

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA QUÍMICA

GUILHERME HENRIQUE SAGA GOMES

APLICAÇÃO DE *MACHINE LEARNING* PARA O ESTUDO DE COMPORTAMENTO
DE UMA PLANTA DE PRODUÇÃO DE ESTIRENO

TRABALHO DE CONCLUSÃO DE CURSO

UBERLÂNDIA – MG
JUNHO 2021

GUILHERME HENRIQUE SAGA GOMES

APLICAÇÃO DE *MACHINE LEARNING* PARA O ESTUDO DE COMPORTAMENTO
DE UMA PLANTA DE PRODUÇÃO DE ESTIRENO

Trabalho de Conclusão de Curso
apresentado ao Curso de Engenharia
Química, da Universidade Federal de
Uberlândia, para obtenção do grau de
Engenheiro Químico.

Orientador: Prof. Sérgio Mauro da Silva Neiro

UBERLÂNDIA – MG

JUNHO 2021

RESUMO

A Inteligência Artificial (IA) e seus algoritmos de *Machine Learning (ML)* estão ganhando destaque em diversas áreas nos últimos anos, e as indústrias não ficam de fora disso. Aliás, o setor representado por elas tende a ganhar muito com estes algoritmos de IA, visto que a base deles é o recebimento de uma quantidade massiva de dados, o que há em grande quantidade nas indústrias, para o treinamento e aprendizado de processos. Dessa forma, a partir do uso desses algoritmos, empresas e indústrias conseguem realizar previsões de produção e financeiras, otimizações, entre outras. Por conta disso, há a possibilidade de um ganho significativo por parte dessas companhias, já que essas previsões auxiliam diretamente em melhorias nas tomadas de decisão baseadas nos dados gerados pelo próprio processo. Levando em consideração este contexto, o objetivo deste trabalho é a aplicação de dois algoritmos de ML com aprendizado supervisionado: a Rede Neural e a Regressão Linear Múltipla. Tais algoritmos foram utilizados em uma planta digital de Estireno, de forma a entender como alguns parâmetros do seu processo influenciam o produto final esperado. Assim, a partir do treinamento de dados, é possível realizar a previsão do Estireno por meio de parâmetros pré-definidos. Como consequência, há a possibilidade da realização de otimizações no processo, ao passo que se sabe quais os parâmetros mais influenciam no produto final desejado. Por fim, a avaliação do melhor modelo é realizada a partir de medidas estatísticas específicas.

Palavras-Chaves: Aprendizado Supervisionado, Estireno, *Machine Learning*, Otimização, Rede Neural, Regressão Linear Múltipla.

ABSTRACT

Artificial Intelligence (AI) and its Machine Learning (ML) algorithms has been calling attention in what it is concerned to several areas lately, and industries are not left out of this. In fact, the sector represented by them tends to gain a lot with those algorithms, since they are based on a massive amount of data, which exists in great amount in the industries, which can help with the training and learning processes. Thus, using those algorithms, companies and industries are able to make production and financial predictions, optimizations, among others. Because of that, there is the possibility of a significant gain by some of these companies, since those predictions directly assist in improvements in what it is concerned with making decisions based on the data generated by the process itself. Taking into consideration this context, the objective of this work is the application of two ML algorithms with supervised learning: Neural Network and Multiple Linear Regression. Such algorithms were used in a digital blueprint of Styrene, in order to understand how some parameters of its process influence the expected final product. From the data training, it is possible to perform the prediction of Styrene by pre-defined parameters. As a consequence, there is the possibility of optimizing the process, while knowing which parameters most influence the desired final product. Concluding, the evaluation of the best model is carried out using specific statistical measures.

Keywords: Supervised Learning, Styrene, Machine Learning, Optimization, Artificial neural network, Multiple Linear Regression.

SUMÁRIO

RESUMO	3
ABSTRACT	4
SUMÁRIO	5
LISTA DE FIGURAS	6
LISTA DE TABELAS	7
1. INTRODUÇÃO	7
2. REVISÃO BIBLIOGRÁFICA	9
3. METODOLOGIA	14
4. RESULTADOS E DISCUSSÕES	25
4.1. GERAÇÃO DE NOVOS DADOS	25
4.2. DIVISÃO DOS DADOS EM TREINO E TESTE	30
4.3. PREDIÇÃO DOS MODELOS NO <i>ORANGE</i>	35
4.4. AVALIAÇÃO ESTATÍSTICA DO MELHOR MODELO.....	40
4.5. VISUALIZAÇÃO GRÁFICA DE CORRELAÇÕES DE VARIÁVEIS DO MODELO	42
5. CONCLUSÃO	45
6. REFERÊNCIAS	46

LISTA DE FIGURAS

Figura 1 – Três tipos de aprendizagem	11
Figura 2 – Exemplo de Aprendizagem Supervisionada	12
Figura 3 – Diferenciando cão e gato usando algoritmo de rede neural.....	13
Figura 4 – Planta digital completa do estireno.....	16
Figura 5 – Parâmetros da cinética da reação	17
Figura 6 – Parâmetros (Reactions - Overall) do PFR-100	18
Figura 7 – Parâmetros (Rating - Sizing) do PFR-100	18
Figura 8 – Parâmetros (Reactions - Overall) do PFR-101	19
Figura 9 – Parâmetros (Rating - Sizing) do PFR-101	19
Figura 10 – Parâmetros (Design - Connections) da T-401	20
Figura 11 – Parâmetros (Design - Monitor) da T-401	20
Figura 12 – Parâmetros (Design - Connections) da T-402	21
Figura 13 – Parâmetros (Design - Monitor) da T-402	21
Figura 14 – Estrutura utilizada no simulador Orange.....	35
Figura 15 – Seleção dos dados de treino	36
Figura 16 – Seleção dos dados de teste	37
Figura 17 – Seleção dos parâmetros da Rede Neural.....	38
Figura 18 – Seleção dos parâmetros da Regressão Linear.....	38
Figura 19 – Gráfico vazão 6 (kmol/h) x vazão 20 (produto) (kmol/h)	43
Figura 20 – Gráfico temp 7 (alim reator 1) (°C) x vazão 20 (produto) (kmol/h)	44

LISTA DE TABELAS

Tabela 1 – Dados da simulação inicial	22
Tabela 2 – Novos dados gerados	25
Tabela 3 – Dados utilizados para o treinamento do modelo	30
Tabela 4 – Dados utilizados para o teste do modelo	34
Tabela 5 – Predições da Rede Neural e Regressão Linear em comparação com os dados de teste.....	39
Tabela 6 – Medidas obtidas em Predição (5) da Figura 13.....	41

1. INTRODUÇÃO

A quantidade de dados cresce a cada dia, assim como a possibilidade para sua utilização, segundo Demchenko, Grosso, De Laat, & Membrey (2013). Conseqüentemente, a Inteligência Artificial (IA) que é definida como o estudo de agentes que recebem percepções do ambiente e executam ações (RUSSEL e NORVIG, 2013) têm ganhado maior destaque devido ao crescente volume de dados em empresas e indústrias. Tendo em vista que os algoritmos de IA têm como base o recebimento de dados, a fim de que possam aprender com as informações recebidas, tais algoritmos conseguem auxiliar em novas tomadas de decisões de acordo com seu aprendizado contínuo.

Nesse sentido, este trabalho teve como base o uso da IA a partir de algoritmos de *Machine Learning* (ML) - Rede Neural e Regressão Linear Múltipla - com Aprendizado Supervisionado, a fim de realizar o estudo do comportamento de uma planta digital de produção de estireno. Para isso, foi construída uma base de dados para que parte dela fosse utilizada como treinamento dos algoritmos e o restante dos dados como teste, com o intuito de avaliar se os modelos seriam capazes de prever novas entradas de dados.

Dessa forma, é possível aprimorar os estudos de otimização na área, visto que os algoritmos de ML têm um grande potencial para resolução de problemas (RUSSEL e NORVIG, 2013). Isso se dá pela dinâmica do funcionamento do mecanismo da IA, já que ela consegue aprender como o processo funciona a partir do treinamento com dados do próprio processo. Em outras palavras, ao passo que novos dados são processados e alimentam a IA, mais ela aprende e pode realizar novas previsões, a fim de otimizar os resultados do objeto em questão.

Sendo assim, o objetivo deste trabalho é conseguir unificar o estudo da IA a partir dos algoritmos de ML Rede Neural Artificial e Regressão Linear Múltipla com a otimização de uma planta digital de Estireno. Com efeito, trazer uma abordagem que vem crescendo

nos últimos anos sobre como a indústria 4.0 com suas ramificações pode ser benéfica, a fim de gerar otimizações, se bem implementada nas indústrias e empresas.

2. REVISÃO BIBLIOGRÁFICA

A indústria 4.0 vem ganhando muito destaque no cenário mundial nos últimos anos, isso porque ela tem como objetivo proporcionar eficiências financeiras significativas e bens ou serviços de melhor qualidade para as empresas. Para isso, utiliza-se tecnologias de ponta junto à internet, com o intuito de que os sistemas de produção sejam mais flexíveis e colaborativos (BAHRIN; OTHMAN; AZLI; TALIB, 2016). Nesse contexto, a Inteligência Artificial ganha evidência, devido ao seu potencial de realizar tarefas complexas.

Dessa forma, tais tarefas complexas englobam a utilização de dispositivos ou métodos computacionais análogos à capacidade de raciocínio do ser humano, a fim de resolver problemas de maneira mais eficiente possível. Com efeito, por meio da Inteligência Artificial, ainda há a possibilidade de fornecer sugestões para o auxílio nas tomadas de decisões (SACOMANO; GONÇALVES et al., 2018).

Nesse sentido, estes ambientes complexos com potencial para utilização da Inteligência Artificial são cada vez mais comuns hoje, visto que lidam com uma quantidade muito densa de dados e informações. Esses, por sua vez, devem ser analisados por *softwares* com grande embasamento estatístico e geração de correlações entre os parâmetros dos negócios com potencial de gerar otimizações e melhores adequações para as empresas. Para este volume de dados utilizados na Inteligência Artificial é dado o nome de *Big Data* (SACOMANO; GONÇALVES et al., 2018).

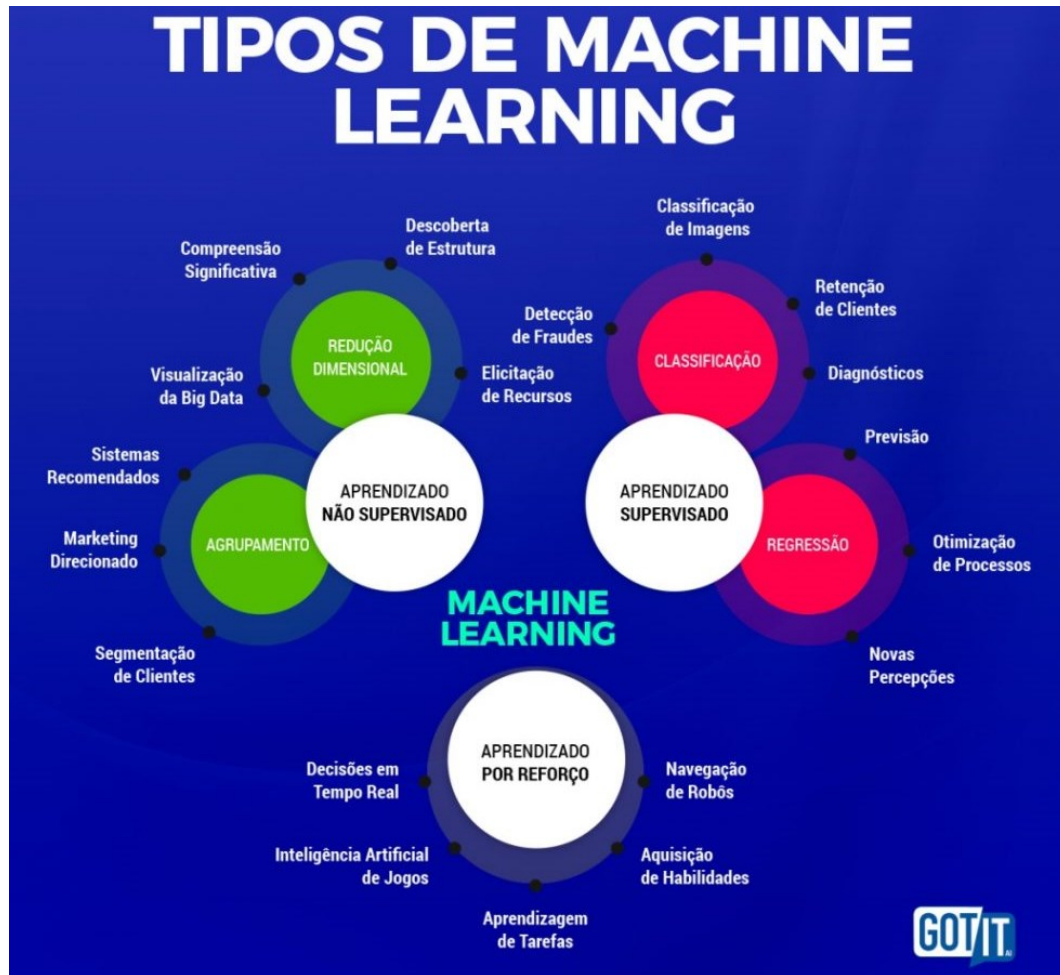
Sendo assim, é possível entender o contexto da indústria 4.0 hoje no mercado e sua importância cada vez maior para os negócios, já que independentemente do ramo de atuação, há um gigantesco volume de dados gerados. Os quais, se bem analisados e tratados com *softwares* com algoritmos de Inteligência Artificial, podem resultar em otimizações financeiras e de tempo.

