



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
CURSO DE BACHARELADO EM ENGENHARIA QUÍMICA

JOSÉ LEONARDO GOMES DE ARAÚJO

**MODELAGEM DE REFORMA CATALÍTICA SECA DE METANO A GÁS DE  
SÍNTESE, UTILIZANDO MACHINE LEARNING E REDES NEURAIS**

Recife  
2022

**JOSÉ LEONARDO GOMES DE ARAÚJO**

**MODELAGEM DE REFORMA CATALÍTICA SECA DE METANO A GÁS DE  
SÍNTESE, UTILIZANDO MACHINE LEARNING E REDES NEURAIS**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Graduação em Engenharia Química da Universidade Federal de Pernambuco, como requisito parcial à obtenção do grau de Bacharel em Engenharia Química.

Orientador: Prof. Dr. Sérgio Lucena

Recife  
2022

Ficha de identificação da obra elaborada pelo autor,  
através do programa de geração automática do SIB/UFPE

Araújo, José Leonardo Gomes de.

Modelagem de reforma catalítica seca de metano a gás de síntese, utilizando  
machine learning e redes neurais / José Leonardo Gomes de Araújo. - Recife,  
2022.

49 p. : il., tab.

Orientador(a): Sérgio Lucena

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de  
Pernambuco, Centro de Tecnologia e Geociências, Engenharia Química -  
Bacharelado, 2022.

1. Gás de síntese. 2. Reforma catalítica. 3. Indústria 4.0. 4. Machine  
Learning. 5. Redes Neurais Artificiais. I. Lucena, Sérgio. (Orientação). II. Título.

660 CDD (22.ed.)

JOSÉ LEONARDO GOMES DE ARAÚJO

**MODELAGEM DE REFORMA CATALÍTICA SECA DE METANO A GÁS DE SÍNTESE, UTILIZANDO MACHINE LEARNING E REDES NEURAIS**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Graduação em Engenharia Química da Universidade Federal de Pernambuco, como requisito parcial à obtenção do grau de Bacharel em Engenharia Química.

Aprovado em: 03/11/2022

**BANCA EXAMINADORA**

---

Prof. Dr. Sérgio Lucena (Orientador)  
Universidade Federal de Pernambuco

---

Prof. Dr. Nelson Medeiros de Lima Filho (Examinador)  
Universidade Federal de Pernambuco

---

Prof. Dr. Humberto Dória Silva (Examinador)  
Universidade Federal de Pernambuco

## **AGRADECIMENTOS**

Agradeço aos meus pais, Severino e Sandra, por todo apoio e por tudo que me ensinaram. Sem eles, não teria me tornado a pessoa que sou hoje.

Ao meu orientador Sérgio Lucena pelo suporte no desenvolvimento deste trabalho, e pelo incentivo em explorar novas tecnologias.

Aos meus amigos e colegas de faculdade, em especial a Karolina, Joyce e Wesley, por todo o companheirismo e apoio durante a jornada da graduação.

Aos docentes e funcionários do Departamento de Engenharia Química da UFPE, por seu trabalho e por participar da minha formação.

## RESUMO

Estamos vivenciando o início da Era da Informação, onde enormes quantidades de dados são geradas a cada instante, principalmente em decorrência da rápida modernização e desenvolvimento tecnológico que têm acontecido nos mais diversos setores da sociedade. Com isso, no âmbito das indústrias estamos presenciando o desenrolar da 4<sup>a</sup> Revolução Industrial, ou Indústria 4.0, fortemente marcada pela automação, inteligência artificial e *Big Data*. Essa revolução e as novas tecnologias trazidas por ela, podem trazer grandes melhorias à indústria, modificando as relações de trabalho e aumentando significativamente a eficiência, segurança e qualidade dos processos industriais. Uma das tecnologias que trazem grande impacto em diversas áreas, é o campo da inteligência artificial, um conjunto de recursos computacionais que podem ser usados para modelar e resolver problemas de alta complexidade utilizando muito menos recursos do que métodos tradicionais. Na Indústria Química, inteligências artificiais podem ser usadas por exemplo na modelagem de processos complexos, não-lineares e que contenham grande número de variáveis associadas, tornando difícil desenvolver modelos fenomenológicos robustos e precisos o suficiente para prever esses processos. Um exemplo de processo complexo que pode se valer de tais tecnologias é a produção de gás de síntese, através da reforma catalítica de metano. O gás de síntese é uma mistura gasosa composta por gás hidrogênio e monóxido de carbono, de grande interesse econômico pois pode ser utilizado para obter vários produtos com alto valor agregado, como amônia, hidrogênio combustível, e hidrocarbonetos incluindo combustíveis sintéticos. Além disso o processo de reforma catalítica seca, consome o metano e o gás carbônico, os dois principais causadores de efeito estufa, transformando-os em compostos que terão menos impacto ambiental. Porém, devido à complexidade e baixa linearidade presente nestes processos, torna-se difícil gerar modelos fenomenológicos precisos, e que possam se valer dos dados disponíveis. Assim a utilização de inteligências artificiais, nesse caso, de métodos *Machine Learning* podem trazer avanços no desenvolvimento, controle e entendimento desses processos. Neste trabalho foram utilizados métodos de *Machine Learning*, dentre eles as redes neurais artificiais, para modelagem da reforma catalítica seca do metano para a produção de gás de síntese, a partir de dados experimentais obtidos na literatura. Os modelos desenvolvidos foram avaliados através de métricas de regressão e comparados entre si buscando aqueles que apresentaram melhor desempenho. Os modelos de redes neurais obtiveram o melhor desempenho entre os modelos testados, apresentando coeficientes de determinação acima de 0,97 na previsão das taxas de produção dos componentes do gás de síntese e erros relativos médios de até 6,22%. Os resultados apontam para a utilidade desses modelos para uso industrial e acadêmico na modelagem de processos químicos complexos.

**Palavras-chave:** Gás de síntese; Reforma Catalítica; Indústria 4.0; Machine Learning; Redes Neurais Artificiais.

## ABSTRACT

We live in the beginning of the Information Age, where huge amounts of data are generated at every moment, mainly as a result of the rapid modernization and technological development that has been taking place in the most diverse sectors of society. As a result, within industries we are witnessing the unfolding of the 4th Industrial Revolution, or Industry 4.0, strongly marked by automation, artificial intelligence and Big Data. This revolution and the new technologies brought about by it can bring great improvements to the industry, changing work relationships and significantly increasing the efficiency, safety and quality of industrial processes. One of the technologies that has a great impact in several areas is the field of artificial intelligence, a set of computational resources that can be used to model and solve highly complex problems using much less resources than traditional methods. In the Chemical Industry, artificial intelligences can be used, for example, in the modeling of complex, non-linear processes that contain a large number of associated variables, making it difficult to develop phenomenological models that are robust and accurate enough to predict these processes. An example of a complex process that can make use of such technologies is the production of synthesis gas (syngas), through the catalytic reform of methane. Syngas is a gaseous mixture composed of hydrogen gas and carbon monoxide, and is of great economic interest, as it can be used to obtain several products with high added value, such as ammonia, hydrogen fuel and hydrocarbons, including synthetic fuels. In addition, the dry catalytic reform process consumes methane and carbon dioxide, the two main causes of the greenhouse effect, transforming them into compounds that will have less environmental impact. However, due to the complexity and low linearity of these processes, it is difficult to generate accurate phenomenological models that can make use of the available data. Thus, the use of artificial intelligence, in this case, Machine Learning methods can bring advances in the development, control and understanding of these processes. In this work, Machine Learning methods, including artificial neural networks, were used to model the dry catalytic reform of methane for syngas production, based on experimental data obtained in the literature. The developed models were evaluated using regression metrics and compared to each other, looking for those that presented the best performance. The neural network models obtained the best performance among the tested models, presenting coefficients of determination above 0.97 in the prediction of the production rates of the syngas components and average relative errors of up to 6.22%. The results point to the usefulness of these models for industrial and academic use in modeling complex chemical processes.

**Keywords:** Syngas; Catalytic Reform; Industry 4.0; Machine Learning; Artificial Neural Networks.

## **LISTA DE FIGURAS**

Figura 1 - Diagrama esquemático do experimento de reforma catalítica seca.....	18
Figura 2 - Fluxograma da gaseificação e co-gaseificação de carvão e biomassa.....	20
Figura 3 - Representação da estrutura de uma rede neural.....	24
Figura 4 - Esquema de representação de um neurônio genérico. ....	24
Figura 5 – Gráficos de distribuição dos dados de entrada.....	35
Figura 6 - Matriz de correlação linear entre dados de entrada e saída. ....	36
Figura 7 - Gráficos Q-Q para os modelos de machine learning simples.....	40
Figura 8 - Gráficos Q-Q para as redes neurais. ....	42
Figura 9 - Gráficos Q-Q por variável alvo, para previsões da rede neural desenvolvida por Ayodele e Cheng (2015).....	44

## **LISTA DE TABELAS**

Tabela 1 - Nível de correlação para o coeficiente de Pearson.....	31
Tabela 2 – Informações numéricas dos dados experimentais.....	36
Tabela 3 – Desempenho de treino por modelo, para as variáveis $Y_1$ e $Y_2$ .....	38
Tabela 4 – Desempenho de treino por modelo, para as variáveis $Y_3$ e $Y_4$ .....	39
Tabela 5 - Desempenho por modelo selecionado frente ao conjunto de teste.....	39
Tabela 6 - Número de camadas e neurônios por camada para as ANNs, por variável alvo.....	41
Tabela 7 - Desempenho das ANNs frente ao conjunto de teste. ....	42

## LISTA DE ABREVIATURAS E SIGLAS

ANN	<i>Artificial Neural Network</i>
ERM	Erro Relativo Médio
GBR	<i>Gradient Boost Regressor</i>
HPO	<i>HyperParameter Optimization</i>
IA	Inteligência Artificial
IDE	<i>Integrated Development Environment</i>
IoT	<i>Internet of Things</i>
MAE	<i>Mean Absolute Error</i>
MIMO	<i>Multiple Input Multiple Output</i>
MISO	<i>Multiple Input Single Output</i>
ML	<i>Machine Learning</i>
MSE	<i>Mean Squared Error</i>
RWGS	<i>Reverse Water Gas Shift</i>
SVM	<i>Support Vector Machine</i>
SVR	<i>Support Vector Regression</i>
WGS	<i>Water Gas Shift</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>13</b>
2.1	GÁS DE SÍNTESE	14
<b>2.1.1</b>	<b>Reforma catalítica de gás natural</b>	<b>15</b>
2.1.1.1	Reforma catalítica a vapor	15
2.1.1.2	Reforma catalítica seca	16
<b>2.1.2</b>	<b>Processos de gaseificação</b>	<b>18</b>
2.2	MACHINE LEARNING	20
<b>2.2.1</b>	<b>Redes Neurais</b>	<b>22</b>
2.2.1.1	Estrutura de redes neurais	23
2.2.1.2	Treino de redes neurais	25
2.2.1.3	Construção e hiperparâmetros	26
<b>2.2.2</b>	<b>Machine Learning na Indústria Química</b>	<b>27</b>
2.2.2.1	Machine Learning na produção de Gás de Síntese	28
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>30</b>
3.1	SOFTWARES	30
3.2	CONJUNTO DE DADOS	30
3.3	METODOLOGIA	31
<b>3.3.1</b>	<b>Análise de dados</b>	<b>31</b>
<b>3.3.2</b>	<b>Tratamento de dados</b>	<b>31</b>
3.3.2.1	Normalização	31
3.3.2.2	Separação dos conjuntos teste e treino	32
<b>3.3.3</b>	<b>Desenvolvimento dos modelos de Machine Learning</b>	<b>32</b>
<b>3.3.4</b>	<b>Desenvolvimento das Redes Neurais Artificiais</b>	<b>33</b>
<b>3.3.5</b>	<b>Confronto entre modelos</b>	<b>34</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>35</b>

4.1	ANÁLISE E TRATAMENTO DE DADOS	35
4.1.1	Análise exploratória dos dados	35
4.1.2	Normalização dos dados	37
4.2	DESENVOLVIMENTO DOS MODELOS	38
4.2.1	Modelos de Machine Learning simples	38
4.2.2	Redes Neurais Artificiais	41
4.3	COMPARAÇÃO AO TRABALHO ORIGINAL	43
5	CONCLUSÕES	46
	REFERÊNCIAS	47

## 1 INTRODUÇÃO

Nos últimos anos, o mundo está presenciando grandes avanços tecnológicos nas mais diversas áreas do conhecimento científico. A indústria química está diretamente envolvida nesses avanços, com desenvolvimento de novos materiais, produtos e processos que tornam a produção de bens mais eficiente e limpa. Além disso, a indústria química também é grandemente impactada por avanços no conhecimento em outras áreas tecnológicas, que têm crescido a cada dia e apresentado soluções rápidas, simplificadas e inteligentes para diversos setores.

Um exemplo de conhecimento que está cada vez mais em evidência no campo da Ciência de Dados, e que é estreitamente ligado à Estatística e à Computação, são os métodos de *Machine Learning* (ML), uma subdivisão da área de Inteligência Artificial (IA). Métodos deste tipo, quando aplicados, têm uma grande capacidade para modelar processos complexos, o que por muitas vezes é difícil a partir de equações fenomenológicas conhecidas (AL AANI *et al.*, 2019).

Por vezes, se torna difícil e custoso desenvolver e resolver modelos teóricos que possam descrever precisamente um sistema real, a exemplo das equações de Navier-Stokes para modelagem do comportamento de um fluido. Dessa forma, em alguns casos é necessário simplificar os modelos, para que se tenham soluções analíticas, mas simples, abrindo mão de acurácia para ganhar tempo na resolução do problema. Porém, esse balanço torna-se desafiador quando um problema complexo exige uma alta precisão (DOBBELAERE *et al.*, 2021).

Nesse sentido, métodos de inteligência artificial, em especial os de ML, podem ter grande vantagem na modelagem de problemas complexos, pois são modelos orientados por dados (*data-driven*), que se adaptam às especificidades dos processos e sistemas estudados (AL AANI *et al.*, 2019).

Um dos subtópicos de ML que tem recebido bastante atenção nos últimos tempos, especialmente nas ciências de forma geral, é o *Deep Learning*, também muito conhecido como campo das redes neurais. Uma rede neural é um modelo computacional que tenta simular a estrutura com que neurônios biológicos funcionam em nossos cérebros. São compostas de unidades básicas de processamento, também chamadas neurônios, organizados em múltiplas camadas em uma rede neural, organizadas através de um conceito hierárquico é o que dá a esses modelos a capacidade de aprendizagem (SARKAR; BALI; SHARMA, 2018).

