

IDENTIFICAÇÃO DE COLUNA DE DESTILAÇÃO EXPERIMENTAL POR REDES NEURAIS ARTIFICIAIS

Murilo Silva Colombo¹, Thiago Tonon², Matheus Felipe Gremes³ Cid Marcos Gonçalves Andrade⁴, Oswaldo Curty da Motta Lima⁵

¹Aluno de Doutorado, Mestre, Programa de Pós-Graduação em Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá - UEM. pg401365@uem.br

²Professor, Doutor, Campus Ivaiporã/PR. Instituto Federal do Paraná – IFPR. thiago.tonon@ifpr.edu.br

³Aluno de Mestrado, Programa de Pós-Graduação em Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá – UEM. pg402992@uem.br

⁴Coorientador, Doutor, Departamento de Engenharia Química, Pesquisador do Programa de Pós-Graduação em Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá. cmgandrade@uem.br

⁵Orientador, Doutor, Departamento de Engenharia Química, Pesquisador do Programa de Pós-Graduação em Engenharia Química, Campus Maringá/PR, Universidade Estadual de Maringá. ocmlima@uem.br

RESUMO

Este artigo apresenta um esforço de modelar uma coluna de destilação experimental com modelos empíricos baseados em aprendizado de máquina, especificamente em Redes Neurais Artificiais (RNA). O problema é proposto como um desenvolvimento de um modelo de maneira mais rápida e robusta do que um modelo fenomenológico. Os dados são obtidos por um módulo de aquisição de dados conectado à coluna de destilação. Existem dois conjuntos de dados, de dois experimentos diferentes, que são então preprocessados para torná-los adequados para aprendizado e predição pelas redes neurais. Após a preparação dos dados, os modelos de RNA foram otimizados em relação aos dados. Os modelos obtiveram desempenho suficiente, embora os conjuntos de dados fossem problemáticos e com poucas amostras. Também foi relatado que os modelos mais simples foram os que apresentam melhor desempenho.

PALAVRAS-CHAVE: Aprendizado de máquinas; Redes neurais artificiais; Identificação de sistemas; Colunas de destilação.

1 INTRODUÇÃO

Destilação é um processo físico (operação unitária, na nomenclatura de engenharia química) de apreciável adoção na indústria para separação de misturas de líquidos com diferentes pontos de ebulição. Estima-se que 40 % de toda a energia consumida por uma planta química seja utilizada na seção de destilaria. Estima-se, ainda, que essas colunas operem com eficiência termodinâmica de 5 a 20 % (JANA, 2014). Apesar de serem estáveis na maioria dos casos, implantar um sistema de controle pode levar a uma maior pureza dos produtos e diminuição do consumo energético. Em diversas ocasiões, dependendo da lógica de controle que se deseja implantar, é necessário a construção de um modelo que realize previsões do perfil dinâmico de colunas de destilação.

Na construção de um modelo de coluna(s) de destilação, desenvolver um modelo físico, fenomenológico, é uma tarefa não trivial. São envolvidas diversas equações, de estágio a estágio, de características diferencial, acopladas e, algumas delas, não lineares. Além disso, é necessário mais de uma equação por estágio, representando um balanço de massa, de energia e um modelo termodinâmico. O resultado é um modelo de alta dimensão, que precisa ser sintonizado e, preferencialmente, seja validado com dados experimentais (SKOGESTAD, 1997).

No entanto, em plantas que se desejam usar este tipo de modelo, é comum a(s) coluna(s) já estarem instrumentadas, isto é, com sensores e atuadores instalados. Nesta perspectiva, é proposto como alternativa o uso de modelos de identificação, ou seja, modelos paramétricos construídos a partir de uma estrutura predefinida e que são otimizados em relação aos dados experimentais coletados pelo sistema de sensores. A vantagem desta metodologia é o consumo menor de tempo e recursos para se chegar a um modelo que alcance os critérios de desempenho preestabelecidos.

O desenvolvimento de modelos a partir de redes neurais para colunas de destilação é prática na engenharia química desde a década de 1980, em especial na década de 1990, quando os principais trabalhos aplicados - que utilizavam dados geralmente obtidos por modelos computacionais - com estruturas de redes neurais ainda simples mas que alcançavam resultados promissores (MACMURRAY e HIMMELBLAU, 1995).