Além disso, há outro conceito importante a ser trabalhado junto à Inteligência Artificial que é o do *Machine Learning*. Este campo é focado em como construir programas de computador que melhorem automaticamente com as experiências, isto é, de acordo com que são alimentados por meio de dados (MITCHELL, 1997).

Vale ressaltar que os dados analisados não são necessariamente números, podem ser também, por exemplo, imagens e tudo que esta tecnologia conseguir identificar. Outro ponto a ser destacado é que softwares tradicionais têm uma estrutura diferente em relação a um algoritmo de *Machine Learning*. Enquanto nos primeiros é criado um conjunto de regras, com o objetivo de gerar respostas a partir dos dados introduzidos, no segundo, o próprio sistema cria regras, a partir dos dados introduzidos para a análise ou dos resultados que são esperados. Outrossim, o algoritmo de *Machine Learning* tem a vantagem de melhorar automaticamente e gradualmente de acordo com o número de experiência em que ele é colocado para treinar. Essa melhoria se dá, pois o algoritmo tem a oportunidade de aprender mais, de acordo que é fornecido novos dados para ele. Assim, com mais aprendizado, maior a chance do algoritmo ser aperfeiçoado (HURWITZ; KIRSCH, 2018).

Ainda dentro do *Machine Learning*, há uma ramificação importante a ser apresentada, a qual se trata dos tipos de aprendizado que o algoritmo poderá ter: Aprendizado Supervisionado, Aprendizado Não Supervisionado e Aprendizado por Reforço (HONDA; FACURE et al., 2017), os quais são apresentadas na Figura 1.

Figura 1 – Três tipos de aprendizagem

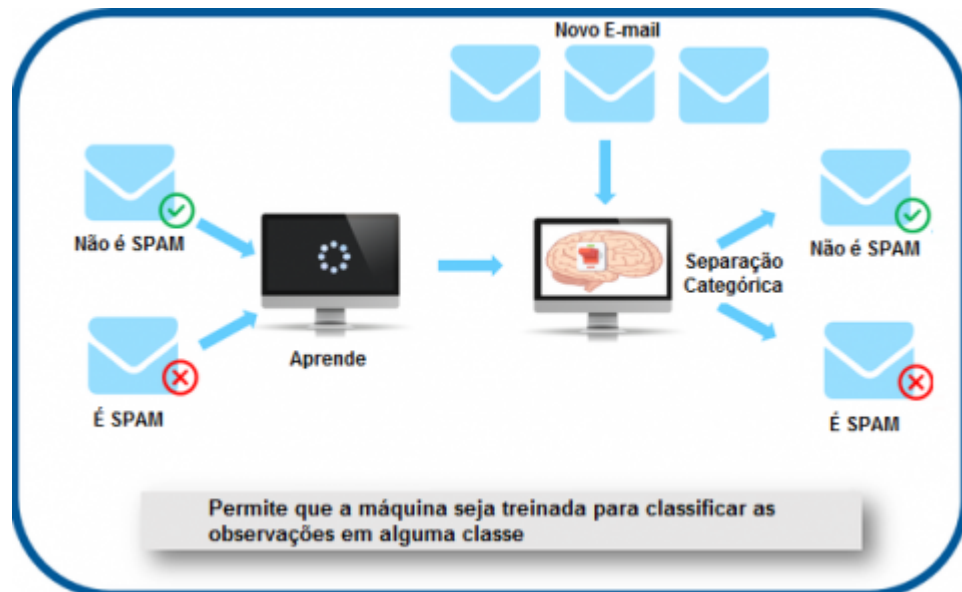


Fonte: Got It AI, 2020¹.

Neste trabalho o enfoque é no Aprendizado Supervisionado, no qual se tenta prever uma variável dependente a partir de uma lista de variáveis independentes. A característica básica destes sistemas é que os dados utilizados para o treino contêm a resposta desejada, ou seja, contêm a variável dependente resultante das independentes observadas (HONDA; FACURE et al., 2017). Um exemplo disso é a detecção de e-mails de SPAM, o qual é apresentado na Figura 2. Dentre as técnicas mais conhecidas para resolver os problemas de Aprendizado Supervisionado pode-se citar a Regressão Linear e a Rede Neural Artificial (RNA).

¹ Disponível em: <<https://www.programaria.org/quais-sao-os-tipos-de-aplicacoes-de-inteligencia-artificial-mais-comuns/>>. Acesso em: 04 jun. 2021.

Figura 2 – Exemplo de Aprendizagem Supervisionada



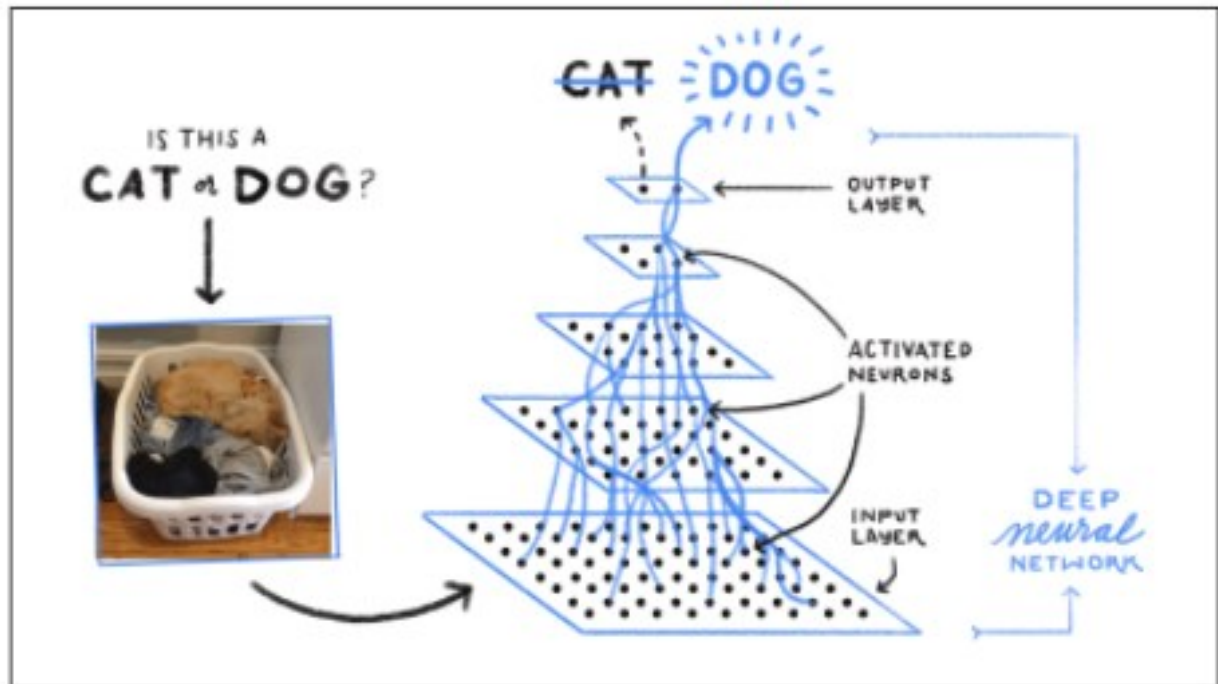
Fonte: Ayush Pant, 2019².

Primeiramente, quanto às Redes Neurais Artificiais, elas são sistemas de *Machine Learning* dentro do *Deep Learning* - o qual possui métodos de natureza dinâmica que objetivam a concretização de sistemas preditivos, de modo a mostrar capacidade contínua de melhora e adaptação às mudanças de padrões, onde um modelo não é adaptado, mas sim, treinado. Assim, a RNA funciona como os neurônios do cérebro humano, sendo que por meio de algoritmos, elas podem reconhecer padrões escondidos e correlações em dados brutos, agrupá-los e classificá-los. Além disso, com o tempo, conseguem aprender mais e melhor continuamente (ALVES, 2020).

Ainda em relação a RNA, o objetivo deste trabalho é utilizar o neurônio do tipo Perceptron Multicamadas treinadas com o algoritmo *Backpropagation*. Este modelo é composto por várias unidades de processamento, as quais, geralmente, são conectadas por canais de comunicação que estão associados a determinados pesos (GRÜBLER, 2018). Um exemplo disso é apresentado na Figura 3:

² Disponível em: <<https://towardsdatascience.com/introduction-to-machine-learning-for-beginners-ee6024fdb08>>. Acesso em: 12 mai. 2021.

Figura 3 – Diferenciando cão e gato usando algoritmo de rede neural



Fonte: site Selman Design³.

Por outro lado, o outro modelo de *Machine Learning* supervisionado que será tratado neste trabalho é o da Regressão Linear Múltipla. Este, por sua vez, trata de modelos com, no mínimo, três variáveis, sendo uma delas dependente e as outras duas explanatórias (independentes). Isto acontece em contraposição aos modelos de Regressão Linear Simples, já que, em cenários reais, não raro, as variáveis dependentes encontradas nos sistemas dependem de mais de uma variável independente (GUJARATI; PORTER, 2011).

Por fim, tem-se como objetivo neste trabalho a aplicação de *Machine Learning* para o estudo de comportamento de uma planta de produção de estireno. Sendo assim, foi tomado como base uma planta digital de Estireno - usado na fabricação de embalagens plásticas e materiais descartáveis, borracha sintética, em isolamento térmico e resinas para fabricação de barcos, chuveiros, acessórios para automóveis e muitos outros produtos (CETESB, 2012).

³ Disponível em: < <https://selmandesign.com/qa-on-machine-learning/>>. Acesso em: 12 mai. 2021.