Redes neurais têm sido usadas em diversos campos das ciências, no tratamento de variados problemas, como processamento de linguagem natural e imagens, agrupamento e classificação de dados, previsão e regressão de dados. Em especial, na indústria química, tem sido usada na

área de controle ajudando a prever o comportamento de processos e na detecção de falhas, redução de dimensionalidade de modelos e modelagem de dinâmicas complexas, permitindo gerar previsões acuradas para as variáveis de um processo a partir de dados históricos (LEE; SHIN; REALFF, 2018).

Um exemplo de processo que pode ser modelado a partir de redes neurais é a reforma catalítica de metano ( $\text{CH}_4$ ) e gás carbônico ( $\text{CO}_2$ ) para produção de gás de síntese, mistura gasosa composta essencialmente de monóxido de carbono (CO) e gás hidrogênio ( $\text{H}_2$ ). O gás natural é um recurso de grande importância energética a nível mundial, mas apesar disso poucos produtos podem ser sintetizados diretamente a partir do metano (SALHI *et al.*, 2011).

Dessa forma, a maior parte das sínteses a partir do metano são realizadas através da sua conversão para gás de síntese, que servirá de base para a produção de diversos outros produtos químicos com alto valor agregado, tais como o metanol, amoníaco, e uma variedade de hidrocarbonetos inclusive combustíveis sintéticos (SALHI *et al.*, 2011).

A conversão do gás natural ( $\text{CH}_4$ ) em outros produtos através da reforma a gás de síntese, além de agregar mais valor a esse composto é também um método mais amigável ao meio ambiente, pois a utilização de combustíveis sintetizados a partir desse processo podem apresentar uma redução de cerca de 90% em emissão de carcinogênicos e 50% de redução na emissão de  $\text{CO}_2$  quando comparada a queima combustíveis fósseis convencionais (AYODELE; CHENG, 2015).

Assim, o objetivo principal deste trabalho é desenvolver um sistema de redes neurais artificiais capazes de modelar o processo de reforma catalítica seca do metano a gás de síntese, com base nos dados obtidos por Ayodele e Cheng (2015), e avaliar seu desempenho frente a outros métodos mais simples de machine learning e também aos resultados no trabalho original dos autores citados.

De maneira geral, os objetivos específicos são:

- Treinar modelos de machine learning simples para previsão dos resultados no processo de reforma catalítica seca, e selecionar o de melhor desempenho.
- Otimizar hiperparâmetros e treinar redes neurais para previsão dos resultados no processo de reforma catalítica seca do metano.
- Comparar desempenho dos modelos de machine learning simples com o desempenho das redes neurais desenvolvidas.
- Comparar desempenho das redes neurais desenvolvidas com o desempenho das redes treinadas originalmente por Ayodele e Cheng (2015).

## 2 FUNDAMENTAÇÃO TEÓRICA

Todos os dias presenciamos o desenvolvimento de novas tecnologias que têm grande capacidade para mudar a forma com que conhecemos e encaramos o mundo e estas estão sendo rapidamente incorporadas na sociedade, como em aspectos simples do dia-a-dia, a exemplo de ferramentas de comunicação e entretenimento, ou na tomada de decisões em grandes empresas, que passam a ser orientadas por grandes volumes de dados, e na indústria de maneira geral, que se beneficia de tais tecnologias para melhoria na qualidade dos produtos e eficiência dos processos. Quando olhamos diretamente para indústrias, a inclusão dessas novas tecnologias faz parte da chamada Indústria 4.0, ou 4<sup>a</sup> Revolução Industrial (PERASSO, 2016).

As revoluções industriais, foram momentos de grande mudança nos meios de produção, marcados por avanços tecnológicos que mudaram definitivamente o funcionamento das indústrias. A 1<sup>a</sup> Revolução Industrial, ocorreu por volta de 1765 e foi marcada pela mecanização dos processos, que permitiu a acelerar a produção e substituir parte do trabalho humano, resultando numa produção em larga escala. Foi seguida pela 2<sup>a</sup> Revolução em 1870, marcada principalmente pelo desenvolvimento da Indústria Química e de novas formas de energia como o petróleo e a eletricidade, que permitiram a criação de novos inventos, como automóveis, aparelhos de rádio e telefones (A VOZ DA INDÚSTRIA, 2020).

Após a 2<sup>a</sup> Guerra Mundial, o conhecimento na área nuclear e quântica e também a corrida pela exploração espacial, permitiram não só o desenvolvimento de uma nova forma de geração de energia, mas também a criação de aparelhos e máquinas mais avançadas, com o desenvolvimento da microeletrônica, automação e biotecnologia. Com isso, iniciou-se a 3<sup>a</sup> Revolução Industrial, que originou boa parte dos produtos e métodos de produção que temos hoje (PORTAL DA INDÚSTRIA, 2021).

A 4<sup>a</sup> Revolução Industrial, ou Indústria 4.0, engloba diversas novas tecnologias avançadas como inteligência artificial, robótica, *Big Data*, Internet das Coisas (IoT, do inglês *Internet of Things*), computação em nuvem, entre outras. Essa nova etapa da evolução industrial tem mudado rapidamente e em grande escala os modelos de produção atuais, aumentando a eficiência de dos processos e melhorando a tomada de decisões baseando-se no grande volume de dados e *insights* permitidos por essas ferramentas (PORTAL DA INDÚSTRIA, 2021).

Com isso, a Indústria Química tem, cada vez mais à sua disposição, ferramentas que permitem a melhoria e desenvolvimento dos processos existentes. Ao utilizar recursos como o *machine learning* e inteligências artificiais para estudar e modelar processos complexos, não-lineares e que são complicados de analisar através dos modelos tradicionais, há bons resultados

com predições acuradas e que demandam menos tempo e esforços, uma vez que não é necessário conhecer a fundo os detalhes e particularidades das reações, equipamentos e condições experimentais do processo estudado (DOBBELAERE *et al.*, 2021).

Porém isso não significa que a utilização desses métodos substitui o estudo do processo por meios tradicionais, pelo contrário, esses conhecimentos são essenciais para entender e tomar decisões sobre o processo e entender o que os dados analisados e gerados por essas ferramentas significam. Não se trata de uma substituição, e sim da união de conhecimentos e ferramentas para tornar a indústria ainda mais eficiente e produtiva.

Um dos temas que está em evidência nos últimos anos é a produção de combustíveis, impulsionada pela crescente demanda de energia e pela preocupação com a preservação ambiental. Busca-se reduzir a emissão de gases estufa e a dependência de combustíveis fósseis, bem como alternativas na produção de combustíveis sustentáveis e formas de energia renovável. Porém a utilização de energias renováveis nem sempre é capaz de suprir completamente as necessidades de países industrializados por conta da dificuldade na extração desses combustíveis (YE, 2019).

Naturalmente partir de tais demandas surgem soluções não só para produzir combustíveis de fontes renováveis, mas também utilizar combustíveis fósseis de maneira mais inteligente e menos prejudicial ao meio ambiente. Um dessas formas é a conversão de gás natural (composto majoritariamente de metano (CH<sub>4</sub>) e gás carbônico (CO<sub>2</sub>) através de uma reforma catalítica para produção de gás de síntese(AYODELE; CHENG, 2015).

O gás de síntese pode ser utilizado para a produção de outros combustíveis e produtos como o metanol, reduzindo a emissão desses gases, principais responsáveis pelo efeito estufa. Além do consumo e transformação desses gases, combustíveis sintéticos produzidos a partir do gás de síntese podem apresentar uma redução de cerca de 90% em emissão de carcinogênicos e 50% de redução na emissão de CO<sub>2</sub> quando comparado a combustíveis fósseis convencionais (AYODELE; CHENG, 2015).

## 2.1 GÁS DE SÍNTESE

O gás de síntese consiste numa mistura gasosa formada por gás hidrogênio (H<sub>2</sub>) e monóxido de carbono (CO), podendo haver traços de outros compostos como gás carbônico (CO<sub>2</sub>), nitrogênio (N<sub>2</sub>) e metano (CH<sub>4</sub>). É utilizado na produção de compostos como metanol e amônia, além de combustíveis, como a gasolina sintética a partir do processo de síntese de Fischer-Tropsch O gás de síntese pode ser obtido através de diferentes processos e matérias primas,

dentre elas: reforma catalítica de gás natural, gaseificação parcial do carvão, gaseificação de biomassa, gaseificação de rejeitos sólidos e coque, entre outras (SALES, 2013).

### 2.1.1 Reforma catalítica de gás natural

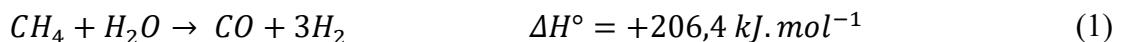
O processo de reforma catalítica do gás natural é bastante utilizado para a produção do gás de síntese, e em algumas refinarias com objetivo principal de produção de hidrogênio. Existem dois processos principais que são a reforma a vapor e a reforma a seco. Consiste num processo termoquímico com utilização de catalisador para promover a reação (SALHI *et al.*, 2011).

O grande desafio desses processos é o desgaste e desativação dos catalisadores devido às reações secundárias e deposição de carbono, principalmente pelo fato de que os catalisadores que apresentam melhor desempenho e mais resistências aos efeitos negativos, utilizam em sua composição metais preciosos, encarecendo e por vezes inviabilizando o processo (SALHI *et al.*, 2011).

#### 2.1.1.1 Reforma catalítica a vapor

O processo de reforma a vapor do metano é um dos mais antigos e mais utilizados para a produção de gás de síntese rico em H<sub>2</sub>, com foco na produção de hidrogênio combustível. A reforma a vapor normalmente é conduzida em fornos industriais de altas temperaturas, operando com reatores tubulares com leito catalítico. Uma consequência dessas condições de operação é o alto custo, refletido pelo grande consumo energético (em muitos casos parte do gás natural é usado como combustível para o processo) e também pelos custos do equipamento, uma vez que o reator precisa suportar as altas temperaturas (MELONI; MARTINO; PALMA, 2020).

A reação de reforma a vapor do metano é altamente endotérmica, sendo necessário operar a temperaturas entre 700 e 800 °C, e ocorre segundo a reação:



Essa reação é normalmente seguida pela reação de deslocamento gás-água (*water gas shift*, WGS) produzindo ainda mais H<sub>2</sub> e CO<sub>2</sub> que acaba sendo um subproduto nesse processo:



Com isso, o processo de reforma a vapor do metano apresenta grande custo-benefício para produção de hidrogênio a baixo custo, e resulta num gás de síntese com alta proporção de H<sub>2</sub>/CO, necessário para produção de H<sub>2</sub> puro. Foram desenvolvidos vários tipos de catalisadores, e um dos mais promissores para utilização em escala industrial é o catalisador de níquel depositado sobre alumina, por conta da sua eficiência e menor custo quando comparado aos demais catalisadores que utilizam em sua maioria metais preciosos (SALHI *et al.*, 2011).

Porém, o desenvolvimento de catalisadores eficientes e baratos para este processo ainda é desafiador, devido ao grande desgaste sofrido durante o processo, principalmente por conta da corrosão promovida pelo vapor. A deposição de carbono, e consequente desativação do catalisador, também é um problema para esse tipo de processo, mas pode ser bastante reduzida utilizando uma maior proporção de vapor em relação ao metano (de 2.5 a 3), o que por outro lado, aumenta os efeitos de corrosão (MELONI; MARTINO; PALMA, 2020).

Outros metais também são ativos como catalisadores do processo, mas acabam apresentando diversas desvantagens. O cobalto, por exemplo, acaba sofrendo desgaste intenso devido as altas pressões e a presença do vapor; o ferro, é rapidamente oxidado e assim acaba inativado; outros metais nobres, como platina, paládio e ródio, apresentam boa resistência mas são demasiadamente custosos para uma operação comercial em escala, inviabilizando sua utilização (MELONI; MARTINO; PALMA, 2020).

#### 2.1.1.2 Reforma catalítica seca

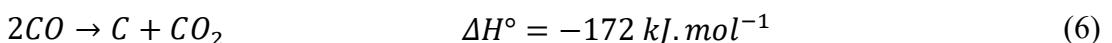
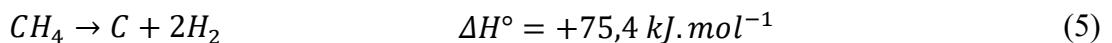
A reforma catalítica seca do metano, gera um gás de síntese com menor proporção de gás hidrogênio quando comparada à reforma a vapor, com proporção média de H<sub>2</sub>/CO = 1. Esse baixo rendimento é obtido principalmente pois além da reação de reforma, descrita pela Equação 3, ocorre simultaneamente a reação de deslocamento gás-água reversa (*reverse water gas shift*, RWGS), descrita pela Equação 4 (USMAN; WAN DAUD; ABBAS, 2015).



Mesmo com a baixa proporção de gás hidrogênio, essa reação de reforma é de grande interesse pois os dois principais gases causadores de efeito estufa são convertidos em compostos de maior valor agregado, permitindo não só uma redução na emissão de gases poluentes, mas

também a geração de valor a partir desses gases. Esse processo tende a produzir um gás de síntese numa proporção apropriada para a produção de gasolina sintética através da reação de síntese de Fischer-Tropsch, e nesse sentido apresenta uma redução de até 50% na emissão de CO<sub>2</sub> quando comparado à queima de combustíveis fósseis (AYODELE; CHENG, 2015).

Porém, mesmo com a importância do ponto de vista ambiental, a operação da reforma seca do metano pode ser ainda mais problemática, pois sofre grandemente com deposição de carbono, induzindo a desativação de catalisadores. Ocasionada pela decomposição do metano e pela reação de Boudouard, descritas pelas Equações 5 e 6, respectivamente (USMAN; WAN DAUD; ABBAS, 2015).



O grande desafio para esse processo é a obtenção de catalisadores apropriados, uma vez que a maioria dos catalisadores com boa atividade são desenvolvidos com a utilização de metais preciosos, a exemplo dos catalisadores de ródio e rutênio que apresentam um bom desempenho e resistência a desativação, porém seu alto preço e baixa disponibilidade limitam seu desenvolvimento e ampla utilização (CHEN; ZAFFRAN; YANG, 2020).