Na anos de 2000 a 2010, destacam-se trabalhos que utilizaram redes neurais para inferir a composição dos produtos purificados após processo de destilação a partir de dados produzidos por modelo computacional. A ideia era que modelos de redes neurais, alimentadas com dados de operação como temperatura dos estágios, poderiam ser utilizadas para estimar a composição dos produtos – já que obter essas concentrações é mais difícil, a partir de sensores que estimam concentração de forma online (SINGH, GUPTA e GUPTA, 2005). Posteriormente, em outro trabalho, são apresentados resultados de redes com mais variáveis de entrada, incluindo agora dados de pressão, vazão de refluxo e taxa de aquecimento no revededor, além do perfil de temperaturas da coluna. Ademais, foi utilizado um otimizador (algoritmo de Levenberg-Marquardt) mais eficiente em relação ao trabalho anterior, para treinar as redes neurais (SINGH, GUPTA e GUPTA, 2007). Os autores desenvolvem tais modelos e demonstram que, após treinamento, com os recursos computacionais da época, as redes gastavam pouco tempo para computar a previsão de composição dos produtos. Isso indicava que a aplicação desses modelos em controle era oportuna.

Em artigos de revisão, são documentados os principais trabalhos de modelos por redes neurais, para colunas de destilação ou na engenharia química em geral, incluindo ou não exemplos e até mesmo descrevendo, caso tenha sido realizada, aplicação do modelo desenvolvido para algum propósito (HUSSAIN 1999; HIMMELBLAY 2008; ALI *et al.* 2015)

Destacam-se trabalhos mais recentes, de escopo comparável com os trabalhos de Singh *et al.* 2005 e 2007, em que se usa o termo sensores virtuais, cujos dados foram coletados de uma coluna experimental (DE CANETE *et al.*, 2012). Há, ainda, alguns trabalhos que utilizam os modelos a alguma finalidade, principalmente controle. Por exemplo, quando é desenvolvido modelo adaptativo e aplicação em controle inferencial de uma coluna de destilação multicomponente(RANI, SINGH e GUPTA, 2013). Com efeito e configurações semelhantes, realizou-se controle inferencial a partir de rede neural recorrente para modelar, desta vez, uma coluna de destilação reativa (KATARIA; SINGH, 2018).

O objetivo deste trabalho foi desenvolver um modelo de identificação de uma coluna de destilação experimental, com base em dados coletados por um sistema de aquisição de dados, em que se usa não todo o perfil de temperatura da coluna inteira, mas uma pequena porção de dados de temperatura de alguns pratos selecionados previamente.

2 MATERIAIS E MÉTODOS

Os dados para identificação de sistemas a partir de redes neurais tem como origem uma coluna de destilação experimental separadora de etanol e água. Essa coluna possui 14 pratos (estágios), um revededor elétrico e um condensador a base de água encanada. Quando em operação em estado estacionário, etanol é coletado no topo com concentração de aproximadamente 90° GL sendo a alimentação realizada em concentração em torno de 30°GL. A figura 1 a seguir representa um diagrama de informação da coluna de destilação.