3. METODOLOGIA

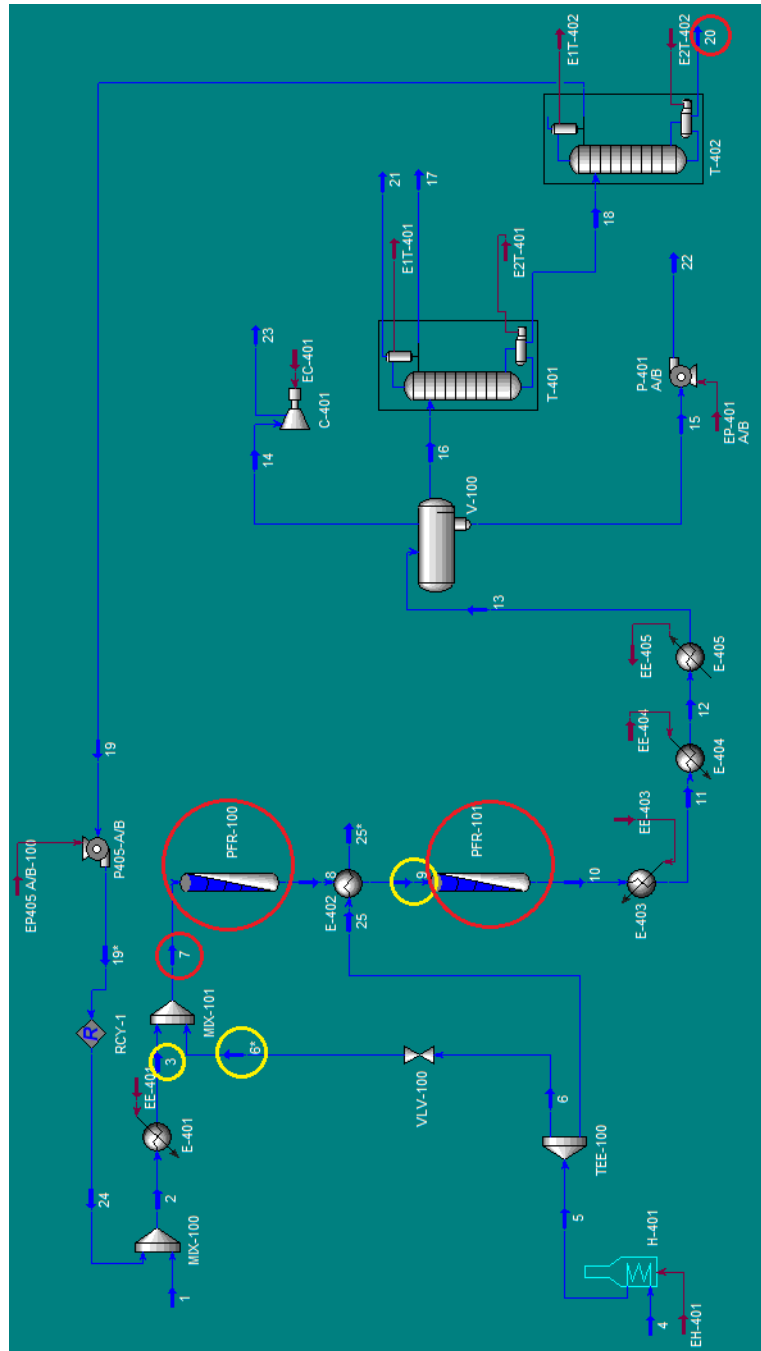
O trabalho tem como objetivo utilizar técnicas de *Machine Learning*, a fim de estabelecer previsões de comportamento da produção de uma planta de Estireno. Tendo em vista que não temos dados reais dessa planta, foi utilizado o simulador de processos comerciais *UniSim - software* para simulação e otimização de processos no qual os usuários podem melhorar a lucratividade da fábrica e maximizar o retorno de seus investimentos em simulação (HONEYWELL, 2021). Dessa forma, a partir da planta digital desenvolvida, foi possível fazer uma geração de dados, com o intuito de utilizá-los posteriormente no algoritmo supervisionado de *Machine Learning*.

Nesse sentido, a simulação desenvolvida neste trabalho utiliza os parâmetros apresentados no projeto preliminar de um processo de produção de Estireno desenvolvido por Turton em *Analysis, Synthesis, and Design of Chemical Processes* (TURTON; BAILIE et al., 2012).

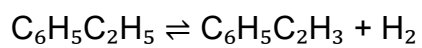
Assim sendo, acontece o seguinte processo que é apresentado na Figura 4: a alimentação é feita com uma corrente de Etilbenzeno (1) misturada a uma corrente de reciclo (24) rica em Etilbenzeno, posteriormente esta corrente contendo a mistura (2) é aquecida e misturada com uma corrente de vapor superaquecido de alta temperatura (6'). Vale destacar que o vapor é um inerte na reação, que direciona o equilíbrio mostrado na Equação (01) para a direita, reduzindo as concentrações de todos os componentes. Como a formação de estireno é altamente endotérmica, o vapor superaquecido também fornece energia para impulsionar a reação. Feito isso, a corrente resultante (7) alimenta dois reatores tubulares em série (PFR-100 e PFR-101). O produto resultante na corrente (10) é resfriado até alimentar um vaso trifásico (V-100), do qual uma corrente é rica em Hidrogênio, uma em Água e a outra nos líquidos orgânicos (Benzeno, Etilbenzeno, Estireno, Hidrogênio e Tolueno). Assim, a corrente rica em Hidrogênio (23) é posteriormente purificada como fonte de Hidrogênio em outras partes da planta. Já a rica nos líquidos orgânicos é destilada duas vezes (T-401 e T-402), sendo que a partir da primeira destilação, são geradas as correntes (17) e (21) com uma pequena quantidade

de compostos orgânicos e a corrente (18) com o Estireno e Etilbenzeno. Por outro lado, a partir da segunda destilação (T-402), alimentada pela corrente (18), obtém-se uma corrente com o Estireno (20), o produto desejado, e outra com Etilbenzeno que dará origem a corrente de reciclo (24). É importante destacar que o Estireno pode polimerizar espontaneamente a temperaturas mais altas. Dessa forma, desde que sua temperatura seja mantida abaixo de 125 °C, não há problema de polimerização espontânea (TURTON; SHAEIWITZ et al., 2018). Por conta disso, se o vaso de flash (V-100) e as colunas (T-401 e T-402) fossem operadas a pressão ambiente, temperaturas acima de 125 °C seriam alcançadas no sistema de separação. Sendo assim, o sistema de separação é projetado para operar com vácuo.

Figura 4 – Planta digital completa do estireno



Fonte: *print screen* da planta digital completa do estireno no simulador *UniSim*



Etilbenzeno Estireno Hidrogênio

(01)

Dessa forma, a partir dessas configurações da planta, obteve-se a primeira simulação para uma produção do Estireno. Vale ressaltar que alguns valores foram ajustados junto às especificações dos equipamentos a fim de conseguir a simulação do processo.

Além disso, é importante destacar outras configurações que foram utilizados no *UniSim*. Em relação aos reatores, foram utilizados reatores tubulares (*PFR*) com a reação (01). Posteriormente, são apresentadas as especificações dos reatores e torres utilizados:

Foram utilizados os parâmetros apresentados na Figura 4 na cinética da reação:

Figura 5 – Parâmetros da cinética da reação

The screenshot shows a software window titled "Simple Rate Reaction: Rxn". It contains the following elements:

- Forward Reaction:** A table with three rows: A (7,9e+05), E (9,1e+04), and B (<empty>).
- Reverse Reaction:** A table with four rows: A' (1,6e+01), B' (-1,5e+04), C' (<empty>), and D' (<empty>).
- Equation Help:** A text box containing the following equations:

$$r = k * \{f(\text{Basis}) - f(\text{Basis}) / K'\}$$

$$k = A * \exp \{ -E / RT \} * T^{\beta}$$

$$\ln (K') = A' + B'/T + C' \ln(T) + D' * T$$
 Below these equations, it states "T in Kelvin".
- Parameters Tab:** At the bottom, there are tabs for "Stoichiometry", "Basis", and "Parameters" (which is selected). Below the tabs are a "Delete" button, a "Name" field containing "Rxn", and a green "Ready" button.

Fonte: *print screen* dos parâmetros da cinética da reação no simulador *UniSim*

Para o PFR-100 foram utilizados os parâmetros apresentados nas Figuras 5 e 6:

Figura 6 – Parâmetros (Reactions - Overall) do PFR-100

Reactions

Overall

Details

Results

Reaction Info

Reaction Set: Global Rxn Set [View]

Initialize segment reactions from:

Current Previous Re-init

Integration Information

Number of Segments	20
Minimum Step Fraction	1,0e-06
Minimum Step Length	9,3e-06 m

Catalyst Data

Particle Diameter	0,02500 m
Particle Sphericity	1,000
Solid Density	2500,0 kg/m ³
Bulk Density	1287,5 kg/m ³
Solid Heat Capacity	1,000 kJ/kg-C

Design Reactions Rating Worksheet Performance Dynamics Cost

Delete OK Ignored

Fonte: print screen dos parâmetros (Reactions - Overall) do PFR-100 no simulador UniSim

Figura 7 – Parâmetros (Rating - Sizing) do PFR-100

Rating

Sizing

Nozzles

Heat Loss

Tube Dimensions

Total Volume	25,000 m ³
Length	9,260 m
Diameter	1,8540 m
Number of Tubes	1
Wall Thickness	0,0050 m

Tube Packing

Void Fraction	0,485
Void Volume	12,125 m ³

Catalyst Data

Particle Diameter	0,02500 m
Particle Sphericity	1,000
Solid Density	2500,0 kg/m ³
Bulk Density	1287,5 kg/m ³
Solid Heat Capacity	1,000 kJ/kg-C

Design Reactions Rating Worksheet Performance Dynamics Cost

Delete OK Ignored

Fonte: print screen dos parâmetros (Rating - Sizing) do PFR-100 no simulador UniSim

Para o PFR-101 foram utilizados os parâmetros apresentados nas Figuras 7 e 8:

Figura 8 – Parâmetros (Reactions - Overall) do PFR-101

The screenshot shows the 'Reactions - Overall' configuration window. The 'Reaction Info' section includes a dropdown for 'Global Rxn Set' and radio buttons for 'Current', 'Previous', and 'Re-int'. The 'Integration Information' section contains a table with the following values:

Number of Segments	20
Minimum Step Fraction	1,0e-06
Minimum Step Length	9,3e-06 m

The 'Catalyst Data' section contains a table with the following values:

Particle Diameter	0,00100 m
Particle Sphericity	1,000
Solid Density	2500,0 kg/m ³
Bulk Density	1287,5 kg/m ³
Solid Heat Capacity	1,000 kJ/kg-C

At the bottom, there are tabs for 'Design', 'Reactions', 'Rating', 'Worksheet', 'Performance', 'Dynamics', and 'Cost'. A green 'OK' button is visible, along with 'Delete' and 'Ignored' options.

Fonte: *print screen* dos parâmetros (*Reactions - Overall*) do PFR-101 no simulador *UniSim*

Figura 9 – Parâmetros (Rating - Sizing) do PFR-101

The screenshot shows the 'Rating - Sizing' configuration window. The 'Tube Dimensions' section contains a table with the following values:

Total Volume	25,000 m ³
Length	9,260 m
Diameter	1,8540 m
Number of Tubes	1
Wall Thickness	0,0050 m

The 'Tube Packing' section contains a table with the following values:

Void Fraction	0,485
Void Volume	12,125 m ³

The 'Catalyst Data' section contains a table with the following values:

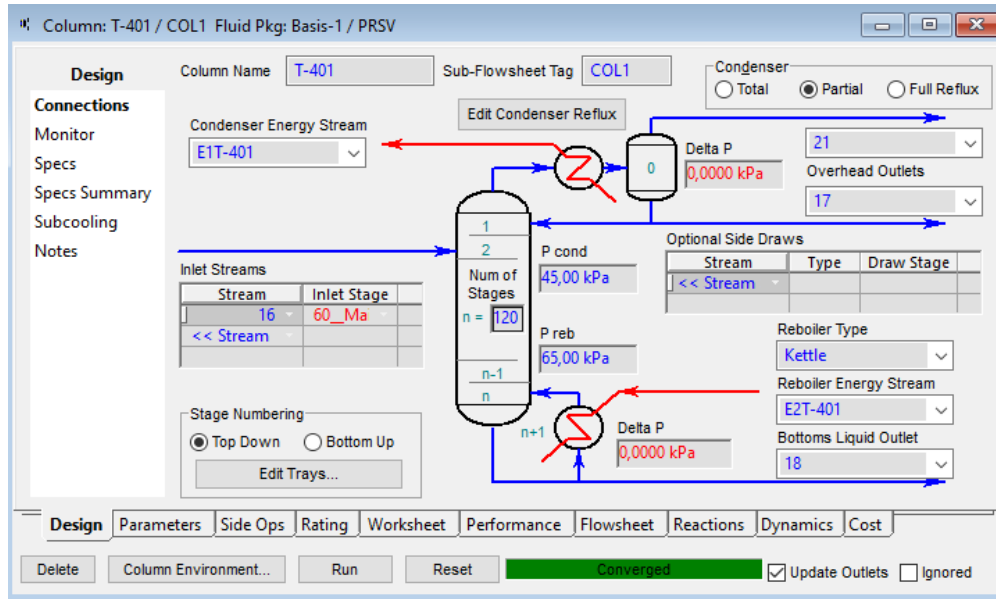
Particle Diameter	0,00100 m
Particle Sphericity	1,000
Solid Density	2500,0 kg/m ³
Bulk Density	1287,5 kg/m ³
Solid Heat Capacity	1,000 kJ/kg-C

At the bottom, there are tabs for 'Design', 'Reactions', 'Rating', 'Worksheet', 'Performance', 'Dynamics', and 'Cost'. A green 'OK' button is visible, along with 'Delete' and 'Ignored' options.