Outros metais mais baratos e abundantes têm sido testados para este processo. Catalisadores de níquel suportado em óxidos de outros metais, tem sido bastante utilizado, principalmente pelo seu baixo preço e facilidade de obtenção. Porém ainda não é o catalisador ideal, pois tem uma tendência a sofrer sinterização e desativação por deposição de carbono (MELONI; MARTINO; PALMA, 2020).

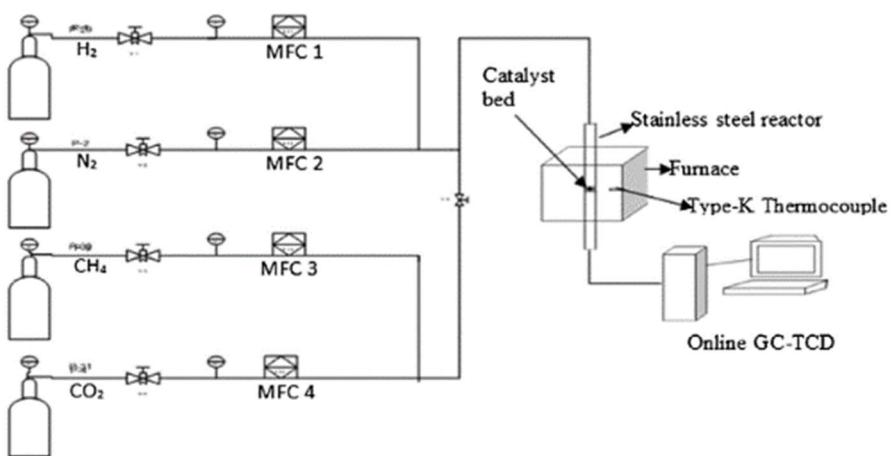
Combinações de cobalto e níquel, além de outros metais também têm sido testadas e apresentam bom desempenho e resistência, porém ainda não há estudos conclusivos sobre esse tipo de catalisador (CHEN; ZAFFRAN; YANG, 2020).

Usman, Wan Daud e Abbas (2015) analisaram a influência de uma série de parâmetros no processo de reforma seca, observando que o tamanho das partículas de catalisador pode ser uma variável útil para o controle da deposição de carbono. A utilização de partículas menores que 15 nm, apontou para uma redução na desativação por deposição, independentemente do tipo de catalisador. Nesse estudo, também foram apontadas combinações de metais nobres e níquel como catalisadores bi metálicos, o que ajudaria a melhorar a resistência e atividade, bem como

diminuiria os custos em comparação a utilização de catalisadores monometálicos de metais nobres.

Ayodele e Cheng (2015) utilizaram um catalisador de cobalto suportado sobre óxido de cério em um reator de leito fixo para realizar a reforma catalítica seca do metano. O experimento foi realizado em escala laboratorial para fins de modelagem e otimização, e os dados obtidos nesse experimento, foram utilizados no presente trabalho. A Figura 1 mostra o diagrama do esquemático do experimento realizado.

Figura 1 - Diagrama esquemático do experimento de reforma catalítica seca.



Fonte: (AYODELE; CHENG, 2015).

A reforma foi executada em um reator de aço inoxidável preenchido com o leito fixo catalítico de cobalto sobre óxido de cério, em um forno com temperatura controlada. A alimentação consiste na corrente gasosa composta pelos reagentes, metano ( $\text{CH}_4$ ) e gás carbônico ( $\text{CO}_2$ ) e os agente carreador gás nitrogênio ( $\text{N}_2$ ). O gás hidrogênio foi usado como agente redutor para o catalisador ao início do experimento, utilizando uma corrente de 20%  $\text{H}_2$  e 80% de  $\text{N}_2$  para iniciar a atividade catalítica (AYODELE; CHENG, 2015).

A temperatura foi medida através de um termopar acoplado ao forno, os fluxos gasosos controlados e medidos através de controladores digitais associados a cada corrente individualmente. A composição do gás de síntese produzido foi medida por cromatografia gasosa com detector de condutividade térmica, e a partir dos dados, calculadas as conversões e taxa de produção dos componentes do gás de síntese (AYODELE; CHENG, 2015).

### **2.1.2 Processos de gaseificação**

Processos de gaseificação consistem na conversão de matéria rica em carbono para gás através de aquecimento em determinadas condições, o que inclui processos como a pirólise,

hidrogenação e oxidação parcial, sendo o último o método mais utilizado para produção de gás de síntese. A composição do gás de síntese gerado nesse processo depende das condições de operação (temperatura, pressão, catalisador) e da composição da matéria prima, que em geral é biomassa (WEN; LU; PHUC, 2021).

A biomassa, em diversas formas tem sido utilizada como matéria prima no processo de gaseificação, sendo decomposta a altas temperaturas em vasos pressurizados. A gaseificação da biomassa exige inicialmente que haja uma pirólise para secagem do material e remoção de compostos voláteis. Então o material restante, predominantemente carbonáceo, sofre oxidação parcial para produção do gás de síntese (KAN *et al.*, 2019).

A pirólise exige temperaturas em torno dos 400 a 600°C e um ambiente livre de oxigênio. Já a gaseificação por oxidação parcial exige temperaturas ainda mais altas, em torno dos 800 a 900 °C e uma quantidade controlada de oxigênio (BAHADAR *et al.*, 2022).

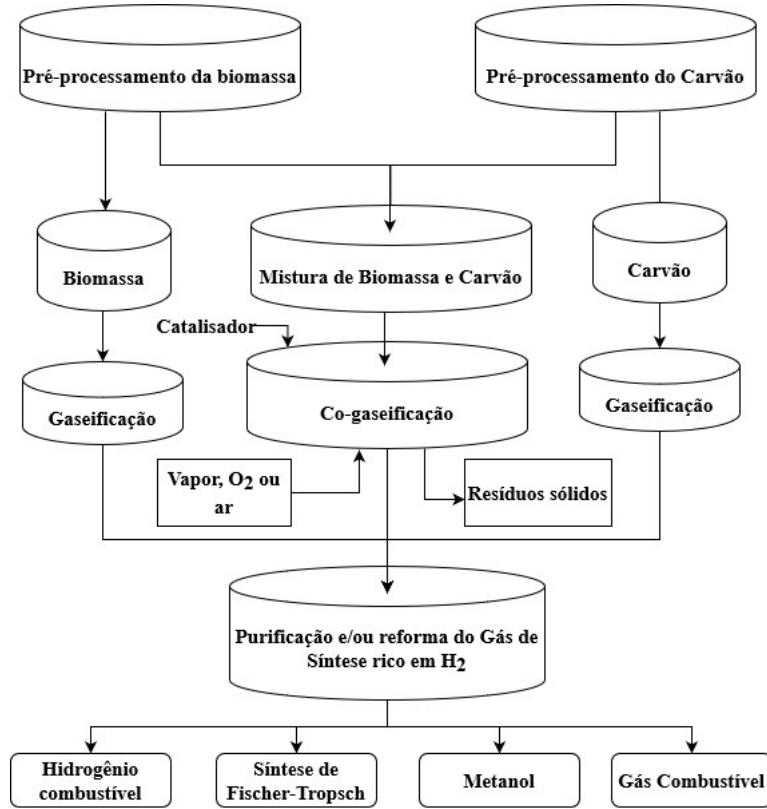
Também são estudados processos de co-gaseificação, onde mais de um tipo de biomassa e/ou outras matérias carbonáceas são utilizadas. Kan *et al.* (2019) apresentaram uma co-gaseificação combinando resíduos de horticultura e lodo de esgoto para produção de gás de síntese, usando CO<sub>2</sub> como oxidante, e utilizando as próprias cinzas como catalisador do processo, o que reduz custos e evita problemas de envenenamento e desgaste do catalisador.

Mansur *et al.*, (2020) fizeram um estudo de co-gaseificação utilizando carvão como matéria carbonácea principal, em *blends* combinados com serragem ou *pellets* de madeira, feitos a partir da serragem, para comparar o desempenho das duas matérias primas. Nesse estudo foi observado que a mistura de carvão e *pellets* produz um gás de síntese com maior proporção de H<sub>2</sub>/CO do que usando a serragem, dessa forma mostrando que um pré-processamento da matéria prima pode gerar um produto de maior valor agregado.

A Figura 2 é uma representação dos processos de gaseificação e co-gaseificação de biomassa e carvão. De maneira geral o processo pode ser dividido em etapas de preparação, gaseificação e processamento de produtos. As etapas de preparação incluem processos como secagem, trituração e compactação da matéria prima (BAHADAR *et al.*, 2022).

Já o processamento dos produtos varia de acordo com a finalidade esperada para o gás de síntese, podendo ir da purificação e separação dos gases, como para a produção de hidrogênio combustível com alto grau de pureza; produção de metanol; armazenamento do gás de síntese para servir de insumo na Síntese de Fischer-Tropsch; ou a utilização do próprio gás de síntese como gás combustível (BAHADAR *et al.*, 2022).

Figura 2 - Fluxograma da gaseificação e co-gaseificação de carvão e biomassa.



Fonte: Adaptado de BAHADAR *et al.* (2022)

## 2.2 MACHINE LEARNING

*Machine learning* (ML) é a denominação dada a um conjunto de técnicas computacionais contidas dentro do campo da Inteligência Artificial (IA), cujas bases são a Ciência da Computação, a Matemática e a Estatística. A ideia por trás do ML é gerar algoritmos capazes de ajudar a máquina a aprender a partir de dados.

Um algoritmo deve aprender a partir da *Experiência* ao processar os dados disponíveis, para executar determinadas *Tarefas*, enquanto a *Performance* desse algoritmo ao desempenhar tal tarefa é mensurada, com base em métricas definidas. É dito que este algoritmo aprende, quando a *Performance* nessas determinadas *Tarefas* aumenta com a *Experiência*, ou seja, ao processar mais dados no decorrer do tempo (SARKAR; BALI; SHARMA, 2018).

Dentre as tarefas definidas para o algoritmo, pode-se destacar: classificação (ou categorização) de dados; detecção de anomalias; regressão; agrupamentos; tradução e transcrição (ambas se aplicam vastamente no processamento de linguagem natural); dentre outras possíveis tarefas (SARKAR; BALI; SHARMA, 2018).

Dentre os principais paradigmas do ML destacam-se:

- **Aprendizado supervisionado**, no qual o modelo é treinado a partir de dados já classificados e tratados com uma estrutura bem definida de *inputs* e *outputs*. Um bom exemplo são modelos de Regressão, como *support vector machines* (SVM), redes neurais artificiais (ANNs) e *Gaussian process regression* (GPR).
- **Aprendizado não supervisionado**, onde o modelo é apresentado a dados sem classificação e sem uma estrutura explícita de *inputs* e *outputs*. O modelo deve então, encontrar os padrões dos dados e agrupá-los. Técnica utilizada principalmente quando não há relações claras entre *inputs* e *outputs*. Exemplos de métodos não supervisionados são *k-means clustering* e *support vector data description*.
- **Aprendizado por reforço**, onde o modelo é apresentado a um conjunto de dados e aprende com base numa “recompensa”, uma métrica computada através de uma função objetiva que deve ser maximizada, indicando o “caminho correto” que a rede deve tomar.

Olhando para o ML aplicado à engenharia química, pode-se destacar como principais pontos fortes desses métodos a grande velocidade de desenvolvimento junto à boa acurácia de um modelo devidamente treinado. Junto a isso, há grande escalabilidade desse tipo de modelo, permitindo que se aplique a diferentes escalas e processos (DOBBELAERE *et al.*, 2021).

Em contrapartida, esses modelos se comportam como uma “caixa preta”, então se torna difícil entender o funcionamento do modelo e como seus parâmetros se relacionam, conforme a complexidade do modelo escolhido. Um outro fator que deve ser levado em conta é a reproduzibilidade, pois o ajuste e treinamento dos modelos é fortemente influenciado pelos dados disponíveis para o treinamento, bem como pela grande quantidade de combinações de parâmetros que podem ser utilizados (DOBBELAERE *et al.*, 2021).

Técnicas de *machine learning* aplicadas à engenharia química não são novidade. Essa indústria já experimentou os impactos do ML em épocas anteriores, desde os anos de 1980 utilizando redes neurais “rasas” e outros modelos guiados por dados para fazer controle de processos e predição de falhas. (SCHWEIDTMANN *et al.*, 2021)

Um exemplo dessa utilização pioneira foi o sistema CONPHYDE, que auxiliava na previsão das propriedades termo físicas de misturas complexas, desenvolvido em 1983. O sistema era baseado no que se chamava de “sistemas experts”, modelos altamente especializados e impulsionados por heurística e conhecimento técnico dos processos. Vários outros modelos foram desenvolvidos a exemplo do DECADE (1985), para desenvolvimento de catalisadores e do MODELL.LA, uma linguagem de programação para modelagem de processos. Apesar do sucesso dos sistemas experts, os altos investimentos em tempo, dinheiro

e esforço para desenvolver tais ferramentas, acabaram inviabilizando seu uso na época (VENKATASUBRAMANIAN, 2019).

Porém, essas “ondas” de aplicação do ML tiveram seu desenvolvimento limitado por aspectos técnicos, como a baixa disponibilidade e qualidade dos dados, a falta de poder computacional e de linguagens e ambientes de desenvolvimento amigáveis ao usuário. Nos dias atuais, com a rápida evolução da tecnologia e barateamento do poder computacional, ferramentas de ML são muito mais acessíveis, tanto no ambiente industrial quanto no acadêmico (SCHWEIDTMANN *et al.*, 2021).

Além das melhorias de *hardware* e acesso, soluções de *software* também foram aprimoradas e popularizadas, tornando o desenvolvimento de modelos de ML mais simples e amigável para uso. Hoje, linguagens como MATLAB, R, C++ e principalmente Python, são bastante utilizadas para criar modelos de ML, tratar e analisar dados e desenvolver aplicações para uso em produção. A disponibilidade de enormes quantidades de dados também permitiu o desenvolvimento dessa área, introduzindo conceitos como a ciência de dados e o “*Big Data*” (VENKATASUBRAMANIAN, 2019).