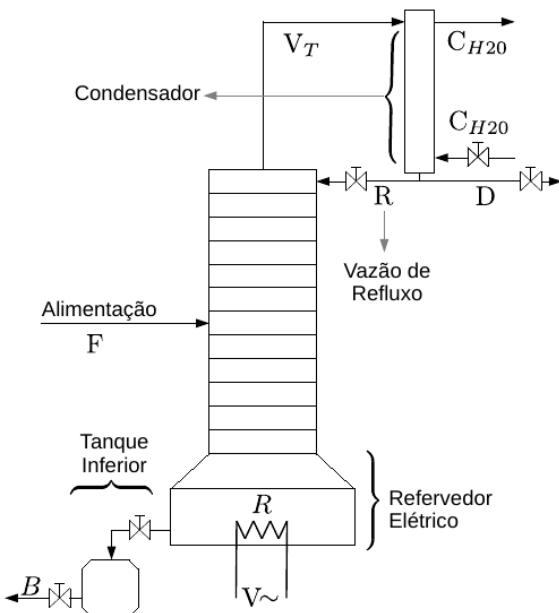


Figura 1: Desenho esquemático da coluna de destilação

Fonte: Autores

A coluna de destilação possui apenas sensores de temperatura, já que sensores de vazão seriam de difícil adaptação na coluna e sensores de pressão são desnecessários já que a coluna opera em pressão atmosférica. A alimentação de reagentes na coluna é manual e, nos experimentos realizados, não havia um sistema de controle atuando sobre a coluna. A operação estável (ou instável, como será descrito) foi obtida manipulando manualmente a corrente elétrica do referedor.



Figura 2: Foto da coluna de destilação
Fonte: Autores

a. SOBRE OS COJUNTOS DE DADOS

Os dados de treinamento correspondem a um experimento realizado em 2017, com duração de 2 horas e meia, coletando dados em um intervalo de 10 segundos, resultando em séries temporais de dimensão 952. Já os dados de teste, compreenderam experimento realizado em 2020, com duração de 1 hora e vinte minutos, intervalo amostral de 3 segundos e formando séries temporais de dimensão 1921. Os gráficos 1 e 2 representam os dados de treinamento e teste respectivamente.

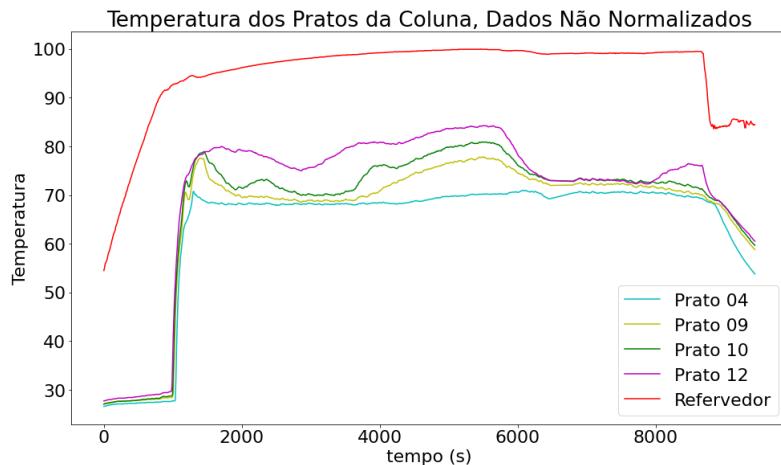


Gráfico 1: Conjunto de dados de treinamento

Fonte: Dados coletados da pesquisa

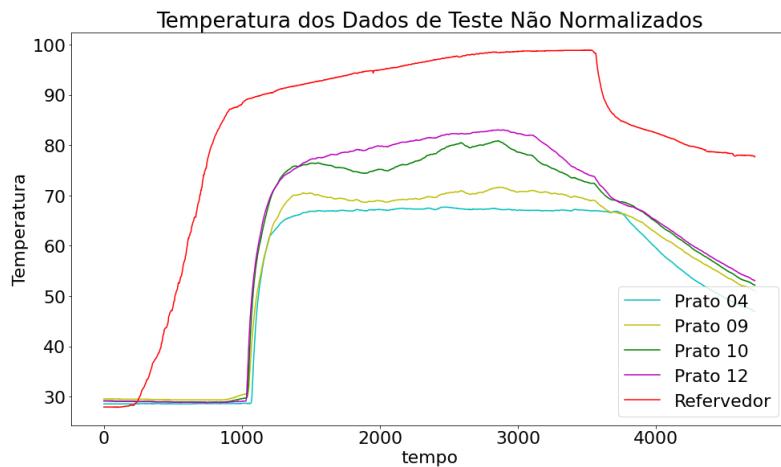


Gráfico 2: Conjunto de dados de validação

Fonte: Dados coletados da pesquisa

O experimento que gerou os dados de teste foi atípico, pois não foi possível obter um estado estacionário rigorosamente, no sentido de se estabilizar as vazões e obter produtos purificados, com vazão aproximadamente constante, nas saídas da coluna. Além disso, alguns pratos não formaram uma camada de líquido, o que descaracteriza a operação da coluna e o que, efetivamente, diminui o número de estágios da coluna disponíveis para separação. Pode-se observar, pelos gráficos 1 e 2, grandes diferenças nos regimes de operação dos experimentos, principalmente na dificuldade de se fazer o referedor atingir o ponto de ebulação da água quando se diminuía a carga térmica para forçar os pratos mais baixos da coluna a acumularem líquido.