Fonte: *print screen* dos parâmetros (*Rating - Sizing*) do PFR-101 no simulador *UniSim*

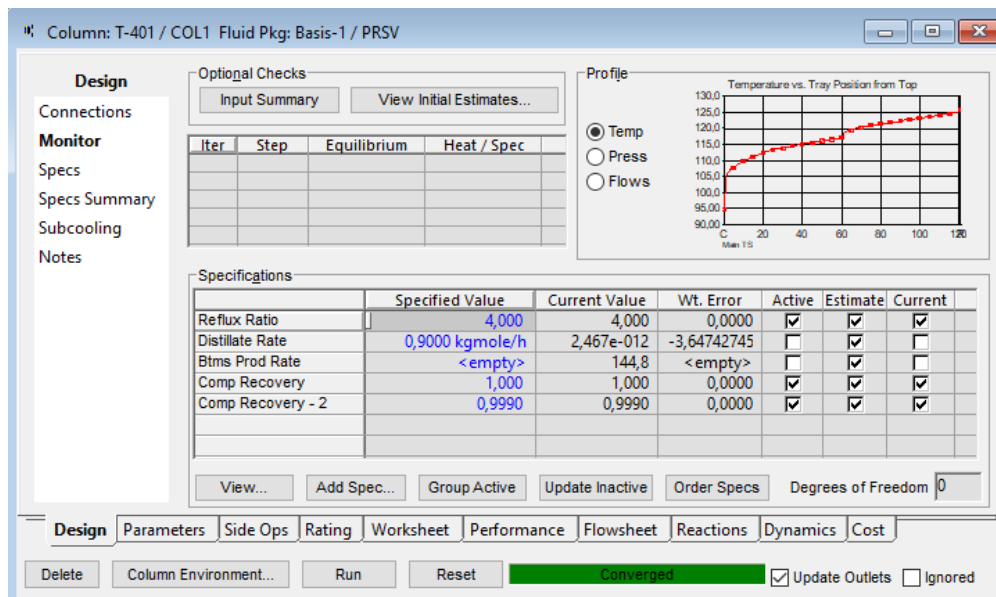
Para a T-401 foram utilizados os parâmetros apresentados nas Figuras 9 e 10:

Figura 10 – Parâmetros (Design - Connections) da T-401



Fonte: *print screen* dos parâmetros (Design - Connections) da T-401 no simulador UniSim

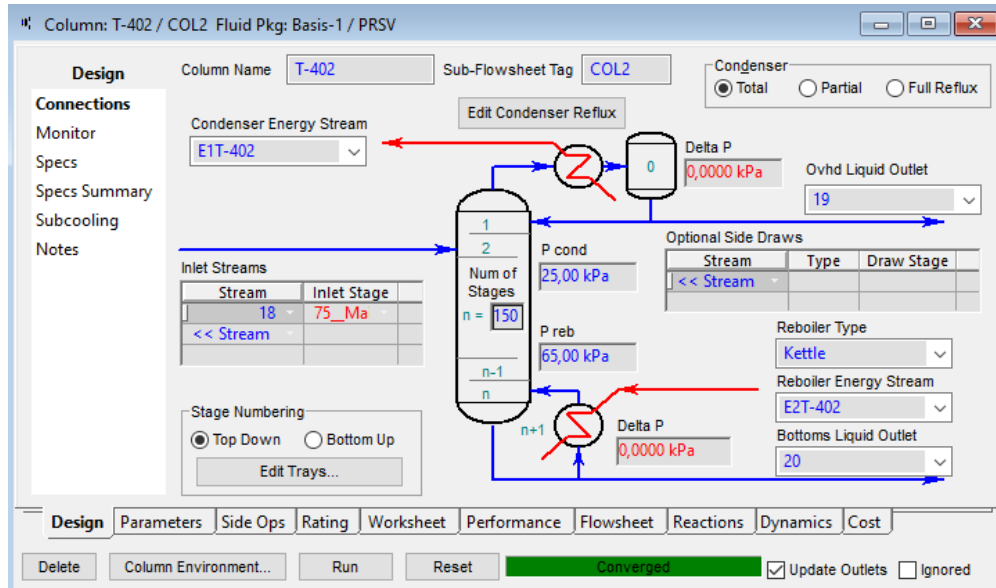
Figura 11 – Parâmetros (Design - Monitor) da T-401



Fonte: *print screen* dos parâmetros (Design - Monitor) da T-401 no simulador UniSim

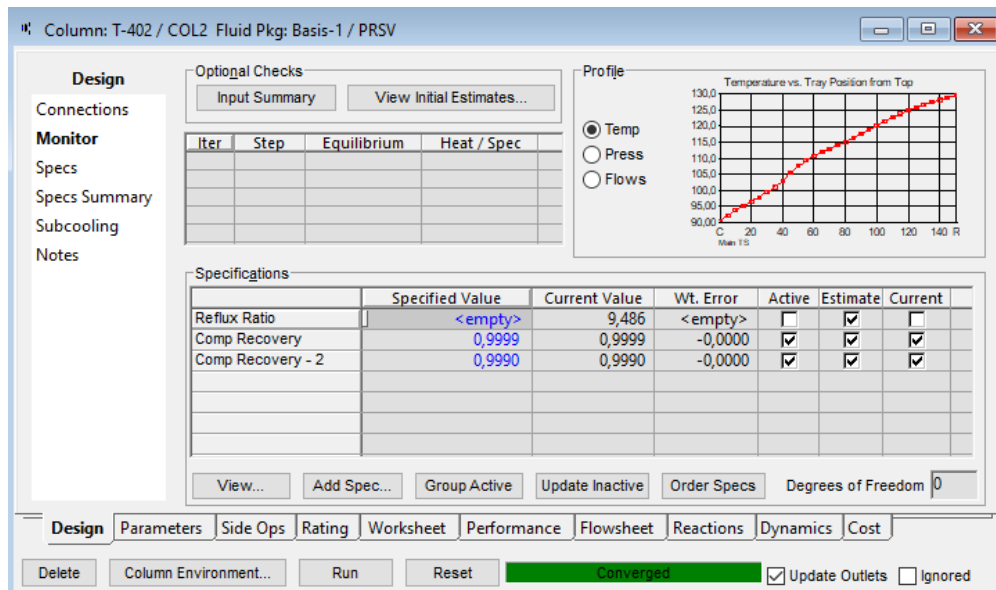
Para a T-402 foram utilizados os parâmetros apresentados nas Figuras 11 e 12:

Figura 12 – Parâmetros (Design - Connections) da T-402



Fonte: print screen dos parâmetros (Design - Connections) da T-402 no simulador UniSim

Figura 13 – Parâmetros (Design - Monitor) da T-402



Fonte: print screen dos parâmetros (Design - Monitor) da T-402 no simulador UniSim

Por fim, os dados da simulação são apresentados na tabela 1:

Tabela 1 – Dados da simulação inicial

Parâmetro	Corrente 1	Corrente 4	Corrente 5	Corrente 6	Corrente 25	Corrente 8	Corrente 9
Fração de vapor	0,0000	0,0000	1,0000	1,0000	1,0000	1,0000	1,0000
Temperatura (°C)	136,00	253,70	800,00	800,00	800,00	545,64	550,00
Pressão (kPa)	200,00	4.237,00	4.202,00	4.202,00	4.202,00	150,00	135,00
Fluxo molar (kmol/h)	123,42	4.016,30	4.016,30	3.000,00	1.016,30	3.277,07	3.277,07
Fluxo de massa (kg/h)	1.305,21	7.235,40	7.235,40	5.404,53	1.830,87	7.239,78	7.239,78
Fração molar de H ₂ O	0,0000	1,0000	1,0000	1,0000	1,0000	0,9155	0,9155
Fração molar de Etilbenzeno	0,9804	0,0000	0,0000	0,0000	0,0000	0,0202	0,0202
Fração molar de Estireno	0,0000	0,0000	0,0000	0,0000	0,0000	0,0316	0,0316
Fração molar de Hidrogênio	0,0000	0,0000	0,0000	0,0000	0,0000	0,0316	0,0316
Fração molar de Benzeno	0,0098	0,0000	0,0000	0,0000	0,0000	0,0004	0,0004
Fração molar de Tolueno	0,0098	0,0000	0,0000	0,0000	0,0000	0,0008	0,0008
Fração molar de Metano	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Fração molar de Etileno	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Continuação da Tabela 1:

Parâmetro	Corrente 25*	Corrente 2	Corrente 3	Corrente 24	Corrente 7	Corrente 10	Corrente 11
Fração de vapor	1,0000	0,0000	0,8832	0,0000	1,0000	1,0000	1,0000
Temperatura (°C)	783,04	123,25	155,80	90,34	621,32	545,90	267,00
Pressão (kPa)	4.202,00	200,00	170,00	200,00	170,00	125,00	110,00
Fluxo molar (kmol/h)	1.016,30	173,54	173,54	50,12	3.173,54	3.282,56	3.282,56
Fluxo de massa (kg/h)	1.830,87	1.835,22	1.835,22	5.300,12	7.239,75	7.239,79	7.239,79
Fração molar de H ₂ O	1,0000	0,0000	0,0000	0,0000	0,9453	0,9139	0,9139
Fração molar de Etilbenzeno	0,0000	0,9774	0,9774	0,9702	0,0534	0,0185	0,0185
Fração molar de Estireno	0,0000	0,0000	0,0000	0,0002	0,0000	0,0332	0,0332
Fração molar de Hidrogênio	0,0000	0,0000	0,0000	0,0000	0,0000	0,0332	0,0332
Fração molar de Benzeno	0,0000	0,0070	0,0070	0,0000	0,0004	0,0004	0,0004
Fração molar de Tolueno	0,0000	0,0155	0,0155	0,0296	0,0008	0,0008	0,0008
Fração molar de Metano	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Fração molar de Etileno	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Continuação da Tabela 1 na próxima página.

Com isso, a fim de obter mais dados para treinar o algoritmo, foram realizadas mais 114 simulações dentro do mesmo processo, com variações dos seguintes parâmetros: temperatura da corrente 3 (°C), pressão da corrente 3 (kPa), vazão da corrente 6 (kmol/h), temperatura da corrente 9 (°C), pressão da corrente 9 (kPa) (apresentados em amarelo na Figura 13). Feito isso, como consequência das variações, foram gerados outros dados de outras correntes e equipamentos da planta, os quais auxiliaram na alimentação dos algoritmos de Regressão Linear Múltipla e Rede Neural.