Além das linguagens de programação, diversos *frameworks* foram criados e aprimorados para esse tipo de trabalho. Para o Python, podem ser citados o TensorFlow do Google e o scikit-learn para criação e treino de modelo; Pandas, numpy e scipy para leitura e tratamento de dados; matplotlib e Seaborn para visualização de dados, dentre outras bibliotecas. O desenvolvimento destes *frameworks*, além de tornar a etapa técnica mais amigável e simplificada, expande as possibilidades do que pode ser feito com ML e reduz drasticamente o tempo e esforço necessário para o seu desenvolvimento.

Hoje, para desenvolver uma rede neural não é mais necessário programa-la do zero, levando em conta toda a matemática e estatística por trás dos modelos, e com algumas linhas de código é possível tratar os dados, criar o modelo, treiná-lo e obter os resultados. Utilizando por exemplo frameworks como Pycaret, uma ferramenta do que é chamado AutoML, que facilita o processo de tratamento dos dados e treinamento de diversos modelos simultaneamente, permitindo a escolha daqueles que apresentam melhor desempenho para serem otimizados e aprofundados.

### **2.2.1 Redes Neurais**

Dentre os métodos de *machine learning*, um dos conjuntos de maior destaque são as redes neurais, devido a sua grande versatilidade e adaptabilidade à diversos tipos de problemas, desde classificação de dados, predição e regressão, até mesmo processamento de imagens e linguagem

natural. O campo em que atuam as redes neurais é chamado de *Deep Learning*, e esses algoritmos são baseados num conceito em que as várias representações dos dados são aprendidas em diferentes camadas na estrutura da rede(SARKAR; BALI; SHARMA, 2018).

A interação e hierarquia entre as diferentes camadas, possibilita construir uma inteligência para a máquina, a partir da forma que os dados são representados. Esse tipo de arquitetura baseado em camadas é, de maneira geral, o “coração” dos algoritmos de *Deep Learning* (SARKAR; BALI; SHARMA, 2018).

Redes neurais imitam o funcionamento de neurônios humanos para representar as formas de relacionamento não-lineares entre os dados, e são compostas por nós (chamados “neurônios”), elementos processadores básicos das redes, com valores associados a cada um (pesos e *bias*) que são “aplicados” aos sinais de entrada da rede, e o conjunto de operações realizadas por toda a rede determinam os valores de saída (BAHADAR *et al.*, 2022).

O objetivo do treino de uma rede neural, é calcular e determinar os melhores pesos e *bias* para cada neurônio da rede de forma a retornar valores o mais próximo dos dados reais utilizados no treino. Uma rede treinada, deve ser capaz de gerar resultados confiáveis para dados fora do conjunto de dados de treino, ou seja, deve ser capaz de fazer previsões acuradas para inputs que ela nunca tenha entrado em contato (VALENTE; VALENTE, 2021).

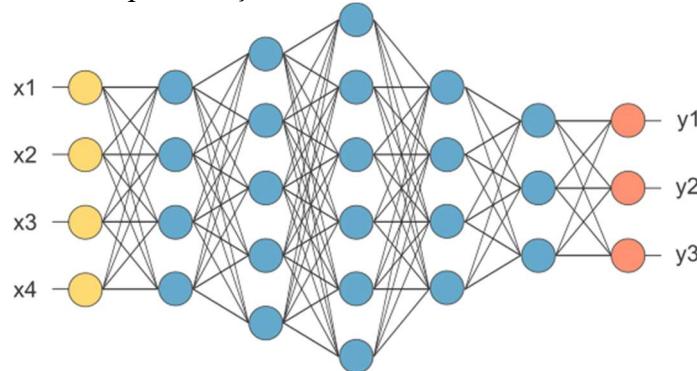
Esse método de treinamento cria um modelo “caixa preta” do ponto de vista do fenômeno estudado, pois não são utilizadas equações fenomenológicas na predição e todos os parâmetros de pesos e *bias* calculados no treino das redes não necessariamente têm sentido físico ou químico atribuído, e são valores compreendidos apenas pela rede (VALENTE; VALENTE, 2021).

#### 2.2.1.1 Estrutura de redes neurais

As redes neurais consistem em nós, também chamados neurônios, que são as unidades mais básicas de processamento em uma rede. O aglomerado de neurônios num mesmo nível é chamado de camada (*layer*). Uma rede pode ser composta por diversas camadas de neurônios, contendo pelo menos uma camada de entrada (*input*) e uma de saída (*output*), as demais camadas intermediárias são chamadas de camadas ocultas (*hidden layers*). Os dados de entrada são repassados para os neurônios da próxima camada, processam esses dados de acordo com os pesos atribuídos a cada neurônio, e passam o valor processado para a próxima camada através de uma função de ativação até alcançar a camada de saída, onde resultados são computados (ELMAZ; YÜCEL; MUTLU, 2020).

Um exemplo de estrutura de rede neural é apresentado na Figura 3, onde os círculos amarelos representam os pontos de entrada dos dados (inputs), os vermelhos os pontos de saída (outputs), e no meio, os azuis representam os neurônios das chamadas “camadas ocultas” da rede neural. Cada coluna de neurônios representa uma camada, e normalmente cada neurônio interage apenas com os neurônios das camadas anterior e posterior, não interagindo com outros neurônios da própria camada.

Figura 3 - Representação da estrutura de uma rede neural.

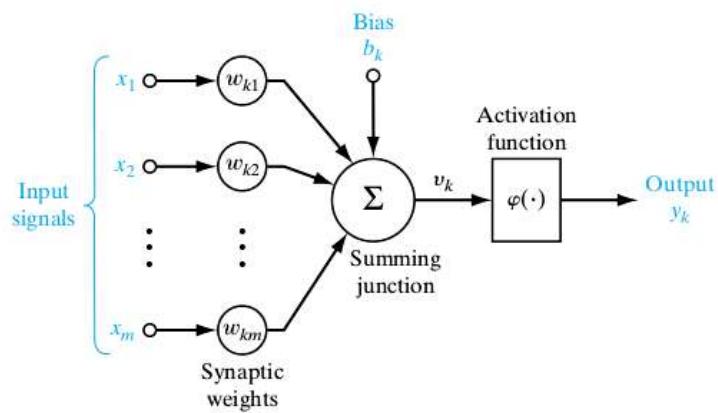


Fonte: (COELHO, 2017)

Cada neurônio recebe os valores de todos os neurônios da camada anterior, que são ponderados por pesos associados a cada neurônio, e então somado com um parâmetro *bias* e então esse valor é passado para uma função de ativação responsável por combinar linearmente os valores computados nos neurônios. As funções de ativação têm como objetivo gerar o sinal de saída de cada neurônio, através de uma transformação escalar e determinar o “estado ativado” de cada neurônio. Essas funções são o que dão às redes neurais a capacidade de se ajustar a dados não-lineares (DING; QIAN; ZHOU, 2018).

A estrutura de um neurônio genérico é ilustrada na Figura 4.

Figura 4 - Esquema de representação de um neurônio genérico.



Fonte: (COELHO, 2017)

Generalizando, a equação do neurônio  $k$  pode ser descrita como:

$$v_k = b_k + \sum x_i \omega_{k,i} \quad (7)$$

e o sinal de saída é definido através da função de ativação ( $\varphi$ ) da camada:

$$y_k = \varphi(v_k) \quad (8)$$

A função de ativação pode assumir diversas expressões, dentre as mais comuns podem-se destacar:

- Sigmoidal (sigmoid):

$$\varphi(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

- ReLU:

$$\varphi(x) = \max(0, x) = \begin{cases} x, & \text{se } x \geq 0 \\ 0, & \text{se } x \leq 0 \end{cases} \quad (10)$$

### 2.2.1.2 Treino de redes neurais

O objetivo do treino de uma rede neural, é calcular e determinar os melhores pesos e *bias* para cada neurônio da rede de forma a retornar valores o mais próximo dos dados reais utilizados no treino. Uma rede treinada, deve ser capaz de gerar resultados confiáveis para dados fora do conjunto de dados de treino, ou seja, deve ser capaz de fazer previsões acuradas para *inputs* que ela nunca tenha entrado em contato (VALENTE; VALENTE, 2021).

Esse método de treinamento cria um modelo “caixa preta” do ponto de vista do fenômeno estudado, pois não são utilizadas equações fenomenológicas na predição e todos os parâmetros de pesos e *bias* calculados no treino das redes não necessariamente têm sentido físico ou químico atribuído, e são valores compreendidos apenas pela rede (VALENTE; VALENTE, 2021).

Um dos métodos mais populares para eficientes para o treino de redes neurais é o *Backpropagation* (ou retropropagação). Algoritmos de treino baseados neste método, em geral são construídos em duas etapas: a propagação e o ajuste de pesos. Na etapa de propagação, a

rede neural é alimentada pelos *inputs* do conjunto de dados de treino, esses dados são processados pela rede, inicializada com valores aleatórios ou pré-definidos para os pesos e *biases*. O processamento desses dados gera valores de saída, que são então comparados aos *outputs* reais do conjunto de treino. Então são computados os erros entre os valores gerados pela rede e os valores reais, desejados (SARKAR; BALI; SHARMA, 2018).

Esses erros são então retropropagados pela rede, refazendo os cálculos em cada neurônio e gerando um erro para o ajuste de pesos. O algoritmo de treino computa o gradiente dos pesos com base no erro para cada neurônio, e aplica a “taxa de aprendizagem” para definir a porcentagem do valor do gradiente que será subtraída do peso daquele neurônio. O cálculo do gradiente normalmente é feito a partir de algoritmos de otimização e funções similares, como por exemplo o gradiente descendente estocástico (SARKAR; BALI; SHARMA, 2018).

Essas etapas são repetidas diversas vezes, até que se obtenha um baixo erro, que seja satisfatório para o problema. Essas múltiplas repetições/iterações são as chamadas épocas (SARKAR; BALI; SHARMA, 2018).

#### 2.2.1.3 Construção e hiperparâmetros

Ao treinar modelos de *machine learning*, os parâmetros internos do modelo são ajustados para se adaptar bem aos dados. Porém, além de parâmetros internos do modelo (pesos, *bias*, etc.), também existem os chamados hiperparâmetros, que não podem ser ajustados a partir dos dados, e devem ser definidos antes mesmo do treinamento. Hiperparâmetros são a configuração do modelo, definindo sua estrutura e controlando o processo de aprendizagem (YANG; SHAMI, 2020).

Todos os modelos de machine learning possuem hiperparâmetros, na maioria dos casos diferentes entre si, mas em especial para redes neurais é possível citar exemplos como:

- **Número de camadas ocultas**
- **Número de neurônios por camada** (pode haver diferente quantidade de neurônios em cada camada)
- **Função de erro**
- **Algoritmo de otimização** para o ajuste dos pesos
- **Taxa de aprendizagem**: taxa que será aplicada sobre o gradiente ao recalcular pesos e bias dos neurônios.
- **Tamanho da amostra por propagação**: define o volume de amostras que será usada em cada propagação completa, antes de reajustar os parâmetros da rede.

- **Número de épocas:** quantidade de iterações/ciclos de treinamento que serão executados
- **Métodos de parada antecipada:** essa função avalia a variação do erro, num determinado número de épocas, caso não haja melhoria, é feita uma parada antecipada no treinamento.

A definição dos hiperparâmetros (processo de *hyperparameter tuning* ou *hyperparameter optimization*, HPO) pode ser feita manualmente, ajustando os valores, treinando o modelo e avaliando o desempenho do treino. Esse método funciona bem para modelos com poucos hiperparâmetros, porém para modelos mais complexos fazer esse processo manualmente se torna muito custoso e demorado. Além disso, exige um conhecimento mais aprofundado de cada um deles, para entender sua influência sobre o modelo e definir os novos valores que devem ser testados. Por vezes, a interação entre diferentes hiperparâmetros tem uma relação não linear, o que torna o processo ainda mais difícil e não analítico (YANG; SHAMI, 2020).

Visando tornar o processo de HPO mais simples e menos custoso, é possível utilizar algoritmos que executarão esse processo repetidas vezes avaliando o desempenho de cada conjunto de hiperparâmetros testados para escolher aquele que apresenta melhor performance. Como não existe um método analítico para definir a combinação ótima de hiperparâmetros para cada modelo, em geral esses modelos fazem uma busca pseudoaleatória num intervalo determinado de valores. Pode-se citar como exemplo os métodos de *Grid Search* e *Randomized Search*, que testam os conjuntos de hiperparâmetros dentro de um intervalo de valores de maneira sequencial e randomizada, respectivamente (SARKAR; BALI; SHARMA, 2018).

### 2.2.2 Machine Learning na Indústria Química

Redes neurais são reconhecidas por ter uma vasta gama de aplicação e adaptabilidade aos mais diversos problemas, abrangendo áreas como processamento de imagens, linguagem natural, classificação de dados e regressão de sistemas não lineares. Na indústria química, um bom exemplo de seu uso são os sistemas de controle preditivo e de predição de falhas. Yao, Wang e Xu (2014) desenvolveram um sistema de monitoramento de falhas em processos em batelada, utilizando para isso o método de *support vector data description*. O método se mostrou válido quando testado no controle de um processo de produção de semicondutores, apresentando algumas melhorias.

Petsagkourakis *et al.* (2020) desenvolveram um método utilizando um modelo de aprendizado por reforço de uma rede neural recorrente para ajuste de parâmetros de controle

em processos bioquímicos complexos, sendo um dos casos de estudo a produção de ficocianina, pigmento azul produzido por cianobactérias que tem propriedades anti-alergênicas e antioxidantes. Aliado aos modelos fenomenológicos reais para a simulação do processo, este método obteve desempenho superior ao método tradicional de controle preditivo não linear, mesmo com uma quantidade reduzida de amostras para o treinamento do modelo.

O aprendizado supervisionado tem sido usado também para modelar sistemas dinâmicos e fazer previsões e otimização dos processos, auxiliando na tomada de decisão. Processos como cristalização, polimerização, destilações e outros, podem ser simulados como modelos orientados por dados (*data-driven*), permitindo um estudo mais rápido dos processos, levando em conta as especificidades de cada planta (SCHWEIDTMANN *et al.*, 2021).

#### 2.2.2.1 Machine Learning na produção de Gás de Síntese

A produção de gás de síntese pode ser feita a partir de diversas matérias primas de origem carbonácea, e as características dessas matérias primas impactam diretamente na eficiência do processo. Por muitas vezes é difícil determinar precisamente a composição dessas matérias primas para fazer uso de equações fenomenológicas para modelar o processo. Nessas situações, o uso de redes neurais e outros métodos de machine learning podem ser úteis pois não dependem de conhecer composições exatas, o modelo vai se adaptar aos dados e medições disponíveis, se elas forem significantes para os resultados (BAHADAR *et al.*, 2022).