Com uma fundamentação mais rigorosa, a escolha das variáveis de saída seriam a temperatura dos pratos 12 ou 4 (SKOGESTAD, 2007). No entanto, o prato 13 não formou líquido, e, portanto, não haveria um prato entre o 12 e o referedor para modelar a coluna. Por conta desses fatos, é justificada a escolha dos pratos 4, 9, 12 e o referedor como variáveis de entrada para um modelo com a temperatura do prato 10 como saída.

b. PREPROCESSAMENTO DE DADOS

Foi realizado neste trabalho preprocessamento de dados, com o objetivo de extrair, normalizar e agrupar variáveis de interesse, já que isso é necessário para alimentar as redes neurais. Foi realizada a aquisição dos dados de documento de texto a partir do módulo *pandas* do *python* e em seguida os dados foram transformados para *arrays* do *numpy*, por último retirando caracteres alfabéticos. Em seguida, os dados foram

normalizados por meio da função *minmax*, e agrupados os dados de entrada em uma array de 4 variáveis

c. METODOLOGIA REFERENTE A IDENTIFICAÇÃO DE SISTEMAS

Após a aquisição e análise dos dados, a metodologia de identificação de sistemas requer: a escolha de estruturas do modelo que será ajustado aos dados experimentais, o número de parâmetros desse modelo, a escolha do otimizador que irá realizar o ajuste perante os dados e, por último, a validação do modelo perante um novo conjunto de dados. Destaca-se que são testadas diversas estruturas de modelos, até que o melhor modelo perante os dados de validação é selecionado e, seguidamente, é utilizado conforme necessidade do usuário.

Isso vem ao encontro da teoria existente para se construir uma rede neural artificial, descritas na próxima subseção. Além disso, na intersecção dessas ciências, há um conflito com o que se quer dizer por validação. Em aprendizado de máquinas, validação compreende uma fração dos dados de treinamento que são usados para avaliar a rede durante o treinamento. Já em identificação de sistemas, validação significa etapa de verificação se o modelo consegue generalizar o objeto em outros conjuntos de dados, ou intervalos de operação. Essa etapa, em aprendizagem de máquina, compreende avaliação da rede em um conjunto de dados de testes.

d. REDES NEURAIS ARTIFICIAIS

Neste trabalho, seguiu-se as recomendações da ciência de redes neurais artificiais, uma ciência experimental, em que a construção dos modelos tem uma bases teóricas, mas que os hiperparâmetros - número de camadas ocultas, números de neurônios em cada camada, critérios de aprendizagem - são ajustados pelo usuário para atender os critérios do uso futuro das redes (HAYKIN, 1999).

As topologias de redes neurais utilizadas neste trabalho foram a rede neural alimentada à frente (FFNN, do inglês *Feed-Forward Neural Network*) e uma rede neural recorrente (RNN, *Recurrent Neural Network*). A escolha do número de camadas segue a observação que a coluna de destilação, com a configuração de dados de entrada selecionados, em um trabalho anterior, demandavam 2 camadas ocultas (COLOMBO, 2020). Em seguida, foram testados, numa configuração padrão, o melhor otimizador e melhor função de ativação para cada rede. A partir daí, o número de neurônios foi ajustado de forma a obter os melhores resultado possíveis dentro daquela configuração determinada anteriormente. Minilote e número de épocas foram manipulados interativamente conforme se observava o curso do treinamento das redes.