Vale ressaltar que estes parâmetros foram escolhidos levando em consideração que a concentração, a pressão e a temperatura têm influência na reação principal de equilíbrio, logo, no sistema de reação. Portanto, essas variações influenciam na produção final do Estireno e, assim, é possível analisar como cada uma delas contribui na formação do produto. Com isso, os parâmetros para teste captados com as variações foram: temperatura da corrente 7 (°C), pressão da corrente 7 (kPa), conversão do reator 1, conversão do reator 2 e vazão da corrente 20 (kmol/h) - produção do Estireno (apresentados em vermelho na Figura 13).

Sendo assim, após a geração dos dados, foi utilizado o *Orange*, *software* para aplicação de *Machine Learning* e visualização de dados (ORANGE, 2021), a fim de realizar a aplicação de duas técnicas de *Machine Learning* a partir de algoritmos supervisionados para o aprendizado de máquina e auxiliar na visualização de dados. Tal *software* foi utilizado levando em conta que ele tem uma boa interface gráfica e é bastante didático. Como parte dos modelos, os dados gerados foram divididos entre dados de treino e dados de teste, visto que isso é uma das etapas do processo de aprendizado de máquina. Por fim, seu uso teve como objetivo comparar os desempenhos de uma Regressão Linear Múltipla e uma Rede Neural com os dados obtidos nas simulações do *UniSim*.

4. RESULTADOS E DISCUSSÕES

4.1. GERAÇÃO DE NOVOS DADOS

Para a parte prática deste trabalho, primeiramente, foi construída a planta digital do Estireno no *UniSim*, baseada no modelo e nos dados apresentados em *Analysis, Synthesis, and Design of Chemical Processes* (TURTON; SHAEIWITZ et al., 2018). Feito isso, foram feitas mais 114 simulações, a fim de colher dados para o treino do algoritmo de *Machine Learning* que foi feito posteriormente na plataforma do *Orange*. Dessa forma, são apresentados na Tabela 2 a seguir os novos dados originados das simulações a partir da planta digital apresentada na Figura 13:

Tabela 2 – Novos dados gerados

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
155,80	170,00	3000	550,00	135,00	625,80	170,00	0,6230	0,0639	94,46
95,00	170,00	3000	550,00	135,00	560,80	170,00	0,4275	0,2474	93,74
105,00	170,00	3000	550,00	135,00	563,10	170,00	0,4337	0,2431	94,06
110,00	170,00	3000	550,00	135,00	564,30	170,00	0,4368	0,2409	94,21
115,00	170,00	3000	550,00	135,00	565,50	170,00	0,4399	0,2386	94,37
120,00	170,00	3000	550,00	135,00	567,50	170,00	0,4458	0,2352	94,33
130,00	170,00	3000	550,00	135,00	571,10	170,00	0,4568	0,2289	94,31
135,00	170,00	3000	550,00	135,00	572,30	170,00	0,4602	0,2262	94,47
140,00	170,00	3000	550,00	135,00	574,30	170,00	0,4661	0,2226	94,42
150,00	170,00	3000	550,00	135,00	577,90	170,00	0,4774	0,2150	94,37
155,00	170,00	3000	550,00	135,00	579,20	170,00	0,4809	0,2120	94,54
159,00	170,00	3000	550,00	135,00	626,50	170,00	0,6248	0,0613	94,67
160,00	170,00	3000	550,00	135,00	626,60	170,00	0,6253	0,0605	94,69

Continuação da Tabela 2 na próxima página.

Continuação da Tabela 2:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
161,00	170,00	3000	550,00	135,00	626,80	170,00	0,6258	0,0597	94,71
162,00	170,00	3000	550,00	135,00	627,20	170,00	0,6271	0,0580	94,63
162,00	165,00	3000	550,00	135,00	627,20	165,00	0,6272	0,0579	94,64
162,00	140,00	3000	550,00	135,00	627,20	140,00	0,6277	0,0570	94,66
162,00	130,00	3000	550,00	135,00	627,20	130,00	0,6281	0,0564	94,67
162,00	120,00	3000	550,00	135,00	627,20	120,00	0,6282	0,0561	94,68
162,00	110,00	3000	550,00	135,00	627,20	110,00	0,6283	0,0560	94,68
162,00	100,00	3000	550,00	135,00	627,20	100,00	0,6276	0,0571	94,65
162,00	90,00	3000	550,00	135,00	627,20	90,00	0,6277	0,0570	94,65
162,00	80,00	3000	550,00	135,00	627,20	80,00	0,6279	0,0567	94,66
162,00	70,00	1400	550,00	135,00	437,10	70,00	0,0870	0,2134	82,69
162,00	70,00	1600	550,00	135,00	467,00	70,00	0,1291	0,2184	86,45
162,00	70,00	1700	550,00	135,00	483,40	70,00	0,1573	0,2203	87,79
162,00	70,00	1800	550,00	135,00	499,30	70,00	0,1890	0,2212	89,01
162,00	70,00	1900	550,00	135,00	515,00	70,00	0,2244	0,2206	90,07
162,00	70,00	2100	550,00	135,00	543,90	70,00	0,3021	0,2144	91,73
162,00	70,00	2200	550,00	135,00	557,00	70,00	0,3430	0,2083	92,36
162,00	70,00	2300	550,00	135,00	568,90	70,00	0,3837	0,1997	92,88
162,00	70,00	2400	550,00	135,00	579,80	70,00	0,4235	0,1885	93,32
162,00	70,00	2500	550,00	135,00	589,90	70,00	0,4631	0,1740	93,58
162,00	70,00	2600	550,00	135,00	598,90	70,00	0,5008	0,1562	93,92
162,00	70,00	2700	550,00	135,00	607,00	70,00	0,5359	0,1358	94,16
162,00	70,00	2900	550,00	135,00	621,20	70,00	0,6004	0,0852	94,49
162,00	70,00	3000	550,00	135,00	626,70	70,00	0,6260	0,0585	94,96
162,00	70,00	3005	550,00	135,00	626,90	70,00	0,6269	0,0574	95,02
162,00	70,00	3005	555,00	135,00	630,20	70,00	0,6425	0,0748	94,82

Continuação da Tabela 2 na próxima página.

Continuação da Tabela 2:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
162,00	70,00	3005	560,00	135,00	632,60	70,00	0,6541	0,0967	94,97
162,00	70,00	3005	565,00	135,00	634,90	70,00	0,6651	0,1213	95,12
162,00	70,00	3005	580,00	135,00	640,40	70,00	0,6918	0,2037	95,68
162,00	70,00	3005	585,00	135,00	642,10	70,00	0,7001	0,2320	95,78
162,00	70,00	3005	590,00	135,00	643,70	70,00	0,7078	0,2617	95,84
162,00	70,00	3005	595,00	135,00	645,10	70,00	0,7148	0,2921	95,91
162,00	70,00	3005	600,00	135,00	646,40	70,00	0,7209	0,3215	95,94
162,00	70,00	3005	605,00	135,00	647,70	70,00	0,7268	0,3518	96,02
162,00	70,00	3005	610,00	135,00	648,80	70,00	0,7325	0,3827	96,08
162,00	70,00	3005	615,00	135,00	649,90	70,00	0,7377	0,4129	96,12
162,00	70,00	3005	620,00	135,00	650,90	70,00	0,7425	0,4428	96,17
162,00	70,00	3005	625,00	135,00	651,80	70,00	0,7469	0,4718	96,22
162,00	70,00	3005	630,00	135,00	652,70	70,00	0,7511	0,5023	96,28
162,00	70,00	3005	635,00	135,00	653,50	70,00	0,7554	0,5343	96,34
162,00	70,00	3005	635,00	120,00	653,50	70,00	0,7554	0,5297	96,24
162,00	70,00	3005	635,00	115,00	653,50	70,00	0,7554	0,5310	96,26
162,00	70,00	3005	635,00	110,00	653,50	70,00	0,7554	0,5323	96,29
162,00	70,00	3005	635,00	105,00	653,50	70,00	0,7554	0,5324	96,29
162,00	70,00	3005	635,00	100,00	653,50	70,00	0,7554	0,5311	96,27
162,00	70,00	3005	635,00	95,00	653,50	70,00	0,7554	0,5330	96,31
162,00	70,00	3005	635,00	90,00	653,50	70,00	0,7554	0,5330	96,31
162,00	70,00	3005	635,00	85,00	653,50	70,00	0,7554	0,5332	96,31
162,00	70,00	3005	635,00	80,00	653,50	70,00	0,7554	0,5340	96,33
162,00	70,00	3005	635,00	75,00	653,50	70,00	0,7554	0,5340	96,33
162,00	70,00	3005	635,00	70,00	653,50	70,00	0,7554	0,5340	96,33

Continuação da Tabela 2 na próxima página.

Continuação da Tabela 2:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
162,00	70,00	3005	635,00	65,00	653,50	70,00	0,7554	0,5341	96,33
162,00	70,00	3005	635,00	60,00	653,50	70,00	0,7554	0,5335	96,32
162,00	70,00	3005	635,00	45,00	653,50	70,00	0,7554	0,5332	96,31
162,00	70,00	3005	635,00	40,00	653,50	70,00	0,7554	0,5329	96,31
162,00	70,00	3005	635,00	35,00	653,50	70,00	0,7554	0,5330	96,31
162,00	70,00	3005	635,00	30,00	653,50	70,00	0,7554	0,5341	96,33
162,00	70,00	3005	635,00	25,00	653,50	70,00	0,7554	0,5353	96,36
162,00	70,00	3005	635,00	20,00	653,50	70,00	0,7554	0,5361	96,37
162,00	70,00	3005	635,00	200,00	653,50	70,00	0,7543	0,5293	96,37
162,00	70,00	3005	635,00	300,00	653,50	70,00	0,7543	0,5223	96,23
162,00	70,00	3005	635,00	350,00	653,50	70,00	0,7543	0,5199	96,18
162,00	70,00	3005	635,00	400,00	653,50	70,00	0,7539	0,5186	96,19
162,00	70,00	3005	635,00	550,00	653,50	70,00	0,7534	0,5135	96,19
162,00	70,00	3005	635,00	600,00	653,50	70,00	0,7530	0,5117	96,23
162,00	60,00	3005	635,00	25,00	653,60	60,00	0,7552	0,5343	96,36
162,00	70,00	3005	635,00	35,00	653,60	70,00	0,7555	0,5330	96,29
162,00	80,00	3005	635,00	45,00	653,60	80,00	0,7555	0,5327	96,28
162,00	90,00	3005	635,00	55,00	653,60	90,00	0,7554	0,5327	96,28
162,00	100,00	3005	635,00	65,00	653,60	100,00	0,7562	0,5333	96,32
162,00	110,00	3005	635,00	75,00	653,60	110,00	0,7560	0,5336	96,32
162,00	140,00	3005	635,00	105,00	653,60	140,00	0,7562	0,5314	96,28
162,00	150,00	3005	635,00	115,00	653,60	150,00	0,7554	0,5311	96,24
162,00	160,00	3005	635,00	125,00	653,60	160,00	0,7542	0,5332	96,23
162,00	170,00	3005	635,00	135,00	653,60	170,00	0,7550	0,5344	96,29
162,00	180,00	3005	635,00	145,00	653,60	170,00	0,7547	0,5316	96,28

Continuação da Tabela 2 na próxima página.