Pandey *et al.* (2016) testaram diferentes arquiteturas no desenvolvimento de redes neurais para modelagem da produção de gás de síntese através da gaseificação de rejeito sólido em um reator de leito fluidizado. As simulações mostraram que as redes são uma alternativa viável na predição dessas composições, com algumas estruturas atingindo coeficientes de determinação de até 0.99.

Já George, Arun e Muraleedharan (2018) utilizaram um processo de gaseificação similar para conversão de diversos tipos de biomassa, e a modelagem a partir de uma rede neural do tipo MIMO resultou em um coeficiente de regressão de 0.987, além de apresentarem um breve estudo sobre a influência dos parâmetros de entrada na composição final do gás de síntese.

Wen, Lu e Phuc (2021) aplicaram métodos de machine learning para prever a composição do gás de síntese produzido a partir da gaseificação de cascas de arroz. Foram testados dois modelos: redes neurais artificiais (ANN) e *gradient boost regressor* (GBR), que chegaram a alcançar coeficientes de determinação de 0.89 e 0.93 respectivamente. Como dados de entrada, além de análise laboratorial das composições na matéria prima, foram usados parâmetros como

umidade, porcentagem de cinzas na amostra e quantidade de voláteis, para a previsão da composição do gás de síntese.

Quanto à reforma catalítica do metano a gás de síntese, Elmaz, Yücel e Mutlu (2020) fizeram uma modelagem preditiva da reforma seca de metano sobre um catalisador de cobalto, e utilizaram redes neurais, *support vector regression* (SVR) e regressão polinomial para fazer a previsão da conversão de metano e gás carbônico, e nesse caso obtiveram um melhor desempenho utilizando regressão polinomial, seguida de SVR.

Ye (2019) utilizou uma ANN para modelar a produção do metanol a partir do gás de síntese, e investigar o efeito do aumento da pressão parcial na conversão. Para validação foi executado um experimento mostrando que o aumento da pressão houve também um aumento da conversão de metanol, conforme previsto pelo modelo.

Bahadar *et al.*, (2022) testaram diversos algoritmos de machine learning, incluindo ANNs para modelar um conjunto de dados compilado de diversos autores para a modelagem da produção de gás de síntese. A comparação no desempenho de todos os métodos apontou o desempenho superior das ANNs para os dados utilizados, obtendo coeficientes de determinação de até 0.998, superando em todas as situações os demais métodos.

Ayodele e Cheng (2015) por sua vez, realizaram o experimento de reforma catalítica seca do metano em reator de leito fixo, utilizando como catalisador cobalto suportado em cério, e utilizaram redes neurais artificiais e o método de Box-Behnken para investigar os efeitos das pressões parciais de CH<sub>4</sub> e de CO<sub>2</sub> na taxa de produção do gás de síntese e a conversão dos reagentes.

A partir das ANNs, fizeram a previsão das composições atingindo coeficientes de determinação de 0.99, e utilizaram o modelo Box-Behnken, um método de superfície de resposta para determinar os inputs ótimos para maximizar a conversão. Após otimização dos parâmetros de processo, obtiveram conversões de 74,84% e 76,49% para o CH<sub>4</sub> e CO<sub>2</sub>, respectivamente.

### 3 MATERIAIS E MÉTODOS

#### 3.1 SOFTWARES

Para desenvolvimento deste trabalho, foi utilizada a linguagem de programação Python, na versão 3.7. Os *frameworks* e bibliotecas de código utilizados foram:

- Pandas, para manipulação de dados.
- NumPy e SciPy, para utilidades matemáticas e estatísticas.
- Seaborn e matplotlib, para visualização de dados.
- Scikit-learn, para tratamento de dados e desenvolvimento de métodos de machine learning.
- Pycaret, para desenvolvimento de métodos de machine learning.
- Desenvolvimento e treino de redes neurais: Keras, TensorFlow
- Otimização de hiperparâmetros de redes neurais: KerasTuner

Como ambientes de desenvolvimento integrados (IDEs) foram utilizados o VSCode e o Google Colab. Todos os *softwares* citados são de código aberto (*open source*) e gratuitos para uso.

#### 3.2 CONJUNTO DE DADOS

O conjunto de dados experimentais utilizados foram obtidos do trabalho de Ayodele e Cheng (2015). Neste conjunto, existem 57 amostras referentes ao experimento de reforma catalítica seca do metano para produção de gás de síntese realizado em escala laboratorial pelos autores citados, conforme detalhado na Seção 2.1.1.2. Cada amostra no *dataset* conta com 8 parâmetros, sendo 4 deles dados de entrada (*inputs*), variáveis manipuladas no processo:

- X<sub>1</sub>: Pressão parcial de gás metano (CH<sub>4</sub>), em kPa.
- X<sub>2</sub>: Pressão parcial de gás carbônico (CO<sub>2</sub>), em kPa.
- X<sub>3</sub>: Proporção de alimentação entre CH<sub>4</sub> e CO<sub>2</sub>, adimensional.
- X<sub>4</sub>: Temperatura do reator, em °C.

e 4 dados de saída (*outputs*), variáveis medidas, como resultado do experimento, sendo:

- Y<sub>1</sub>: Taxa de produção de gás hidrogênio (H<sub>2</sub>), em mmol.min<sup>-1</sup>.g<sub>cat</sub><sup>-1</sup>.
- Y<sub>2</sub>: Taxa de produção de monóxido de carbono (CO), em mmol.min<sup>-1</sup>.g<sub>cat</sub><sup>-1</sup>.
- Y<sub>3</sub>: Porcentagem de conversão de CH<sub>4</sub>.
- Y<sub>4</sub>: Porcentagem de conversão de CO<sub>2</sub>.

### 3.3 METODOLOGIA

#### 3.3.1 Análise de dados

Os dados foram incialmente analisados utilizando as ferramentas das bibliotecas Pandas, Seaborn e Matplotlib, buscando entender a distribuição dos dados e a possibilidade de outliers, dados faltantes ou errados, ou duplicações de amostras no dataset, para remoção ou correção das amostras afetadas. Foram observados os intervalos dos dados para cada parâmetro para definir os intervalos válidos para predição do modelo.

Utilizou-se uma matriz de correlação para avaliar o nível de correlação linear entre os parâmetros calculada através do método de Pearson, descrito na Equação 11:

$$r_{xy} = \frac{\sum z_x z_y}{N} \quad (11)$$

onde  $z$  é o desvio padrão das variáveis  $x$  e  $y$ , e  $N$  é o número de dados. Para a avaliação, assumiram-se os intervalos descritos na Tabela 1.

Tabela 1 - Nível de correlação para o coeficiente de Pearson.

Módulo Coef. de Pearson ( $r_{xy}$ )	Nível de correlação
$ r_{xy}  = 1$	Correlação perfeita
$0.8 \leq  r_{xy}  < 1$	Muito alta
$0.6 \leq  r_{xy}  < 0.8$	Alta
$0.4 \leq  r_{xy}  < 0.6$	Moderada
$0.2 \leq  r_{xy}  < 0.4$	Baixa
$0 <  r_{xy}  < 0.2$	Muito baixa
$r_{xy} = 0$	Nenhuma correlação

Fonte: Adaptado de SARKAR; BALI; SHARMA (2018).

#### 3.3.2 Tratamento de dados

##### 3.3.2.1 Normalização

Para evitar o efeito da escala e de diferentes dimensões no treinamento dos modelos, o conjunto de dados de entrada foi normalizado utilizando o método de normalização mín-máx, que dimensiona os dados no intervalo de 0 a 1, transformando os dados usando a Equação 12:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (12)$$

Sendo  $x_{min}$  e  $x_{max}$  o menor e o maior valor no intervalo do parâmetro  $x$  a ser normalizado.

Apenas os conjuntos de inputs foram normalizados, pois os outputs, como valores a serem previstos, não sofrem do efeito de escala e não precisam ser normalizados. A normalização foi feita utilizando a função “MinMaxScaler” do framework scikit-learn.

### 3.3.2.2 Separação dos conjuntos teste e treino

Os modelos de machine learning e redes neurais devem ser capazes de prever resultados acurados a partir de dados ainda não vistos pelo modelo. Para isso é necessário separar o conjunto de dados em dois conjuntos:

- **Conjunto de treino:** será utilizado na seleção de modelos, otimização de hiperparâmetros e treino do modelo final. Deve conter a maior quantidade dos dados e ser representar bem sua distribuição.
- **Conjunto de teste:** será utilizado apenas para validação, sendo apresentado ao modelo apenas após o treinamento para avaliar as previsões com dados não vistos até então.

Os dados foram separados em conjuntos de treino e teste usando o algoritmo de seleção pseudoaleatória “train\_test\_split” disponível no *framework* scikit-learn. O conjunto de teste contendo 80% das amostras, e o conjunto de destes os 20% restantes.

### 3.3.3 Desenvolvimento dos modelos de Machine Learning

Para desenvolvimento dos modelos, optou-se por uma abordagem MISO (*multiple input single output*), em que se utiliza todos as variáveis de entrada para prever apenas uma das variáveis de saída por modelo treinado. Dessa forma, para cada uma das variáveis de saída, foram treinados os seguintes modelos:

- Árvores de Decisão
- Florestas Aleatórias
- Gradient Boost Regressor
- Regressão Lasso
- Regressão Linear
- Support Vector Machine

Estes 6 modelos foram treinados utilizando o framework Pycaret, que cria e faz uma otimização prévia dos hiperparâmetros dos modelos, fazendo a avaliação e comparação de cada um com base em métricas predefinidas. Neste caso, as métricas definidas foram o erro médio absoluto (MAE), o erro médio quadrático (MSE), o coeficiente de determinação ( $R^2$ ) e o erro relativo médio (ERM), e também pela construção de um gráfico Q-Q (gráfico de probabilidade normal), para avaliação da reta de ajuste dos dados previstos e reais.

Para cada uma das variáveis de saída, selecionou-se o modelo que apresentou melhor performance com base nas métricas do treinamento. Então o modelo selecionado foi novamente treinado e otimizado, e então avaliado perante os dados do conjunto de teste, gerando as métricas finais para comparação com as redes neurais.

### 3.3.4 Desenvolvimento das Redes Neurais Artificiais

Para desenvolvimento das redes neurais artificiais, foram utilizados os frameworks TensorFlow e Keras, para construção e treinamento, e o KerasTuner para otimização de hiperparâmetros.

Optou-se por uma abordagem MISO (*multiple input single output*), construindo uma rede neural para cada uma das variáveis de saída, totalizando 4 modelos criados. Cada modelo separadamente passou por otimização de hiperparâmetros, treinamento e avaliação frente ao conjunto de dados de teste, para definição da melhor arquitetura de rede para previsão de cada variável de saída.

A otimização de hiperparâmetros foi feita utilizando o método *RandomSearch*, no framework KerasTuner, fazendo um treinamento prévio simplificado de diversas redes, selecionando combinações diferentes de hiperparâmetros. Os hiperparâmetros otimizados por algoritmo foram:

- **Número de camadas**, para a escolha 1 a 3 camadas.
- **Número de neurônios por camada**, utilizando entre 6 e 40 neurônios na primeira camada, e entre 6 e 20 neurônios nas camadas seguintes.
- **Taxa de aprendizagem**, utilizando 10 valores pré-selecionados entre 1 e 0.0005.
- **Função de erro**, utilizando MAE, MSE e Huber Loss.

Outros hiperparâmetros como o número de amostras por propagação, número de épocas e porcentagem de dados usados para validação cruzada durante o treino foram selecionadas manualmente em cada caso.

Após a determinação dos melhores hiperparâmetros, a rede neural foi efetivamente criada, treinada por um maior número de épocas e utilizando critérios de parada antecipada, caso a rede não melhorasse o desempenho após 50 épocas consecutivas. Caso apresentasse boa performance e ajuste aos dados, e as curvas de treino não demonstrassem *overfitting* ou *underfitting*, o modelo seguia para ser testado frente ao conjunto de teste.

O modelo treinado foi usado para previsão dos resultados a partir das amostras do conjunto de teste e os resultados previstos foram comparados aos valores reais, utilizando as métricas: erro médio absoluto (MAE), o erro médio quadrático (MSE), o coeficiente de determinação ( $R^2$ ) e o erro relativo médio (ERM), e também pela construção de um gráfico Q-Q (gráfico de probabilidade normal), para avaliação da reta de ajuste dos dados previstos e reais.

### **3.3.5 Comparação entre modelos**

Incialmente, o desempenho das redes neurais foi comparado ao desempenho dos métodos de machine learning mais simples desenvolvidos previamente, utilizando apenas o melhor modelo produzido para cada uma das variáveis de saída. A comparação foi feita por meio das métricas de cada modelo e dos gráficos de probabilidade normal. Num segundo momento, o desempenho das redes neurais desenvolvidas neste trabalho foi comparado ao desempenho da rede neural desenvolvida no trabalho original de Ayodele e Cheng (2015), que contém os dados utilizados neste trabalho e um modelo de ANN treinado a partir desses dados.