3 RESULTADOS E DISCUSSÕES

Nesta seção são descritos e sumarizados os resultados obtidos conforme métodos detalhados anteriormente. São apresentadas figuras exclusivamente do melhor modelo, e que demonstram resultados excelentes, mesmo com as dificuldades impostas e discutidas na subseção de discussão dos resultados.

a. RESULTADOS

Os parâmetros de desempenho das redes neurais em relação ao conjunto de dados de teste são apresentados na tabela 1 a seguir, com sua topologia também mencionada. A rede neural FNN foi otimizada com *Adagrad* disponível na biblioteca *Keras*, conforme descrito no trabalho (DUCHI, HAZAN e SINGER, 2011). Já a rede RNN foi treinada com o

otimizador *Adam* (KINGMA e BA, 2014). Na maioria das aplicações, os parâmetros dos otimizadores foram sintonizados para cada estrutura de rede utilizada.

Tabela 1: Estrutura das redes e parâmetros de desempenho

Rede Neural	Número de Neurônios	MSE (teste)	MAE (teste)
FFNN	24-12-1	$4,30 \cdot 10^{-4}$	0,0160
RNN	20-12-1	$8,71 \cdot 10^{-4}$	0,0222

Fonte: Dados da pesquisa

A seguir, são reproduzidas, nos gráficos 4, 5 e 6, resultados do melhor modelo (FNN), em específico as figuras do processo de aprendizagem, da predição do modelo, e uma imagem com enfoque no perfil estacionário (apesar de não ser estacionário rigorosamente). Esses resultados demonstram o quanto o modelo erra em relação aos dados (o erro absoluto nas chega, no pior dos casos, a aproximadamente 2,0 °C).