Conclusão da Tabela 2:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
162,00	190,00	3005	635,00	155,00	653,60	170,00	0,7546	0,5310	96,26
162,00	195,00	3005	635,00	160,00	653,60	170,00	0,7543	0,5313	96,32
100,00	170,00	3000	550,00	135,00	562,00	170,00	0,4306	0,2454	93,91
125,00	170,00	3000	550,00	135,00	569,30	170,00	0,4513	0,2320	94,30
145,00	170,00	3000	550,00	135,00	576,10	170,00	0,4719	0,2189	94,39
162,00	70,00	1500	550,00	135,00	447,10	70,00	0,0998	0,2153	85,91
162,00	70,00	2000	550,00	135,00	529,80	70,00	0,2618	0,2187	91,00
162,00	70,00	2800	550,00	135,00	614,40	70,00	0,5689	0,1124	94,37
156,00	170,00	3000	550,00	135,00	625,90	170,00	0,6234	0,0636	94,61
157,00	170,00	3000	550,00	135,00	626,10	170,00	0,6239	0,0628	94,64
158,00	170,00	3000	550,00	135,00	626,30	170,00	0,6243	0,0621	94,65
162,00	160,00	3000	550,00	135,00	627,20	160,00	0,6273	0,0576	94,64
162,00	150,00	3000	550,00	135,00	627,20	150,00	0,6276	0,0572	94,65
162,00	145,00	3000	550,00	135,00	627,20	145,00	0,6263	0,0593	94,60
162,00	70,00	3002	550,00	135,00	626,80	70,00	0,6264	0,0580	94,98
162,00	70,00	3003	550,00	135,00	626,80	70,00	0,6266	0,0578	94,99
162,00	70,00	3005	570,00	135,00	636,50	70,00	0,6732	0,1493	95,51
162,00	70,00	3005	575,00	135,00	638,90	70,00	0,6846	0,1741	95,36
162,00	70,00	3005	635,00	130,00	653,50	70,00	0,7554	0,5332	96,31
162,00	70,00	3005	635,00	125,00	653,50	70,00	0,7554	0,5317	96,28
162,00	70,00	3005	635,00	55,00	653,50	70,00	0,7554	0,5326	96,30
162,00	70,00	3005	635,00	50,00	653,50	70,00	0,7554	0,5323	96,29
162,00	120,00	3005	635,00	85,00	653,60	120,00	0,7558	0,5333	96,30
162,00	70,00	3005	635,00	450,00	653,50	70,00	0,7539	0,5170	96,16
162,00	70,00	3005	635,00	500,00	653,50	70,00	0,7534	0,5152	96,23
162,00	130,00	3005	635,00	95,00	653,60	130,00	0,7548	0,5335	96,27

4.2. DIVISÃO DOS DADOS EM TREINO E TESTE

A partir dos dados gerados, foi possível realizar o estudo com modelos de *Machine Learning* na plataforma do *Orange*. Para isso, os dados foram divididos de forma aleatória, sendo uma parte deles para o treino do modelo e a outra para o seu teste. Esta divisão foi feita na proporção de 80% para treino e os outros 20% para teste, tal divisão é aconselhável no *Machine Learning*, com o intuito de que o modelo possa entender os parâmetros que têm mais influência e ser treinado a partir dos dados de treino (DOMINGOS, 2017). Em seguida, o restante dos dados é utilizado para o teste do modelo criado, a fim de entender estatisticamente se ele funciona com novos parâmetros.

Com isso, é apresentado na Tabela 3 os dados que foram utilizados para o treinamento do modelo, sendo que os outros apresentados na Tabela 4, posteriormente, foram utilizados para o teste.

Tabela 3 – Dados utilizados para o treinamento do modelo

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
155.80	170.00	3000	550.00	135.00	625.80	170.00	0.6230	0.0639	94.46
95.00	170.00	3000	550.00	135.00	560.80	170.00	0.4275	0.2474	93.74
105.00	170.00	3000	550.00	135.00	563.10	170.00	0.4337	0.2431	94.06
110.00	170.00	3000	550.00	135.00	564.30	170.00	0.4368	0.2409	94.21
115.00	170.00	3000	550.00	135.00	565.50	170.00	0.4399	0.2386	94.37
120.00	170.00	3000	550.00	135.00	567.50	170.00	0.4458	0.2352	94.33
130.00	170.00	3000	550.00	135.00	571.10	170.00	0.4568	0.2289	94.31
135.00	170.00	3000	550.00	135.00	572.30	170.00	0.4602	0.2262	94.47
140.00	170.00	3000	550.00	135.00	574.30	170.00	0.4661	0.2226	94.42
150.00	170.00	3000	550.00	135.00	577.90	170.00	0.4774	0.2150	94.37

Continuação da Tabela 3 na próxima página.

Continuação da Tabela 3:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
155.00	170.00	3000	550.00	135.00	579.20	170.00	0.4809	0.2120	94.54
159.00	170.00	3000	550.00	135.00	626.50	170.00	0.6248	0.0613	94.67
160.00	170.00	3000	550.00	135.00	626.60	170.00	0.6253	0.0605	94.69
161.00	170.00	3000	550.00	135.00	626.80	170.00	0.6258	0.0597	94.71
162.00	170.00	3000	550.00	135.00	627.20	170.00	0.6271	0.0580	94.63
162.00	165.00	3000	550.00	135.00	627.20	165.00	0.6272	0.0579	94.64
162.00	140.00	3000	550.00	135.00	627.20	140.00	0.6277	0.0570	94.66
162.00	130.00	3000	550.00	135.00	627.20	130.00	0.6281	0.0564	94.67
162.00	120.00	3000	550.00	135.00	627.20	120.00	0.6282	0.0561	94.68
162.00	110.00	3000	550.00	135.00	627.20	110.00	0.6283	0.0560	94.68
162.00	100.00	3000	550.00	135.00	627.20	100.00	0.6276	0.0571	94.65
162.00	90.00	3000	550.00	135.00	627.20	90.00	0.6277	0.0570	94.65
162.00	80.00	3000	550.00	135.00	627.20	80.00	0.6279	0.0567	94.66
162.00	70.00	1400	550.00	135.00	437.10	70.00	0.0870	0.2134	82.69
162.00	70.00	1600	550.00	135.00	467.00	70.00	0.1291	0.2184	86.45
162.00	70.00	1700	550.00	135.00	483.40	70.00	0.1573	0.2203	87.79
162.00	70.00	1800	550.00	135.00	499.30	70.00	0.1890	0.2212	89.01
162.00	70.00	1900	550.00	135.00	515.00	70.00	0.2244	0.2206	90.07
162.00	70.00	2100	550.00	135.00	543.90	70.00	0.3021	0.2144	91.73
162.00	70.00	2200	550.00	135.00	557.00	70.00	0.3430	0.2083	92.36
162.00	70.00	2300	550.00	135.00	568.90	70.00	0.3837	0.1997	92.88
162.00	70.00	2400	550.00	135.00	579.80	70.00	0.4235	0.1885	93.32
162.00	70.00	2500	550.00	135.00	589.90	70.00	0.4631	0.1740	93.58
162.00	70.00	2600	550.00	135.00	598.90	70.00	0.5008	0.1562	93.92
162.00	70.00	2700	550.00	135.00	607.00	70.00	0.5359	0.1358	94.16
162.00	70.00	2900	550.00	135.00	621.20	70.00	0.6004	0.0852	94.49

Continuação da Tabela 3 na próxima página.

Continuação da Tabela 3:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
162.00	70.00	3000	550.00	135.00	626.70	70.00	0.6260	0.0585	94.96
162.00	70.00	3005	550.00	135.00	626.90	70.00	0.6269	0.0574	95.02
162.00	70.00	3005	555.00	135.00	630.20	70.00	0.6425	0.0748	94.82
162.00	70.00	3005	560.00	135.00	632.60	70.00	0.6541	0.0967	94.97
162.00	70.00	3005	565.00	135.00	634.90	70.00	0.6651	0.1213	95.12
162.00	70.00	3005	580.00	135.00	640.40	70.00	0.6918	0.2037	95.68
162.00	70.00	3005	585.00	135.00	642.10	70.00	0.7001	0.2320	95.78
162.00	70.00	3005	590.00	135.00	643.70	70.00	0.7078	0.2617	95.84
162.00	70.00	3005	595.00	135.00	645.10	70.00	0.7148	0.2921	95.91
162.00	70.00	3005	600.00	135.00	646.40	70.00	0.7209	0.3215	95.94
162.00	70.00	3005	605.00	135.00	647.70	70.00	0.7268	0.3518	96.02
162.00	70.00	3005	610.00	135.00	648.80	70.00	0.7325	0.3827	96.08
162.00	70.00	3005	615.00	135.00	649.90	70.00	0.7377	0.4129	96.12
162.00	70.00	3005	620.00	135.00	650.90	70.00	0.7425	0.4428	96.17
162.00	70.00	3005	625.00	135.00	651.80	70.00	0.7469	0.4718	96.22
162.00	70.00	3005	630.00	135.00	652.70	70.00	0.7511	0.5023	96.28
162.00	70.00	3005	635.00	135.00	653.50	70.00	0.7554	0.5343	96.34
162.00	70.00	3005	635.00	120.00	653.50	70.00	0.7554	0.5297	96.24
162.00	70.00	3005	635.00	115.00	653.50	70.00	0.7554	0.5310	96.26
162.00	70.00	3005	635.00	110.00	653.50	70.00	0.7554	0.5323	96.29
162.00	70.00	3005	635.00	105.00	653.50	70.00	0.7554	0.5324	96.29
162.00	70.00	3005	635.00	100.00	653.50	70.00	0.7554	0.5311	96.27
162.00	70.00	3005	635.00	95.00	653.50	70.00	0.7554	0.5330	96.31
162.00	70.00	3005	635.00	90.00	653.50	70.00	0.7554	0.5330	96.31
162.00	70.00	3005	635.00	85.00	653.50	70.00	0.7554	0.5332	96.31
162.00	70.00	3005	635.00	80.00	653.50	70.00	0.7554	0.5340	96.33

Continuação da Tabela 3 na próxima página.

Continuação da Tabela 3:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
162.00	70.00	3005	635.00	75.00	653.50	70.00	0.7554	0.5340	96.33
162.00	70.00	3005	635.00	70.00	653.50	70.00	0.7554	0.5340	96.33
162.00	70.00	3005	635.00	65.00	653.50	70.00	0.7554	0.5341	96.33
162.00	70.00	3005	635.00	60.00	653.50	70.00	0.7554	0.5335	96.32
162.00	70.00	3005	635.00	45.00	653.50	70.00	0.7554	0.5332	96.31
162.00	70.00	3005	635.00	40.00	653.50	70.00	0.7554	0.5329	96.31
162.00	70.00	3005	635.00	35.00	653.50	70.00	0.7554	0.5330	96.31
162.00	70.00	3005	635.00	30.00	653.50	70.00	0.7554	0.5341	96.33
162.00	70.00	3005	635.00	25.00	653.50	70.00	0.7554	0.5353	96.36
162.00	70.00	3005	635.00	20.00	653.50	70.00	0.7554	0.5361	96.37
162.00	70.00	3005	635.00	200.00	653.50	70.00	0.7543	0.5293	96.37
162.00	70.00	3005	635.00	300.00	653.50	70.00	0.7543	0.5223	96.23
162.00	70.00	3005	635.00	350.00	653.50	70.00	0.7543	0.5199	96.18
162.00	70.00	3005	635.00	400.00	653.50	70.00	0.7539	0.5186	96.19
162.00	70.00	3005	635.00	550.00	653.50	70.00	0.7534	0.5135	96.19
162.00	70.00	3005	635.00	600.00	653.50	70.00	0.7530	0.5117	96.23
162.00	60.00	3005	635.00	25.00	653.60	60.00	0.7552	0.5343	96.36
162.00	70.00	3005	635.00	35.00	653.60	70.00	0.7555	0.5330	96.29
162.00	80.00	3005	635.00	45.00	653.60	80.00	0.7555	0.5327	96.28
162.00	90.00	3005	635.00	55.00	653.60	90.00	0.7554	0.5327	96.28
162.00	100.00	3005	635.00	65.00	653.60	100.00	0.7562	0.5333	96.32
162.00	110.00	3005	635.00	75.00	653.60	110.00	0.7560	0.5336	96.32
162.00	140.00	3005	635.00	105.00	653.60	140.00	0.7562	0.5314	96.28
162.00	150.00	3005	635.00	115.00	653.60	150.00	0.7554	0.5311	96.24
162.00	160.00	3005	635.00	125.00	653.60	160.00	0.7542	0.5332	96.23
162.00	170.00	3005	635.00	135.00	653.60	170.00	0.7550	0.5344	96.29

Continuação da Tabela 3 na próxima página.

