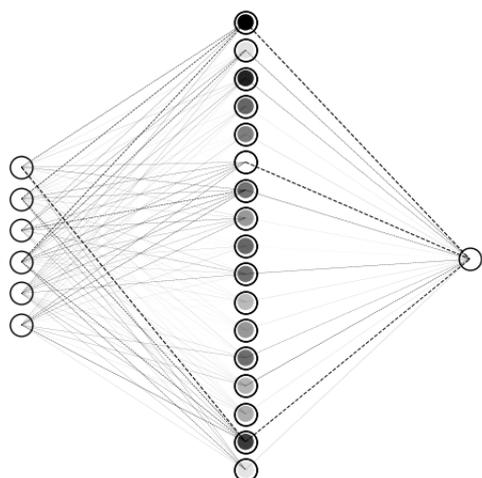


Figura 5 – Topologia da rede neural com 5 variáveis de entrada, 1 variável de saída e 17 neurônios na única camada oculta presente.

obtidos, expõem-se a topologia da rede neural descrita acima (Figura 5) e o gráfico de comparação entre o valor da variável rendimento (analisada no caso) predita *versus* esperada (Figura 6).

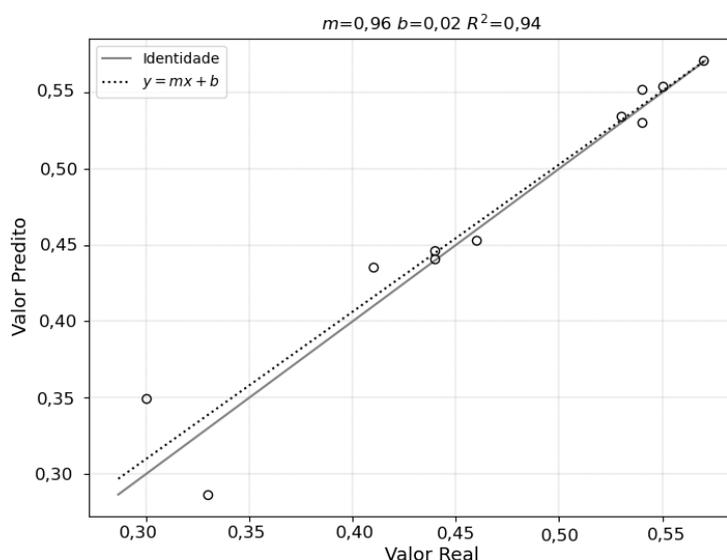


Figura 5 – Gráfico de comparação entre os valores preditos e reais para a variável rendimento (Moayedi *et al*, 2020)

## CONCLUSÕES:

A formulação das ferramentas de automação de uso de modelos *multilayer perceptron* (MLP) ao caso do uso do engenheiro químico, permite a utilização das ferramentas computacionais sem a necessidade de prévios conhecimentos de programação. Em quesitos de melhoria, atualmente, o *software* encontra-se em constante desenvolvimento buscando aprimorá-lo além das opções descritas e alimentando-o com funcionalidades que fogem do olhar da formação de engenharia por si só, como é o caso de preenchimento de dados faltantes e utilização de outros métodos de regressão.

Em relação aos resultados oriundos da aplicação do *software* em conjunto com os dados, verificou-se questões sensíveis às qualidades dos dados. Além dos resultados expostos, outros casos não puderam ser testados por conta da má qualidade dos dados proveniente do banco de dados. Porém, após algumas consultas no mesmo, encontrou-se casos como estes utilizados.

## BIBLIOGRAFIA

- ARIAS, E. L. M. *et al.* **Continuous synthesis and in situ monitoring of biodiesel production in different microfluidic devices.** Industrial and Engineering Chemistry Research, v. 51, n. 33, p. 10755–10767, 2012
- SANTACESARIA, E. *et al.* **Biodiesel process intensification in a very simple microchannel device.** Chemical Engineering and Processing: Process Intensification, Elsevier B.V., v. 52, p. 47–54, 2012
- AGHEL, B. *et al.* **Using a wire coil insert for biodiesel production enhancement in a microreactor.** Energy Conversion and Management, Elsevier Ltd, v. 84, p. 541–549, 2014
- BHOI, R. *et al.* **Transesterification of sunflower oil in microreactors.** International Journal of Chemical Reactor Engineering, v. 12, n. 1, p. 47–62, 2014
- SANTANA, H. S.; SILVA, J. L.; TORTOLA, D. S.; TARRANTO, O. P. **Transesterification of sunflower oil in microchannels with circular obstructions.** Chinese Journal of Chemical Engineering, v. 26, n. 4, p. 852–863, 2018
- GALAN, E. A. *et al.* **Intelligent Microfluidics: The Convergence of Machine Learning and Microfluidics in Materials Science and Biomedicine.** Matter, Elsevier Inc., v. 3, n. 6, p. 1893–1922, 2020
- MOAYEDI, H.; AGHEL, B.; FOONG, L. K.; BUI, D. T. **Feature validity during machine learning paradigms for predicting biodiesel purity.** Fuel, Elsevier, v. 262, November 2019, p. 116498, 2020
- PIRDASHTI, M. *et al.* **Artificial neural networks: applications in chemical engineering.** Rev Chem Eng, v. 29, p. 205–239, 2013
- SUZUKI, K. **Artificial neural networks- architectures an applications.** p.265, 2017.
- MENG, Z.; HU, Y.; ANCEY, C. **Using a data driven approach to predict waves generated by gravity driven mass flows.** Water (Switzerland), v. 12, n. 2, 2020
- Dai, X., Zhou, W., Yang, S., Sun, F., Qian, J., He, M., & Chen, Q. **Microchannel process for phenol production via the cleavage of cumene hydroperoxide.** Chemical Engineering Science, 199, 398–404, 2019
- Moayedi, H., Aghel, B., Foong, L. K., & Bui, D. T. **Feature validity during machine learning paradigms for predicting biodiesel purity.** Fuel, 262, 2020
- Zhai, X., Ding, S., Chang, Y., Jin, Y., & Cheng, Y. **CFD simulation with detailed chemistry of steam reforming of methane for hydrogen production in an integrated micro-reactor.** International Journal of Hydrogen Energy, 35(11), 2010
- Deng, X., Yang, T., Zhang, Q., Chu, Y., Luo, J., Zhang, L., & Li, P. **A monolith CuNiFe/Γ-Al<sub>2</sub>O<sub>3</sub>/Al catalyst for steam reforming of dimethyl ether and applied in a microreactor.** International Journal of Hydrogen Energy, 44(5), 2019