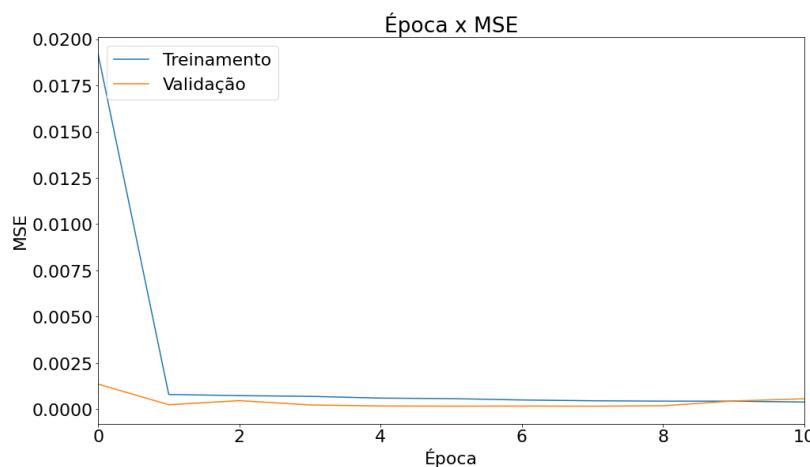


Gráfico 4: Evolução do treinamento da rede FFNN

Fonte: Dados coletados da pesquisa

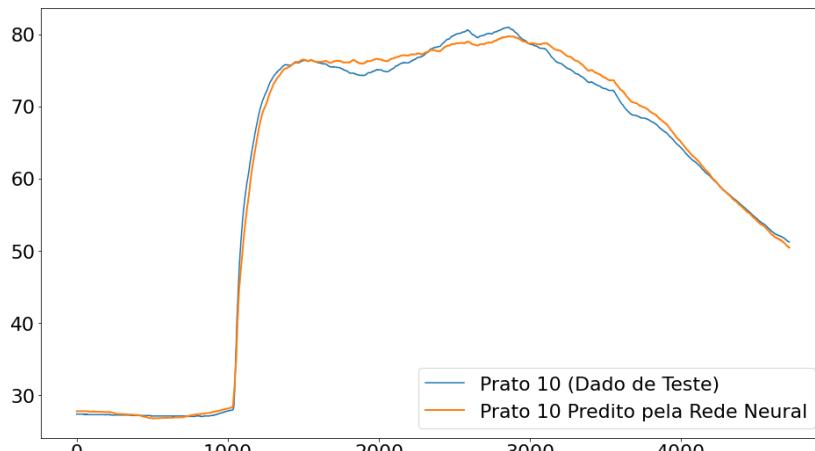


Gráfico 5: Predição da rede neural comparado aos dados

Fonte: Dados coletados da pesquisa

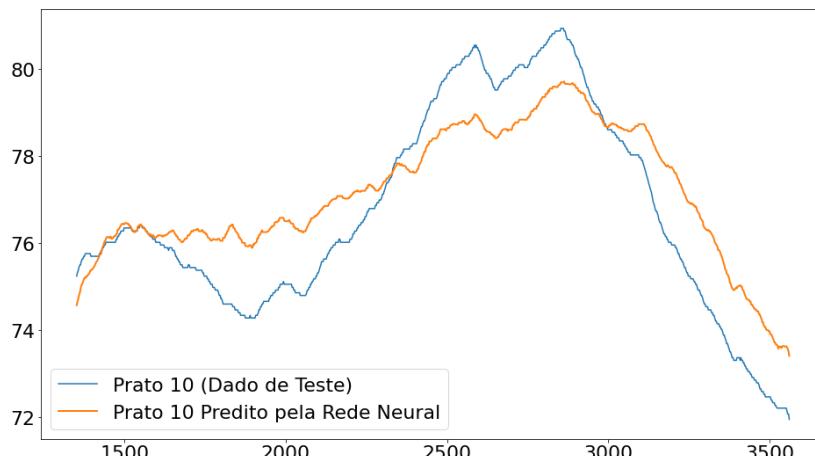


Gráfico 6: Previsão da Rede Neural com enfoque
Fonte: Dados coletados da pesquisa

b. DISCUSSÃO DE RESULTADOS

Não era esperado, do ponto de vista da teoria de redes neurais artificiais, que a rede de estrutura mais simples (FFNN) modelasse melhor a coluna de destilação, já que a rede RNN possui um mecanismo de realimentação e, portanto, deveria ser mais apropriada para os dados de natureza dinâmica. Uma justificativa pode ser a natureza distinta dos conjuntos de dados, com tempos de amostragem distintos. O conjunto de dados de treinamento foi obtido por um experimento válido do ponto de vista de operação da coluna, já que foi obtido um estado estacionário com líquido nos pratos e vazão de saída dos produtos aproximadamente constante e pelo fato de se saber que a coluna estava sendo alimentada pela mistura de etanol e água. Além disso, nos dados de treinamento foram realizadas pequenas perturbações na coluna, e que foram, em seguida, estabilizadas. No conjunto de teste não foi atingido estado estacionário, as vazões não foram estabilizadas e não se tinha a certeza de que a coluna estava sendo alimentada. Em suma, essa discrepância entre os dados, tornando-os impróprios para identificação de sistemas, pode ter levado os modelos mais complexos se adaptarem aos dados de treinamento, porém sem conseguir generalizar os dados de teste.

Poderia ser utilizado como argumento que os dados deveriam passar por melhor preprocessamento dos dados, eliminando os regimes transientes de inicialização e desligamento da coluna. Isso foi levado em consideração, porém, as redes que utilizaram os dados preprocessados dessa forma apresentaram resultados inferiores (com $MAE > 3,5\%$), com uma fuga acentuada em relação aos dados experimentais, quando se observavam os gráficos.

Os resultados obtidos neste trabalho foram melhores em termos de critérios dos dados de teste ($MAE < 2,0\%$) em relação ao melhor resultado de uma rede externamente recorrente (ERN), de trabalho de mestrado, em que foi obtido $MAE = 2,84\%$ em relação aos dados de teste (COLOMBO, 2020). Isso pode ser justificado pelo uso do algoritmo Levenberg-Marquardt, no trabalho anterior a este, que requer a computação da matriz Hessiana e não é eficiente em problemas com maior número de parâmetros, fato que obrigou as redes possuírem um menor número de neurônios. Outra justificativa pode ser a não eliminação, nos resultados aqui apresentados, de coleta dos dados com a coluna desligada. Cabe ressaltar, ademais, que a rede denominada MHLR naquele trabalho, com estrutura semelhante à rede FFNN, apresentou resultados inferiores ($MAE = 3,78\%$) em relação ao que foi alcançado com a rede FFNN e o otimizador *adagrad* ($MAE = 1,60\%$). O que pode também ser justificado pelos argumentos apresentados anteriormente do tipo de otimizador e pelo número de neurônios nas camadas.