Conclusão da Tabela 3:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
162.00	180.00	3005	635.00	145.00	653.60	170.00	0.7547	0.5316	96.28
162.00	190.00	3005	635.00	155.00	653.60	170.00	0.7546	0.5310	96.26
162.00	195.00	3005	635.00	160.00	653.60	170.00	0.7543	0.5313	96.32

Tabela 4 – Dados utilizados para o teste do modelo

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
100.00	170.00	3000	550.00	135.00	562.00	170.00	0.4306	0.2454	93.91
125.00	170.00	3000	550.00	135.00	569.30	170.00	0.4513	0.2320	94.30
145.00	170.00	3000	550.00	135.00	576.10	170.00	0.4719	0.2189	94.39
162.00	70.00	1500	550.00	135.00	447.10	70.00	0.0998	0.2153	85.91
162.00	70.00	2000	550.00	135.00	529.80	70.00	0.2618	0.2187	91.00
162.00	70.00	2800	550.00	135.00	614.40	70.00	0.5689	0.1124	94.37
156.00	170.00	3000	550.00	135.00	625.90	170.00	0.6234	0.0636	94.61
157.00	170.00	3000	550.00	135.00	626.10	170.00	0.6239	0.0628	94.64
158.00	170.00	3000	550.00	135.00	626.30	170.00	0.6243	0.0621	94.65
162.00	160.00	3000	550.00	135.00	627.20	160.00	0.6273	0.0576	94.64
162.00	150.00	3000	550.00	135.00	627.20	150.00	0.6276	0.0572	94.65
162.00	145.00	3000	550.00	135.00	627.20	145.00	0.6263	0.0593	94.60
162.00	70.00	3002	550.00	135.00	626.80	70.00	0.6264	0.0580	94.98
162.00	70.00	3003	550.00	135.00	626.80	70.00	0.6266	0.0578	94.99
162.00	70.00	3005	570.00	135.00	636.50	70.00	0.6732	0.1493	95.51
162.00	70.00	3005	575.00	135.00	638.90	70.00	0.6846	0.1741	95.36
162.00	70.00	3005	635.00	130.00	653.50	70.00	0.7554	0.5332	96.31

Continuação da Tabela 4 na próxima página:

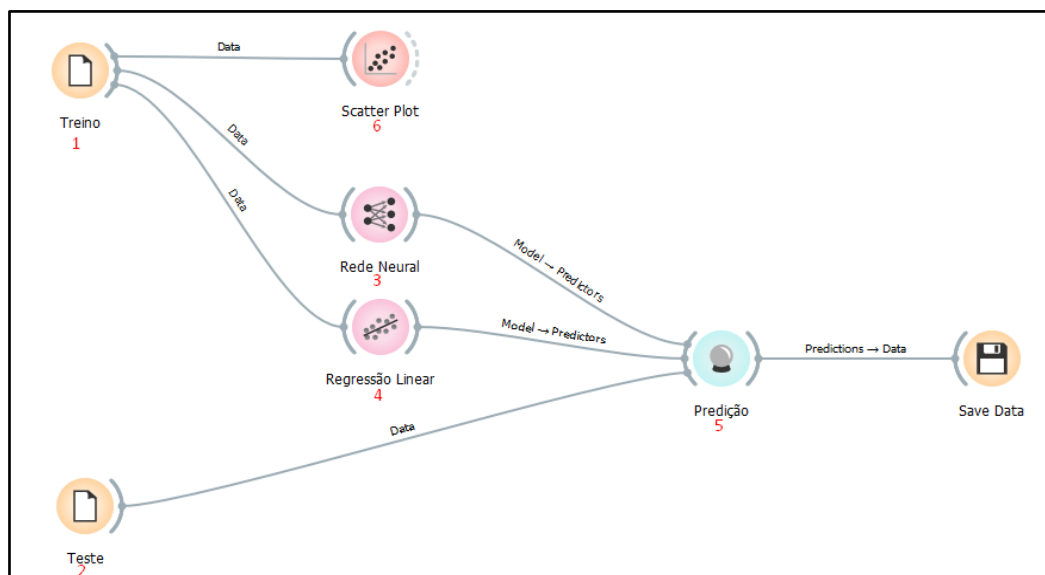
Conclusão da Tabela 4:

Corrente									
3	3	6	9	9	7	7	PFR-100	PFR-101	20
Parâmetro									
T (°C)	P (kPa)	F (kmol/h)	T (°C)	P (kPa)	T (°C)	P (kPa)	Conversão	Conversão	F (kmol/h)
162.00	70.00	3005	635.00	125.00	653.50	70.00	0.7554	0.5317	96.28
162.00	70.00	3005	635.00	55.00	653.50	70.00	0.7554	0.5326	96.30
162.00	70.00	3005	635.00	50.00	653.50	70.00	0.7554	0.5323	96.29
162.00	120.00	3005	635.00	85.00	653.60	120.00	0.7558	0.5333	96.30
162.00	70.00	3005	635.00	450.00	653.50	70.00	0.7539	0.5170	96.16
162.00	70.00	3005	635.00	500.00	653.50	70.00	0.7534	0.5152	96.23
162.00	130.00	3005	635.00	95.00	653.60	130.00	0.7548	0.5335	96.27

4.3. PREDIÇÃO DOS MODELOS NO ORANGE

Por fim, foi utilizado o *Orange* para o treino e predição do modelo, sendo que é apresentado na Figura 14 a estrutura para isso. Posteriormente, é apresentado o que foi utilizado dentro de cada ícone:

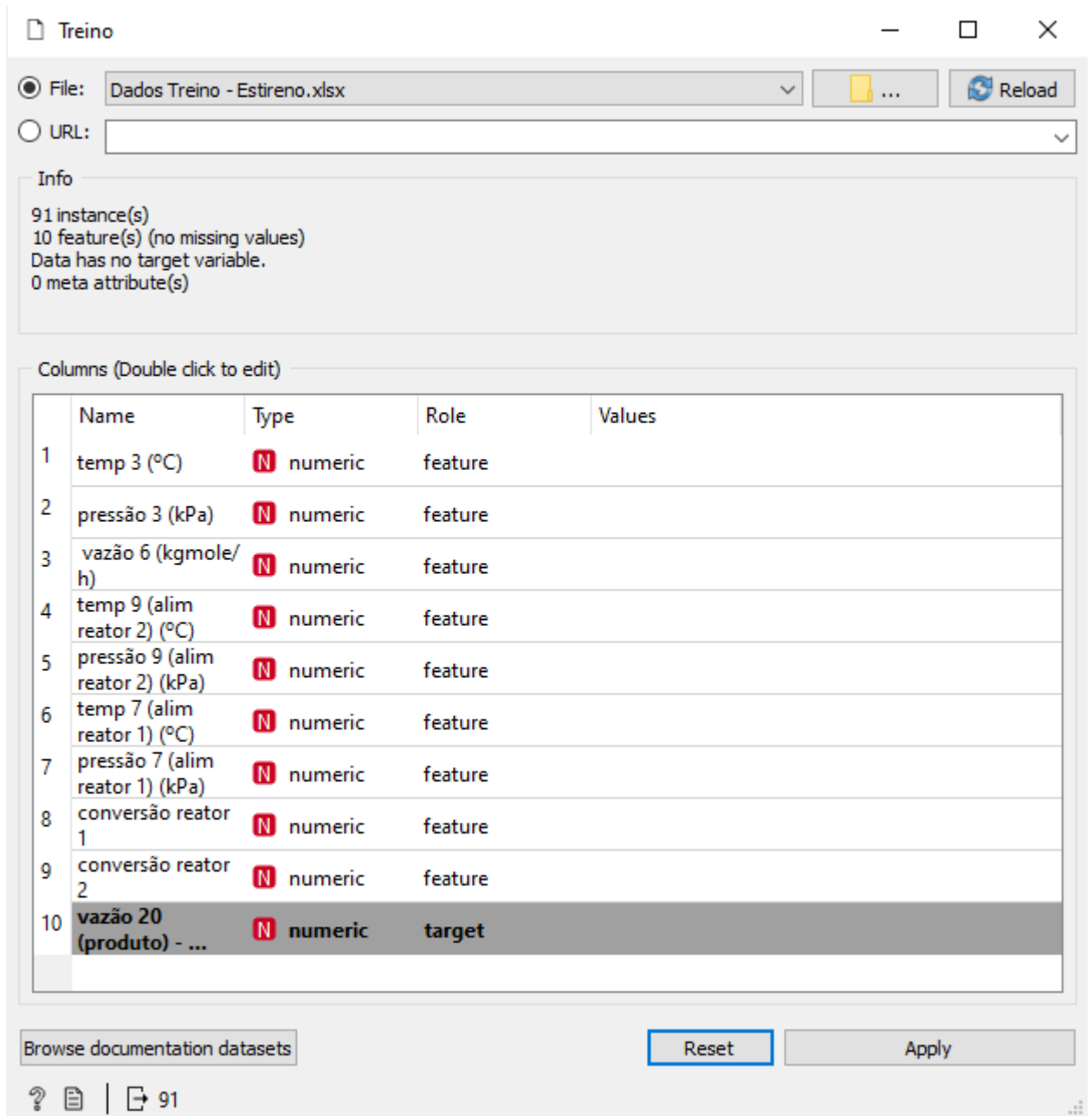
Figura 14 – Estrutura utilizada no simulador Orange



Fonte: *print screen* da estrutura utilizada no simulador *Orange*

Treino (1): Planilha com dados da Tabela 3, em que os parâmetros de treino foram selecionados como *feature* e a vazão do estireno (vazão 20) foi selecionada como *target*, como apresentado na Figura 15:

Figura 15 – Seleção dos dados de treino



The screenshot shows the 'Treino' widget configuration in Orange Data Mining. The 'File' field is set to 'Dados Treino - Estireno.xlsx'. The 'Info' section displays the following information:

- 91 instance(s)
- 10 feature(s) (no missing values)
- Data has no target variable.
- 0 meta attribute(s)

The 'Columns (Double click to edit)' section contains the following table:

	Name	Type	Role	Values
1	temp 3 (°C)	N numeric	feature	
2	pressão 3 (kPa)	N numeric	feature	
3	vazão 6 (kgmole/h)	N numeric	feature	
4	temp 9 (alim reator 2) (°C)	N numeric	feature	
5	pressão 9 (alim reator 2) (kPa)	N numeric	feature	
6	temp 7 (alim reator 1) (°C)	N numeric	feature	
7	pressão 7 (alim reator 1) (kPa)	N numeric	feature	
8	conversão reator 1	N numeric	feature	
9	conversão reator 2	N numeric	feature	
10	vazão 20 (produto) - ...	N numeric	target	

The 'vazão 20 (produto) - ...' row is highlighted in grey, indicating it is the selected target variable. At the bottom of the window, there are buttons for 'Browse documentation datasets', 'Reset', and 'Apply', along with a status bar showing a question mark, a document icon, and the number '91'.

Fonte: *print screen* da seleção dos dados de treino no simulador Orange

Teste (2): Planilha com os dados da Tabela 4, em que os parâmetros de teste foram selecionados como *feature* e a vazão do estireno (vazão 20) foi selecionada como *target*, como apresentado na Figura 16:

Figura 16 – Seleção dos dados de teste

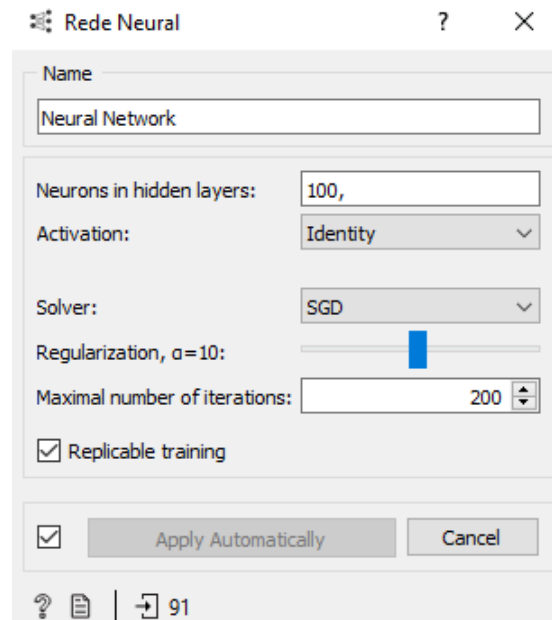
The screenshot shows the 'Teste' window in Orange Data Mining. The file 'Dados Teste - Estireno.xlsx' is loaded. The 'Info' section indicates 24 instances, 10 features, and no target variable. The 'Columns' section shows a table with 10 columns, where the 10th column, 'vazão 20 (produto) - ...', is selected as the target.