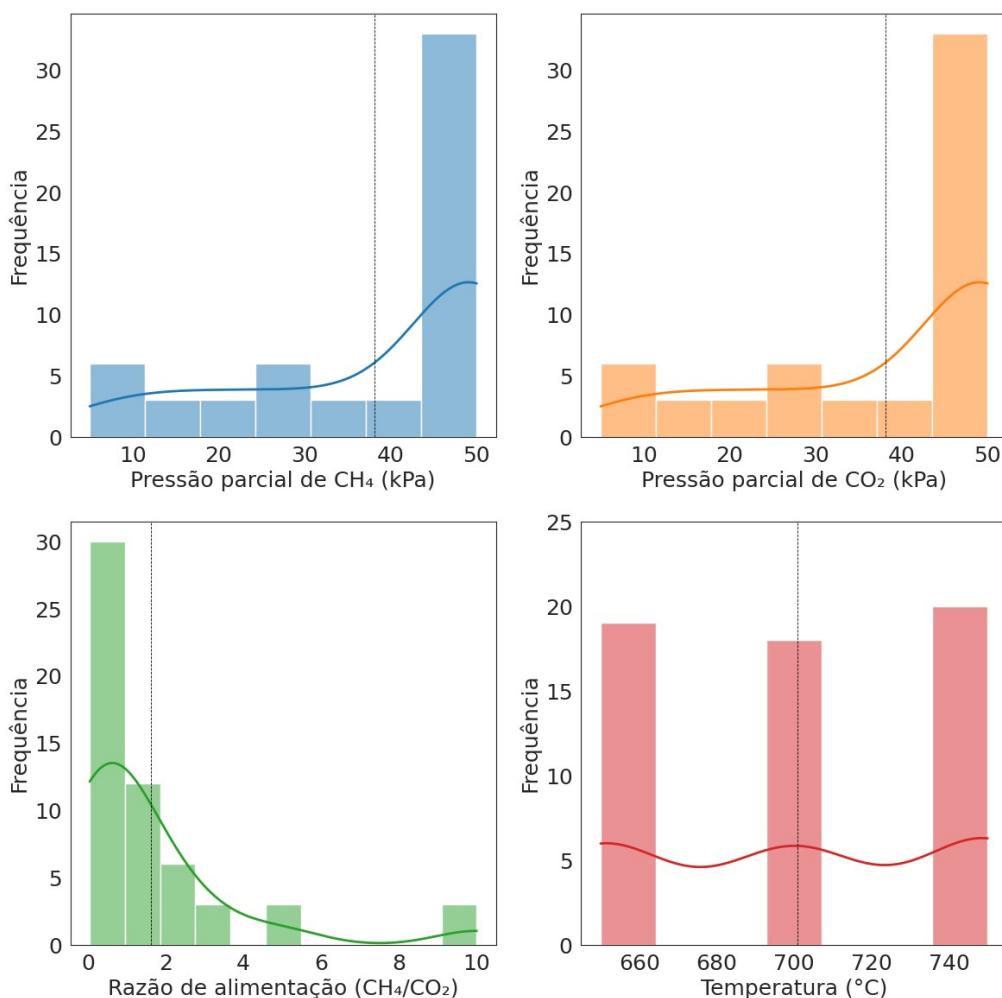
## 4 RESULTADOS E DISCUSSÃO

### 4.1 ANÁLISE E TRATAMENTO DE DADOS

#### 4.1.1 Análise exploratória dos dados

A análise inicial dos dados mostra a distribuição dos inputs como mostrado na Figura 5. Como se trata de condições experimentais dos experimentos, os dados não apresentam uma distribuição Gaussiana, e isso deve ser levado em consideração na hora de selecionar os conjuntos de dados para treino, para que a amostra de treino possa representar bem o dataset. Os dados de pressão parcial de CH<sub>4</sub> e CO<sub>2</sub> apresentam maior quantidade de amostras com valores entre 40 e 50 kPa, onde pouco mais da metade dos dados está contido nesse intervalo. Enquanto os valores de razão de alimentação estão um pouco melhor distribuídos ao redor do valor 1. A temperatura apresenta apenas três valores: 650, 700 e 750 °C.

Figura 5 – Gráficos de distribuição dos dados de entrada.



Fonte: o Autor.

Um resumo das informações numéricas sobre a distribuição de todos os dados é mostrado na Tabela 2.

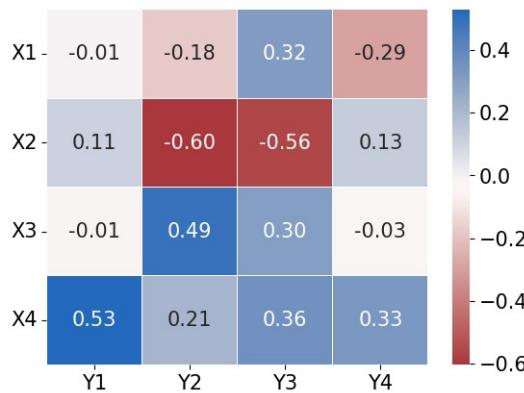
Tabela 2 – Informações numéricas dos dados experimentais.

	Variável	Média	Desvio padrão	Mín.	Máx.
X <sub>1</sub>	<b>Pressão parcial de CH<sub>4</sub> [kPa]</b>	38.16	15.46	5.00	50.00
X <sub>2</sub>	<b>Pressão parcial de CO<sub>2</sub> [kPa]</b>	38.16	15.46	5.00	50.00
X <sub>3</sub>	<b>Razão de alimentação (CH<sub>4</sub>/CO<sub>2</sub>)</b>	1.63	2.36	0.05	10.00
X <sub>4</sub>	<b>Temperatura (°C)</b>	700.88	41.72	650.00	750.00
Y <sub>1</sub>	<b>Taxa de produção de H<sub>2</sub> [mmol.min<sup>-1</sup>.g<sub>cat</sub><sup>-1</sup>]</b>	3.70	2.31	0.15	10.09
Y <sub>2</sub>	<b>Taxa de produção de CO[mmol.min<sup>-1</sup>.g<sub>cat</sub><sup>-1</sup>]</b>	5.12	2.04	1.20	8.38
Y <sub>3</sub>	<b>Conversão de CH<sub>4</sub> [%]</b>	52.61	14.82	23.15	79.46
Y <sub>4</sub>	<b>Conversão de CO<sub>2</sub> [%]</b>	64.06	18.16	17.86	90.68

Fonte: o Autor.

O conjunto de dados também não possui amostras com valores nulos ou outliers que precisam ser removidos do conjunto. Conhecendo os valores máximos e mínimos de cada parâmetro de entrada, definimos o intervalo que será válido para inputs com a intenção de fazer previsões do modelo que será gerado.

Figura 6 - Matriz de correlação linear entre dados de entrada e saída.



Fonte: o Autor.

Ao avaliar o grau de correlação linear dos dados através da Figura 6, observa-se que:

- X1 tem baixa correlação com Y3 e Y4, muito baixa com Y2. Pode-se desconsiderar a correlação com Y1.

- X2 tem uma correlação moderada com Y2 e Y3, e muito baixa com Y1 e Y4.
- X3 tem correlação moderada com Y4 e baixa com Y3. Pode-se desconsiderar a correlação com Y1 e Y4.
- X4 tem correlação moderada com Y1, baixa com Y3 e Y4 e muito baixa com Y2.

Essa informação mostra que os dados não apresentam correlação linear forte, porém é possível observar que algumas variáveis têm impacto significativo nas demais, a exemplo da temperatura (X4), que tem efeito significativo sobre todas as variáveis de saída. O que é esperado, pois a reação de reforma é altamente endotérmica, ou seja, é beneficiada pelo aumento da temperatura, porém não de maneira linear.

Observa-se também que a pressão parcial de CO<sub>2</sub> (X<sub>2</sub>) influencia bastante na taxa de produção do CO (Y<sub>2</sub>). Uma possível explicação para esse maior efeito é a reação de *reverse water gas shift*, representada pela Equação 4, e pode ser estendida também para a conversão de H<sub>2</sub>, pois este produto é consumido nessa reação paralela, tendo sua composição final reduzida no gás de síntese.

É esperado que a quantidade e proporção entre os reagentes também influencie no andamento da reação, e consequentemente nas conversões e produção do gás de síntese. Porém como observado na correlação da pressão parcial de CH<sub>4</sub> (X<sub>1</sub>) com a taxa de produção de H<sub>2</sub> (Y<sub>1</sub>), linearmente não parece haver uma relação estabelecida que possa ser representada por um modelo simples entre estes parâmetros.

Outros efeitos que podem estar associados ao processo não podem ser avaliados diretamente através dos dados experimentais disponíveis, como o desgaste do catalisador durante cada execução do experimento.

#### **4.1.2 Normalização dos dados**

Como visto na seção anterior, os dados de entrada têm dimensões e escalas bastante diferentes, o que vai impactar severamente os resultados do modelo, pois os inputs com maior ordem de grandeza acabam por receber maiores pesos nos parâmetros dos modelos de machine learning. Para minimizar esse efeito, é necessário passar todos os inputs para um intervalo de mesma escala através de um *scaling*.

Os dois principais métodos de *scaling* são a padronização e a normalização. Porém, como visto na Figura, os dados não estão distribuídos numa curva Gaussiana, dessa forma uma padronização não faria sentido, e o melhor método a ser aplicado é a normalização. Com isso, todos os dados de entrada passaram à uma escala de 0 a 1. Além da mudança de escala, nesse processo limita-se o intervalo válido para previsões, pois dados inseridos para fazer previsões

a partir dos modelos desenvolvidos, precisam estar dentro do intervalo de valores mínimos e máximos para cada variável de entrada, conforme mostrado na Tabela 2.

## 4.2 DESENVOLVIMENTO DOS MODELOS

No desenvolvimento dos modelos, optou-se por utilizar uma abordagem MISO ao invés de MIMO, pois os modelos convergiam com mais facilidade. Um modelo do tipo MIMO, ao tentar estimar valores de diferentes escalas (nesse caso as taxas de formação dos produtos, contra a conversão de reagentes) se torna mais complexo para treinar e ajustar os hiperparâmetros de forma a apresentar desempenho razoável, tanto para os modelos simples de machine learning, quanto para as redes neurais. Além disso, por se tratar de um pequeno volume de dados, o esforço computacional e tempo de treino necessário foram muito pequenos, permitindo então gastar um pouco mais de tempo treinando modelos separadamente para cada uma das variáveis de saída.

### 4.2.1 Modelos de Machine Learning simples

A partir do treino inicial, utilizando apenas os dados do conjunto de treino para avaliar e comparar o desempenho de cada modelo, selecionando o melhor com base nas métricas estabelecidas. A primeira métrica observada foi o coeficiente de determinação ( $R^2$ ), seguido do erro médio absoluto (MAE) e erro médio quadrático (MSE). O desempenho por modelo, para cada variável de saída após o treino inicial é mostrado nas Tabelas 3 e 4.

Tabela 3 – Desempenho de treino por modelo, para as variáveis Y<sub>1</sub> e Y<sub>2</sub>.

<b>Modelo</b>	<b>Y<sub>1</sub>: Taxa de produção de H<sub>2</sub></b>			<b>Y<sub>2</sub>: Taxa de produção de CO</b>		
	<b>MAE</b>	<b>MSE</b>	<b>R<sup>2</sup></b>	<b>MAE</b>	<b>MSE</b>	<b>R<sup>2</sup></b>
<b>Árvores de Decisão</b>	0.84	1.35	0.33	0.69	0.75	0.11
<b>Florestas Aleatórias</b>	0.76	1.06	0.54	0.55	0.51	0.73
<b>Gradient Boost Regressor (GBR)</b>	<b>0.66</b>	<b>0.79</b>	<b>0.61</b>	<b>0.46</b>	<b>0.41</b>	<b>0.73</b>
<b>Regressão Lasso</b>	1.66	4.34	-0.35	1.77	4.16	-1.51
<b>Regressão Linear</b>	1.65	4.09	-0.85	0.56	0.55	0.61
<b>Support Vector Machine (SVM)</b>	1.42	3.43	-0.20	0.72	0.89	0.62

Fonte: o Autor.

Tabela 4 – Desempenho de treino por modelo, para as variáveis Y<sub>3</sub> e Y<sub>4</sub>.

<b>Modelo</b>	<b>Y<sub>3</sub>: Conversão de CH<sub>4</sub></b>			<b>Y<sub>4</sub>: Conversão de CO<sub>2</sub></b>		
	<b>MAE</b>	<b>MSE</b>	<b>R<sup>2</sup></b>	<b>MAE</b>	<b>MSE</b>	<b>R<sup>2</sup></b>
<b>Árvores de Decisão</b>	6.27	96.32	0.13	9.75	157.43	-0.63
<b>Florestas Aleatórias</b>	<b>5.07</b>	<b>54.27</b>	<b>0.52</b>	<b>6.29</b>	<b>70.16</b>	<b>0.23</b>
<b>Gradient Boost Regressor (GBR)</b>	5.73	79.85	0.28	5.48	61.27	0.06
<b>Regressão Lasso</b>	10.75	164.04	-0.24	13.73	305.16	-1.06
<b>Regressão Linear</b>	7.71	92.63	0.16	13.28	293.10	-2.84
<b>Support Vector Machine (SVM)</b>	12.51	215.17	-0.52	14.23	364.67	-0.82

Fonte: o Autor.

A partir dos dados mostrados nas Tabelas 3 e 4, os modelos escolhidos foram o *Gradient Boost Regressor* (GBR) para a previsão de Y<sub>1</sub> e Y<sub>2</sub>, e o modelo de Florestas Aleatórias para previsão de Y<sub>3</sub> e Y<sub>4</sub>. Os modelos treinados nessa etapa não conseguiram se ajustar bem aos dados para previsão de Y<sub>4</sub>, a conversão de CO<sub>2</sub>. Como demonstrado na análise dos dados, essa variável tem baixa correlação com as variáveis de entrada estudadas, o que pode explicar a dificuldade desses modelos a, pois mesmo aqueles que usam métodos não lineares, a exemplo das Florestas Aleatórias, podem não apresentar profundidade suficiente para prever essa informação com precisão (BAHADAR *et al.*, 2022).

Após a seleção dos tipos de modelos mais promissores, foi feito um novo treinamento junto à uma otimização de hiperparâmetros mais rígida. Os novos modelos foram validados utilizando o conjunto de teste. A Tabela 5 traz o desempenho final de cada modelo, junto à equação da reta de ajuste do gráfico Q-Q de cada modelo, mostrados na Figura 7.