4 CONCLUSÕES E PERSPECTIVAS FUTURAS

O objetivo deste trabalho foi construir modelos de identificação com estruturas de redes neurais para modelar uma coluna de destilação. Foram obtidos modelos com relativo bom desempenho, perante as dificuldades já mencionadas provenientes dos conjuntos de dados. De característica contraintuitiva, o modelo mais simples modelou melhor os dados, perante os dois tipos testados, redes FFNN e RNN. Em suma, é evidente que estruturas de redes neurais artificiais são pertinentes quando aplicadas em identificação de sistemas, como pode-se inferir, também, de outros trabalhos publicados.

Como sugestões para trabalhos futuros, obter dados mais representativos, com maior frequência de amostragem, poderia melhorar a qualidade dos modelos. Além disso, explorar outras estruturas de redes neurais promissoras, como outras estruturas de redes recorrentes, e como outras topologias de redes, como *Radial Basis Functions* ou *Support Vector Machines*. Seguidamente, seria desejável aplicar os modelos obtidos para alguma finalidade, como simulação, controle ou otimização

REFERÊNCIAS

ALI, J. M. *et al.* Artificial Intelligence techniques applied as estimator in chemical process systems - A literature survey. **Expert systems with Applications**, v. 42, n. 14, p. 5915-5931, 2015.

COLOMBO, M. S. **Identificação de coluna de destilação por redes neurais artificiais.** [s.l.], Universidade Estadual de Maringá, 2020.

DE CANETE, J. F. *et al.* Dual composition control and soft estimation for a pilot distillation column using a neurogenetic design. **Computers & Chemical Engineering**, v. 40, p. 157–170, 2012.

DUCHI, J.; HAZAN, E.; SINGER, Y. Adaptive subgradient methods for online learning and stochastic optimization. **Journal of Machine Learning Research**, v. 12, n. 7, 2011.

HAYKIN, S. **Neural networks:** a comprehensive foundation by Simon Haykin, Macmillan, 1994, ISBN 0-02-352781-7. The Knowledge Engineering Review, 1999.

HIMMELBLAU, D. M. Accounts of experiences in the application of artificial neural networks in chemical engineering. **Industrial & Engineering Chemistry Research**, v. 47, n. 16, p. 5782-5796, 2008.

HUSSAIN, M. A. Review of the applications of neural networks in chemical process control - simulation and online implementation. **Artificial Intelligence in Engineering**, v. 13, n. 1, p. 55–68, 1999.

JANA, A. K. Advances in heat pump assisted distillation column: a review. **Energy Conversion and Management**, v. 77, p. 287-297, 2014.

KATARIA, G.; SINGH, K. Recurrent neural network based soft sensor for monitoring and controlling a reactive distillation column. **Chemical Product and Process Modeling**, v. 13, n. 3, 2018.

KINGMA, D. P.; BA, J. **Adam**: a method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

MACMURRAY, J. C.; HIMMELBLAU, D. M. Modeling and control of a packed distillation column using artificial neural networks. **Computers & Chemical Engineering**, v. 19, n. 10, p. 1077-1088, 1995.

RANI, A.; SINGH, V.; GUPTA, J. R. P. Development of soft sensor for neural network based control of distillation column. **ISA transactions**, v. 52, n. 3, p. 438-449, 2013.

SINGH, V.; GUPTA, I.; GUPTA, H. O. ANN based estimator for distillation-inferential control. **Chemical Engineering and Processing: Process Intensification**, v. 44, n. 7, p. 785-795, 2005.

SINGH, V.; GUPTA, I.; GUPTA, H. O. ANN-based estimator for distillation using Levenberg--Marquardt approach. **Engineering Applications of Artificial Intelligence**, v. 20, n. 2, p. 249–259, 2007.

SKOGESTAD, S. Dynamics and control of distillation columns: A tutorial introduction. **Chemical Engineering Research and Design**, v. 75, n. 6, p. 539-562, 1997.

SKOGESTAD, S.; SKOGESTAD, S. The dos and don'ts of distillation column control. **TRANS ICHEME PART A**, p. 2007, [s. d.].