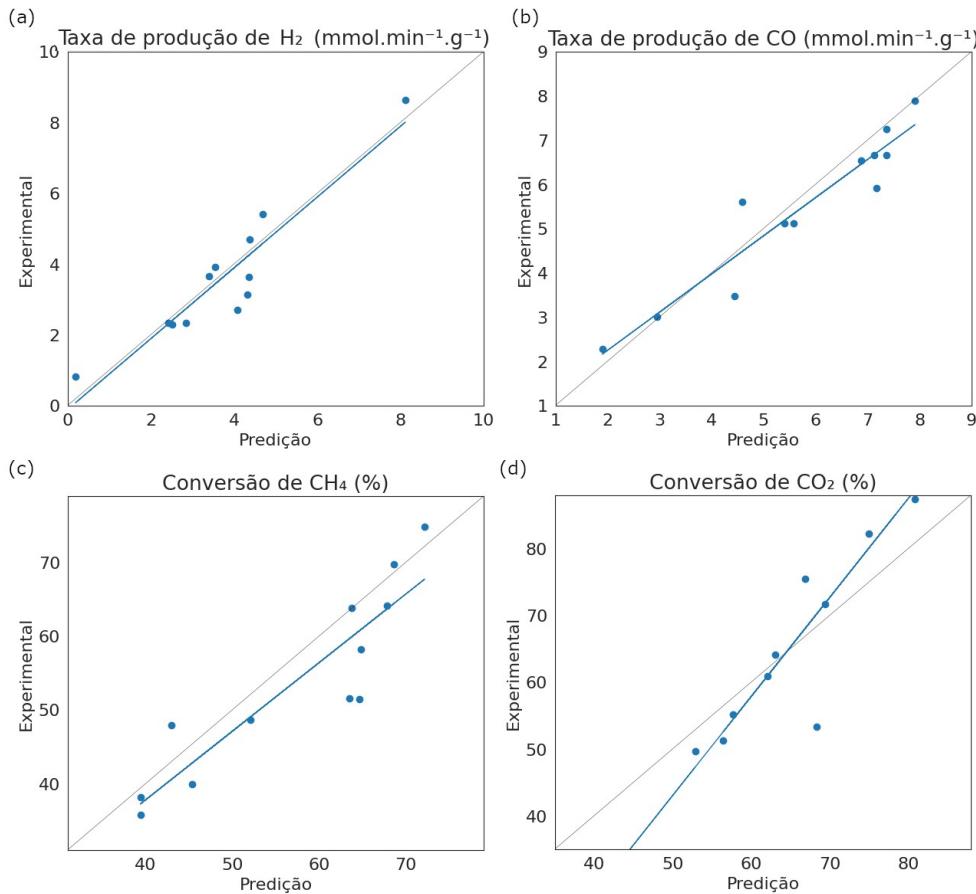
Tabela 5 - Desempenho por modelo selecionado frente ao conjunto de teste.

<b>Variável alvo</b>	<b>Modelo selecionado</b>	<b>MAE</b>	<b>MSE</b>	<b>R<sup>2</sup></b>	<b>ERM</b>	<b>Reta de ajuste</b>
<b>Y<sub>1</sub></b>	GBR	0.58	0.47	0.82	21.95%	$y = 1.00x - 0.11$
<b>Y<sub>2</sub></b>	GBR	0.50	0.40	0.91	10.36%	$y = 0.86x + 0.52$
<b>Y<sub>3</sub></b>	Florestas Aleat.	4.85	38.85	0.82	9.68%	$y = 0.93x + 0.59$
<b>Y<sub>4</sub></b>	Florestas Aleat.	6.21	56.87	0.88	10.43%	$y = 1.48x - 31.10$

Fonte: o Autor.

Figura 7 - Gráficos Q-Q para os modelos de machine learning simples.

Variáveis (a) Y<sub>1</sub>, (b) Y<sub>2</sub>, (c) Y<sub>3</sub>, (d) Y<sub>4</sub>.



Fonte: o Autor.

Nas Figuras 7a e 7b a seguir, percebe-se que os modelos do tipo GBR conseguiram explicar razoavelmente bem a maioria dos dados, para as taxas de produção de H<sub>2</sub> e CO (variáveis Y<sub>1</sub> e Y<sub>2</sub>, respectivamente), como indicado pelos coeficientes de determinação, na Tabela 5. O modelo para Y<sub>1</sub>, se aproximou bastante da reta normal ideal ( $y=x$ ), indicando uma boa precisão na previsão, já para Y<sub>2</sub> a reta teve coeficientes um pouco mais distantes do ideal. Esses modelos foram capazes de se ajustar bem aos dados e fazer previsões razoavelmente acuradas, porém, ainda com uma dispersão moderada.

Já para as conversões de CH<sub>4</sub> e CO<sub>2</sub> (variáveis Y<sub>3</sub> e Y<sub>4</sub>, respectivamente), nas Figuras 7c e 7d, os coeficientes de determinação, também indicaram um ajuste razoavelmente bom aos dados, porém houve uma maior dispersão nos valores preditos. Para Y<sub>3</sub>, apesar dos coeficientes da reta serem também próximos ao ideal, por conta da dispersão, a previsão não é tão precisa quanto a obtida com o GBR para Y<sub>1</sub> e Y<sub>2</sub>.

Para Y<sub>4</sub>, o modelo não teve performance boa, quando comparado aos demais. Apesar do coeficiente de determinação ser bom, o ajuste dos dados não foi preciso como é perceptível na Figura 7d, com a reta de ajuste dos dados se afastando bastante da reta normal ideal. Além de apresentar também uma alta dispersão e um valor de erro alto.

Os resultados de performance para os modelos escolhidos nessa etapa foram utilizados como valores de referência a serem batidos no treinamento das redes neurais, apresentadas na seção seguinte deste trabalho.

#### 4.2.2 Redes Neurais Artificiais

Diferentemente do treino inicial aplicado aos modelos simples de machine learning, o treino inicial aplicado às redes neurais teve como objetivo principal fazer a otimização de hiperparâmetros do modelo. A partir do treino inicial foram definidos os melhores hiperparâmetros para a construção dos modelos finais. O número de camadas e a quantidade de neurônios por camada para cada uma das redes é mostrado na Tabela 6.

Tabela 6 - Número de camadas e neurônios por camada para as ANNs, por variável alvo.

Variável alvo	Nº de camadas	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>
<b>Y1</b>	1	39	-	-
<b>Y2</b>	2	20	8	-
<b>Y3</b>	2	36	9	-
<b>Y4</b>	3	28	12	11

Fonte: o Autor.

É perceptível que não há um padrão nem similaridade na estrutura das redes neurais, mesmo ao comparar variáveis de mesma natureza (Y<sub>1</sub> com Y<sub>2</sub>, e Y<sub>3</sub> com Y<sub>4</sub>). Isso se dá tanto pelo método de HPO utilizado que busca configurações “aleatórias” dentro dos intervalos definidos, como também pela própria natureza das redes neurais, pois não há regras e métodos bem definidos para a escolha do número de neurônios e camadas (YANG; SHAMI, 2020).

Após o treino prévio, cada rede foi reiniciada e novamente treinada usando os hiperparâmetros selecionados a partir do HPO, de forma a obter os modelos finais. Os modelos finais testados frente aos dados do conjunto de teste apresentaram o desempenho mostrado na Tabela 7. Também são mostradas as equações das retas de ajuste das previsões feitas por cada rede para as variáveis alvo.

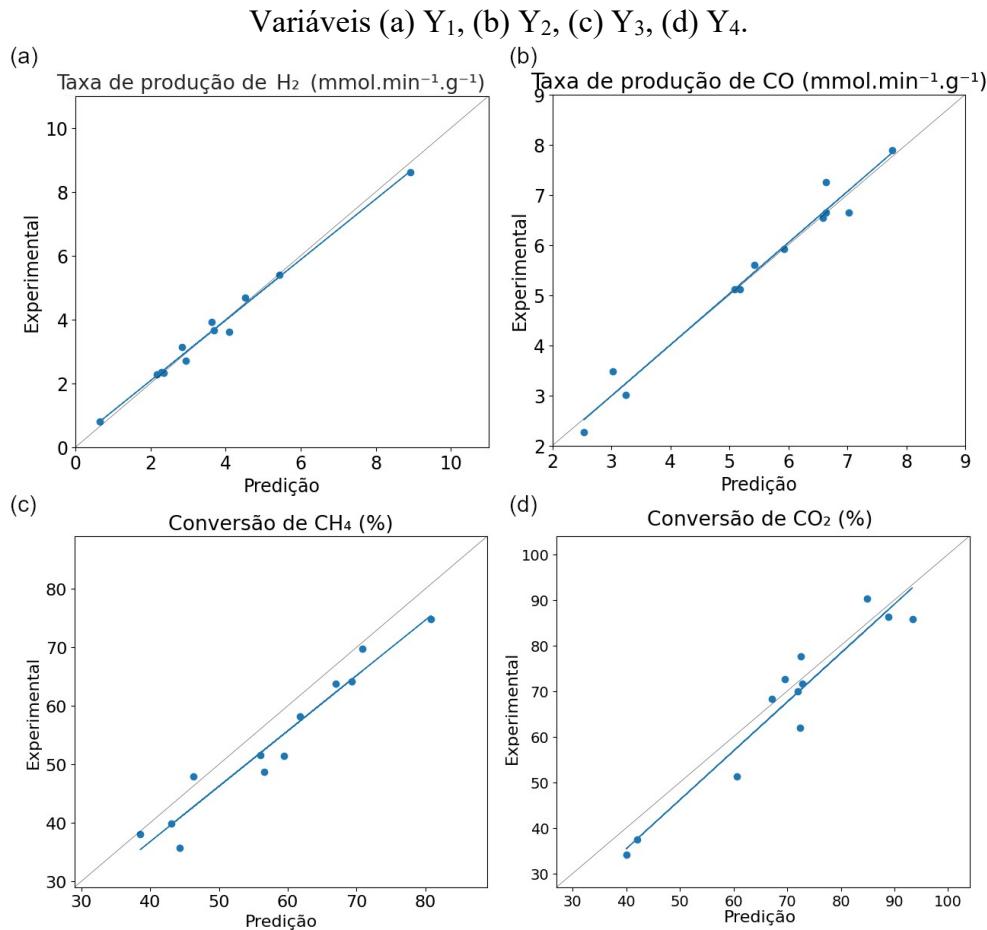
Tabela 7 - Desempenho das ANNs frente ao conjunto de teste.

Variável alvo	MAE	MSE	R <sup>2</sup>	ERM	Reta de ajuste
Y1	0.18	0.05	0.99	6.22%	$y = 0.95x + 0.19$
Y2	0.20	0.07	0.97	4.50%	$y = 1.02x - 0.07$
Y3	4.45	26.76	0.94	8.84%	$y = 0.95x - 1.09$
Y4	4.85	32.35	0.92	7.12%	$y = 1.07x - 7.49$

Fonte: o Autor.

Em comparação aos modelos mais simples, as redes neurais apresentaram uma melhoria na performance, quando comparamos as Tabelas 6 e 7. Observa-se que para todas as variáveis houve diminuição considerável em todas as métricas de erro. Também houve uma melhora no coeficiente de determinação, em especial para Y<sub>1</sub> e Y<sub>3</sub>. A Figura 8 traz os gráficos Q-Q para cada variável, com as previsões feitas com as redes neurais.

Figura 8 - Gráficos Q-Q para as redes neurais.



Fonte: o Autor.

Os coeficientes da reta de ajuste também ficaram bastante próximas do ideal (a reta  $y = x$ ), indicando que houve acurácia na previsão das variáveis alvo. Ao observar a Figura 8, percebe-

se que os dados estão também visualmente menos dispersos em comparação aos dos modelos mais simples, na Figura 7, condizendo com a redução no valor dos erros.

Em comparação ao modelo de Florestas Aleatórias, a rede neural conseguiu fazer previsões e se ajustar muito melhor aos dados, como é observado na comparação entre as Figuras 7d e 8d. Essa melhoria era esperada, dada a maior capacidade das redes neurais de descrever relações não lineares entre dados de maneira mais profunda que outros métodos não-lineares.

Mesmo assim, as previsões para as conversões de CH<sub>4</sub> e CO<sub>2</sub> (Y<sub>3</sub> e Y<sub>4</sub>, respectivamente) apresentaram dispersão um pouco maior em comparação as taxas de produção de H<sub>2</sub> de CO (Y<sub>1</sub> e Y<sub>2</sub>, respectivamente), com os modelos de redes neurais.

Comparando o desempenho entre as redes neurais, para evitar o efeito da escala dos dados, que impacta as métricas MAE e MSE, observasse-se o erro relativo médio (ERM), que normaliza os erros. Houve uma grande melhoria no ERM principalmente para a variável Y<sub>1</sub>, com uma diminuição de quase 16%.

Dessa forma, as redes neurais desenvolvidas tiveram melhor performance que os modelos mais simples de machine learning trabalhados anteriormente, destacando o Gradient Boost Regressor e o modelo de Árvores Aleatórias.

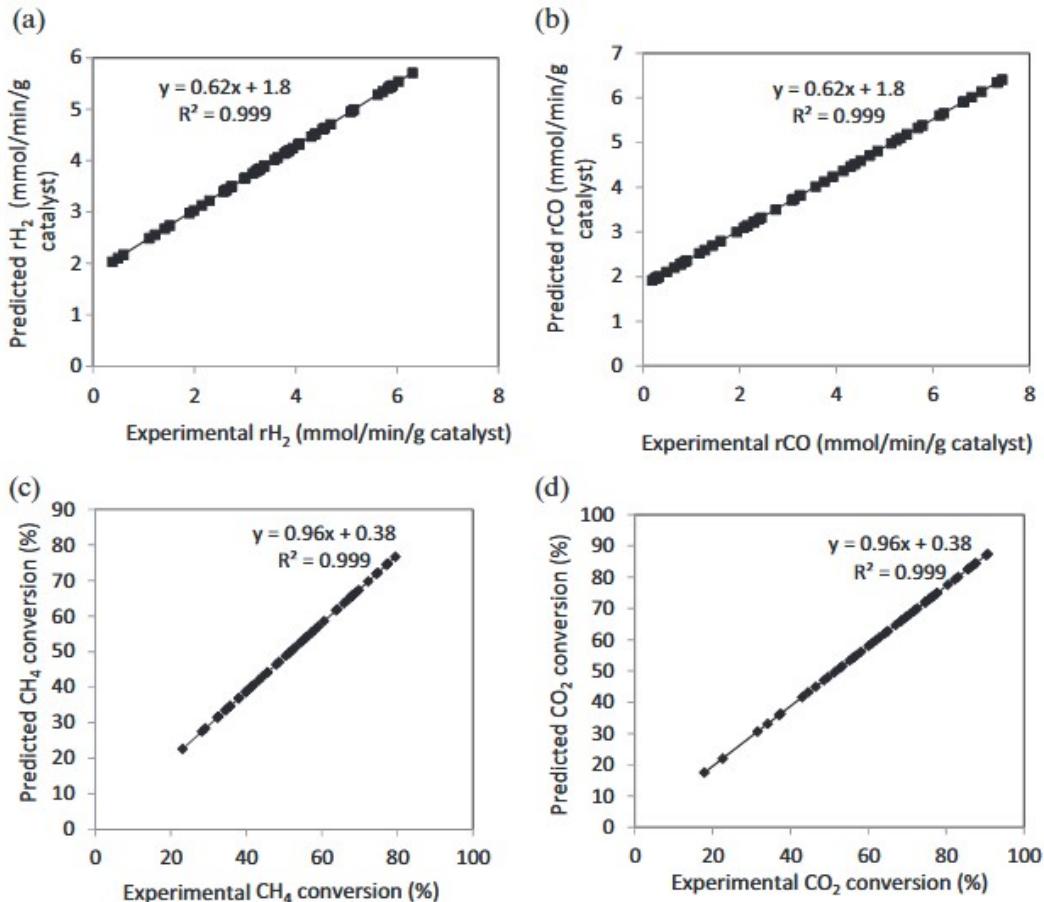
#### 4.3 COMPARAÇÃO AO TRABALHO ORIGINAL

No trabalho desenvolvido por Ayodele e Cheng (2015), foi utilizada uma estrutura de redes neurais MIMO, dessa forma apenas um modelo foi treinado, diferentemente do desenvolvido neste trabalho. A rede foi treinada baseada na métrica de MSE como função erro, porém o valor para essa métrica não foi disponibilizado no trabalho, e também não poderia ser usado como parâmetro de comparação, pois um modelo MIMO apresentaria apenas um valor para essa métrica, referente aos erros em todas as variáveis alvos. O modelo desenvolvido foi construído com uma única camada oculta, com 16 neurônios.

A avaliação do modelo do trabalho original foi feita principalmente a partir dos gráficos Q-Q e também do erro médio relativo. Os gráficos em questão são mostrados na Figura 9.

Figura 9 - Gráficos Q-Q por variável alvo, para previsões da rede neural desenvolvida por Ayodele e Cheng (2015).

Variáveis (a) Y<sub>1</sub>, (b) Y<sub>2</sub>, (c) Y<sub>3</sub>, (d) Y<sub>4</sub>.



Fonte: Adaptado de AYODELE e CHENG (2015).

No entanto, ao observar as Figuras 9a e 9b percebe-se que, apesar do R<sup>2</sup> ser muito próximo do ideal, as retas de ajuste se distanciam da reta normal ideal, indicando que as previsões feitas pelo modelo para Y<sub>1</sub> e Y<sub>2</sub> não foram acuradas. Tomando uma amostra dos dados e as equações das retas de ajustes, estima-se que haja por volta de 20% a 40% de erro relativo nas previsões de Y<sub>1</sub> e Y<sub>2</sub> para este modelo. Porém, para as taxas de produção de H<sub>2</sub> e CO, não foram apresentados os erros relativos médios no trabalho original, logo essa estimativa não pode ser confirmada.

Dessa forma, assume-se os modelos de redes neurais desenvolvidas neste trabalho têm um melhor potencial para prever as taxas de produção de H<sub>2</sub> e CO no processo de reforma catalítica seca do metano estudado, em comparação ao modelo apresentado no trabalho original por Ayodele e Cheng (2015). Isso por que estes novos modelos apresentaram uma baixa dispersão dos dados e uma reta de ajuste mais próxima da ideal, indicando uma maior acurácia na previsão das variáveis do processo.

Ayodele e Cheng (2015) apresentaram os valores de erro relativo médio para as conversões de CH<sub>4</sub> e CO<sub>2</sub>, sendo 3.21% e 3.34%, respectivamente. A partir dessas informações, observa-se que o desempenho do modelo desenvolvido no trabalho original é superior quanto às variáveis Y<sub>3</sub> e Y<sub>4</sub>, com melhor coeficiente de determinação e melhor ajuste.

Dessa o modelo proposto originalmente parece explicar melhor a relação entre os dados, para Y<sub>3</sub> e Y<sub>4</sub>, apresentando menor erro relativo e baixíssima ou quase nenhuma dispersão, como visto nas Figuras 9c e 9d, em comparação aos modelos desenvolvidos neste trabalho. Mesmo assim, o desempenho dos modelos novos se aproximou bastante do modelo original, apresentando uma dispersão razoavelmente baixa e erros relativos baixos.

## 5 CONCLUSÕES

O desenvolvimento de modelos de *machine learning* aplicados à uma reforma catalítica seca de metano, demonstrou o potencial desses métodos como ferramentas que podem ser aplicadas tanto na indústria química como na própria academia. Os resultados obtidos neste trabalho reforçam a ideia de que o uso de modelos de ML pode auxiliar na previsão acurada de dados e do comportamento de processos químicos, enquanto reduz o tempo e esforço necessário para desenvolvimento, em comparação a modelos fenomenológicos complexos.

Mesmo entre os métodos mais simples de ML foi possível obter modelos que se ajustassem aos dados permitindo previsões razoavelmente acuradas. Dentre eles, os modelos de *Gradient Boost Regressor* e de Árvores Aleatórias alcançaram um melhor desempenho entre os demais, atingindo  $R^2$  entre 0,82 e 0,91, porém ainda apresentando um erro relativo médio um pouco alto, principalmente para as variáveis de taxa de produção de H<sub>2</sub> e CO, 21,95% e 10,36% respectivamente.

Os modelos de redes neurais artificiais desenvolvidos, superaram o desempenho dos modelos mais simples, com destaque na previsão das variáveis de taxa de produção do H<sub>2</sub> e de CO, onde os  $R^2$  aumentaram para 0,99 e 0,97 e houve uma redução no erro relativo médio para 6,22% e 4,50%, respectivamente para as duas variáveis. Essa melhora no desempenho demonstra a maior capacidade das redes neurais de se ajustar a dados que possuem baixa correlação linear entre si, quando comparadas aos métodos de ML mais simples.

Porém, redes neurais são mais complexas e difíceis de treinar em relação aos métodos simples, exigindo um treinamento mais cuidadoso e trabalhoso durante a otimização de hiperparâmetros. Os modelos mais simples podem ser definidos com mais facilidade, utilizando ferramentas como o PyCaret. Porém, como evidenciado, a simplicidade desses modelos pode levar a um menor desempenho, a depender do problema trabalhado.

Em comparação ao modelo de rede neural desenvolvido por Ayodele e Cheng, para os mesmos dados utilizados neste trabalho, não foi possível obter coeficientes de determinação tão altos, porém os modelos desenvolvidos aqui apresentam uma maior acurácia na previsão das variáveis de taxa de produção do H<sub>2</sub> e de CO, apresentando certo nível de melhoria apesar de apresentar uma maior dispersão.

## REFERÊNCIAS

- A VOZ DA INDÚSTRIA. **Caminho até a Indústria 4.0: os destaques das revoluções industriais.** 2020. Disponível em: <https://avozdaindustria.com.br/industria-40-totvs/caminho-ate-industria-40-os-destaques-das-revolucoes-industriais>. Acesso em: 17 set. 2022.
- AL AANI, Saif; BONNY, Talal; HASAN, Shadi W.; HILAL, Nidal. Can machine language and artificial intelligence revolutionize process automation for water treatment and desalination? **Desalination**, vol. 458, no. February, p. 84–96, 2019.
- AYODELE, Bamidele V.; CHENG, Chin Kui. Modelling and optimization of syngas production from methane dry reforming over ceria-supported cobalt catalyst using artificial neural networks and Box-Behnken design. **Journal of Industrial and Engineering Chemistry**, v. 32, p. 246–258, 2015.
- BAHADAR, Ali; KANTHASAMY, Ramesh; SAIT, Hani Hussain; ZWAWI, Mohammed; ALGARNI, Mohammed; AYODELE, Bamidele Victor; CHENG, Chin Kui; WEI, Lim Jun. Elucidating the effect of process parameters on the production of hydrogen-rich syngas by biomass and coal Co-gasification techniques: A multi-criteria modeling approach. **Chemosphere**, vol. 287, no. P1, p. 132052, 2022.
- CHEN, Shuyue; ZAFFRAN, Jeremie; YANG, Bo. Dry reforming of methane over the cobalt catalyst: Theoretical insights into the reaction kinetics and mechanism for catalyst deactivation. **Applied Catalysis B: Environmental**, v. 270, n. January, 2020.
- COELHO, Matheus. Fundamentos de Redes Neurais - Laboratório iMobilis. **Universidade Federal de Ouro Preto, 2017**. Disponível em: <http://www2.decom.ufop.br/imobilis/fundamentos-de-redes-neurais/>. Acesso em 09 jul. 2022.
- DING, Bin; QIAN, Huimin; ZHOU, Jun. Activation functions and their characteristics in deep neural networks. Jun. 2018, [S.1.]: **Chinese Control and Decision Conference**, jun. 2018. p. 1836–1841.
- DOBBELAERE, Maarten R.; PLEHIERS, Pieter P.; VAN DE VIJVER, Ruben; STEVENS, Christian V.; VAN GEEM, Kevin M. Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats. **Engineering**, vol. 7, no. 9, p. 1201–1211, 2021.
- ELMAZ, Furkan; YÜCEL, Özgün; MUTLU, Ali Yener. Predictive Modeling of the Syngas Production from Methane Dry Reforming over Cobalt Catalyst with Statistical and Machine Learning Based Approaches. **International Journal of Advances in Engineering and Pure Sciences**, p. 8–14, 2020.
- GEORGE, Joel; ARUN, P.; MURALEEDHARAN, C. Assessment of producer gas composition in air gasification of biomass using artificial neural network model. **International Journal of Hydrogen Energy**, v. 43, n. 20, p. 9558–9568, 2018.
- KAN, Xiang; CHEN, Xiaoping; SHEN, Ye; LAPKIN, Alexei A.; KRAFT, Markus; WANG, Chi Hwa. Box-Behnken design based CO<sub>2</sub> co-gasification of horticultural waste and sewage sludge with addition of ash from waste as catalyst. **Applied Energy**, vol. 242, no. February, p. 1549–1561, 2019.

LEE, Jay H.; SHIN, Joohyun; REALFF, Matthew J. Machine learning: Overview of the recent progresses and implications for the process systems engineering field. **Computers & Chemical Engineering**, v. 114, p. 111–121, jun. 2018.

MANSUR, Fatin Zafirah; FAIZAL, Che Ku Mohammad; MONIR, Minhaj Uddin; SAMAD, Nur Asma Fazli Abdul; ATNAW, Samson Mekbib; SULAIMAN, Shaharin Anwar. Co-gasification between coal/sawdust and coal/wood pellet: A parametric study using response surface methodology. **International Journal of Hydrogen Energy**, vol. 45, no. 32, p. 15963–15976, 2020.

MELONI, Eugenio; MARTINO, Marco; PALMA, Vincenzo. A short review on Ni based catalysts and related engineering issues for methane steam reforming. **Catalysts**, v. 10, n. 3, 2020.

PANDEY, Daya Shankar; DAS, Saptarshi; PAN, Indranil; LEAHY, James J.; KWAPINSKI, Witold. Artificial neural network based modelling approach for municipal solid waste gasification in a fluidized bed reactor. **Waste Management**, vol. 58, p. 202–213, 2016. <https://doi.org/10.1016/j.wasman.2016.08.023>.

PERASSO, Valeria. **O que é a 4ª revolução industrial - e como ela deve afetar nossas vidas**. BBC News Brasil, 2016. Disponível em: <https://www.bbc.com/portuguese/geral-37658309>. Acesso em: 20 set. 2022.

PETSAGKOURAKIS, P.; SANDOVAL, I. O.; BRADFORD, E.; ZHANG, D.; DEL RIO-CHANONA, E. A. Reinforcement learning for batch bioprocess optimization. **Computers and Chemical Engineering**, vol. 133, p. 106649, 2020.

PORTAL DA INDÚSTRIA. **Indústria 4.0: Entenda seus conceitos e fundamentos**. Disponível em: <http://www.portaldaindustria.com.br/industria-de-a-z/industria-4-0/>. Acesso em: 19 set. 2022.

SALES, Deivson Cesar Silva. **Desenvolvimento da tecnologia GTL de produção de metanol a partir do gás de síntese**. 2013. 181f. Tese (Doutorado em Engenharia Química), Programa de Pós-Graduação em Engenharia Química (PPEQ), Universidade Federal de Pernambuco, Recife, PE., 2013.

SALHI, N.; BOULAHOUACHE, A.; PETIT, C.; KIENNEMANN, A.; RABIA, C. Steam reforming of methane to syngas over NiAl<sub>2</sub>O<sub>4</sub> spinel catalysts. **International Journal of Hydrogen Energy**, vol. 36, no. 17, p. 11433–11439, 2011.

SARKAR, Dipanjan; BALI, Raghav; SHARMA, Tushar. **Practical Machine Learning with Python**. Berkeley, CA: Apress, vol. 1, 2017. 530 p.

SCHWEIDTMANN, Artur M.; ESCHE, Erik; FISCHER, Asja; KLOFT, Marius; REPKE, Jens Uwe; SAGER, Sebastian; MITSOS, Alexander. Machine Learning in Chemical Engineering: A Perspective. **Chemie-Ingenieur-Technik**, vol. 93, no. 12, p. 2029–2039, 2021.

USMAN, Muhammad; WAN DAUD, W. M.A.; ABBAS, Hazzim F. Dry reforming of methane: Influence of process parameters - A review. **Renewable and Sustainable Energy Reviews**, v. 45, p. 710–744, 2015.

VALENTE, Emanuele de Oliveira; VALENTE, Gerson de Freitas Silva. Simulação de redes neurais artificiais para estimativa de volume de madeira florestal a partir do DAP / Simulation of artificial neural networks for estimation of forest wood volume from DAP. **Brazilian Journal of Animal and Environmental Research**, v. 4, n. 3, p. 3748–3757, 2021.

VENKATASUBRAMANIAN, Venkat. The promise of artificial intelligence in chemical engineering: Is it here, finally? **AIChe Journal**, v. 65, n. 2, p. 466–478, 2019.

WEN, Hung Ta; LU, Jau Huai; PHUC, Mai Xuan. Applying artificial intelligence to predict the composition of syngas using rice husks: A comparison of artificial neural networks and gradient boosting regression. **Energies**, v. 14, n. 10, p. 1–18, 2021.

YANG, Li; SHAMI, Abdallah. On hyperparameter optimization of machine learning algorithms: Theory and practice. **Neurocomputing**, v. 415, p. 295–316, 2020.

YAO, Ma; WANG, Huangang; XU, Wenli. Batch process monitoring based on functional data analysis and support vector data description. **Journal of Process Control**, v. 24, n. 7, p. 1085–1097, 2014.

YE, Jiansen. Artificial neural network modeling of methanol production from syngas. **Petroleum Science and Technology**, v. 37, n. 6, p. 629–632, 2019.