	Name	Type	Role	Values
1	temp 3 (°C)	N numeric	feature	
2	pressão 3 (kPa)	N numeric	feature	
3	vazão 6 (kgmole/h)	N numeric	feature	
4	temp 9 (alim reator 2) (°C)	N numeric	feature	
5	pressão 9 (alim reator 2) (kPa)	N numeric	feature	
6	temp 7 (alim reator 1) (°C)	N numeric	feature	
7	pressão 7 (alim reator 1) (kPa)	N numeric	feature	
8	conversão reator 1	N numeric	feature	
9	conversão reator 2	N numeric	feature	
10	vazão 20 (produto) - ...	N numeric	target	

Fonte: *print screen* da seleção dos dados de teste no simulador *Orange*

Rede Neural (3): parâmetros utilizados apresentados na Figura 17:

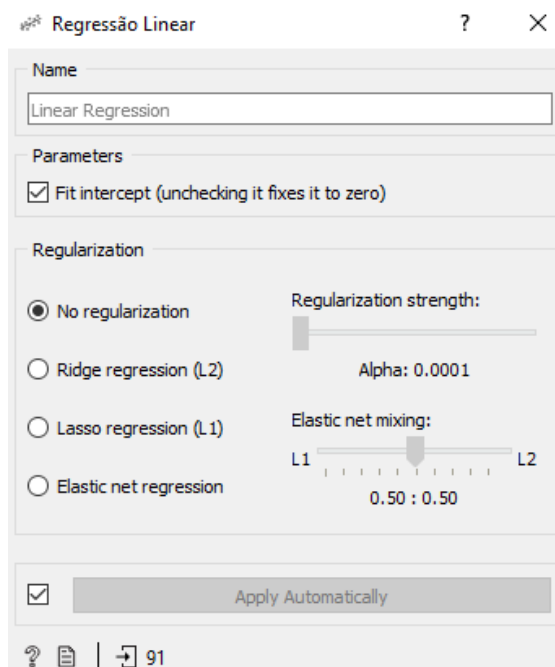
Figura 17 – Seleção dos parâmetros da Rede Neural



Fonte: *print screen* da seleção dos parâmetros da Rede Neural no simulador *Orange*

Regressão Linear (4): parâmetros utilizados apresentados na Figura 18:

Figura 18 – Seleção dos parâmetros da Regressão Linear



Fonte: *print screen* da seleção dos parâmetros da Regressão Linear no simulador *Orange*

Vale destacar que os parâmetros da Rede Neural e da Regressão Linear (Figuras 17 e 18) foram escolhidos em tentativa e erro, a fim de encontrar os melhores resultados para o modelo. Um ponto importante a considerar é que quanto mais camadas (*Neurons in hidden layers* na Figura 17) utilizadas na Rede Neural, mais robusto fica seu algoritmo.

Dessa forma, obtivemos as previsões (dados obtidos no ícone Predição (5) da Figura 14) a partir dos dados de teste da Tabela 4, as quais são apresentadas na Tabela 5, na qual comparamos o valor da vazão de estireno na planta, também apresentado na Tabela 4, com os valores preditos pela Regressão Linear e Rede Neural:

Tabela 5 – Previsões da Rede Neural e Regressão Linear em comparação com os dados de teste

F 20 (kmol/h)	Predição Rede Neural (kmol/h)	Predição Regressão Linear (kmol/h)
93,91	94,05	94,06
94,30	94,19	94,26
94,39	94,30	94,40
85,91	85,25	84,36
91,00	90,30	90,98
94,37	94,46	94,44
94,61	94,65	94,65
94,64	94,65	94,66
94,65	94,65	94,67
94,64	94,67	94,71
94,65	94,70	94,73
94,60	94,74	94,79
94,98	94,94	94,97
94,99	94,94	94,97
95,51	95,17	95,11
95,36	95,24	95,14
96,31	96,32	96,37
96,28	96,31	96,37
96,30	96,32	96,40
96,29	96,31	96,40
96,30	96,17	96,24
96,16	96,21	96,28
96,23	96,20	96,28
96,27	96,15	96,25

É importante destacar que o *range* dos valores apresentados na Tabela 5 é muito baixo. Isso pode ser explicado por conta da planta de Estireno utilizada, já que seus valores não poderiam variar muito, tendo em vista que uma grande variação poderia impossibilitar que a planta digital processasse no *UniSim*.

4.4. AVALIAÇÃO ESTATÍSTICA DO MELHOR MODELO

Junto aos valores da Rede Neural e da Regressão Linear, o *Orange* também forneceu medidas estatísticas, com as quais podemos concluir que o modelo gerado pela Rede Neural atendeu melhor o nosso processo. Vale ressaltar que estas medidas auxiliam a medir a qualidade do modelo, já que correspondem ao quanto a predição se aproximou dos dados observados.

É apresentado, abaixo, uma breve explicação de cada uma das medidas estatísticas de acordo com Legates e McCabe (1999). Em seguida, são exibidos os valores disponibilizados pelo *Orange* por meio da Tabela 6:

A descrição para os índices utilizados na formulação das medidas é apresentada a seguir:

N = número de dados observados

O_i = dados observados

\bar{O} = média dos dados observados

P_i = dados previstos

\bar{P} = média dos dados previstos

Erro Quadrático Médio (MSE): é a média da diferença entre o valor observado e previsto ao quadrado, conforme apresentado na equação (02):

$$MSE = N^{-1} \sum_{i=1}^N (O_i - P_i)^2 \quad (02)$$

Raiz do Erro Quadrático Médio (RMSE): é a raiz da média da diferença entre o valor observado e previsto ao quadrado e possui uma unidade (dimensão) igual ao dos

valores observados e preditos. Dessa forma, sua interpretação é uma medida do desvio médio entre os dados observados e os preditos, como mostrado na equação (03):

$$RMSE = \sqrt{MSE} = \sqrt{N^{-1} \sum_{i=1}^N (O_i - P_i)^2} \quad (03)$$

Erro Médio Absoluto (MAE): representa o desvio médio absoluto entre o observado e o predito. Assim como o RMSE, sua dimensão é a mesma dos valores observados e preditos. Vale destacar que o RMSE dá um peso maior para desvios grandes, já que os valores são elevados ao quadrado. Por outro lado, o MAE dá um peso igual a todos os desvios. Na equação (04) é apresentada sua fórmula:

$$MAE = N^{-1} \sum_{i=1}^N |O_i - P_i| \quad (04)$$

Vale destacar que para as medidas apresentadas até aqui (MSE, RMSE e MAE), valores mais próximos de zero são melhores, já que tais medidas levam em conta o quanto diferiu os valores observados e preditos.

Coefficiente de determinação (R^2): descreve a fração da variância total nos dados observados que pode ser explicada pelo modelo. Esta medida varia de 0,0 a 1,0 sendo que valores mais altos indicam melhor concordância entre o modelo e a observação. A fórmula para seu cálculo é apresentada na equação (05):

$$R^2 = \left\{ \frac{\sum_{i=1}^N (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^N (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^N (P_i - \bar{P})^2}} \right\}^2 \quad (05)$$

Tabela 6 – Medidas obtidas em Predição (5) da Figura 13

Modelo	MSE	RMSE	MAE	R^2
Rede Neural	0,048	0,22	0,125	0,99
Regressão Linear Múltipla	0,113	0,336	0,146	0,976

Com isso, é possível afirmar que a Rede Neural é o modelo que melhor atende a predição, visto que as medidas MSE, RMSE e MAE apresentaram valores mais próximos de zero, quando comparados aos apresentados pela Regressão Linear Múltipla. Além disso, o valor de R^2 da Rede Neural apresentou um valor mais próximo de 1,000 em relação ao da Regressão Linear Múltipla.

Dessa forma, a partir do modelo definido pelo algoritmo da Rede Neural, é possível determinar qual será a produção de Estireno a partir da definição de novos parâmetros, sendo estes os mesmos utilizados no estudo do modelo, os quais foram apresentados na Tabela 4.

4.5. VISUALIZAÇÃO GRÁFICA DE CORRELAÇÕES DE VARIÁVEIS DO MODELO

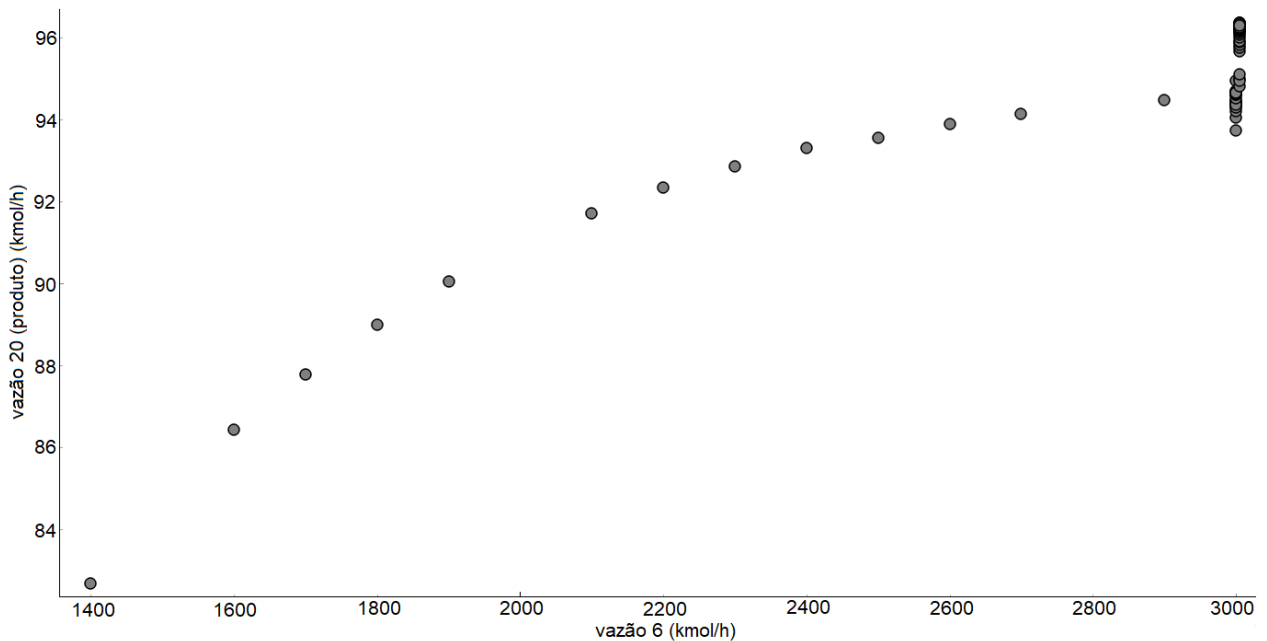
Outro ponto importante para análise é a correlação entre as variáveis do modelo, sendo que a partir disso, é possível identificar quais dos parâmetros utilizados mais influenciam na produção do Estireno. Isso foi possível por meio do *Orange* com a utilização de uma ferramenta que permite plotar gráficos de dispersão correlacionando as variáveis do modelo.

Um ponto importante a ser destacado em relação aos gráficos é que são correlacionados todos os parâmetros que foram utilizados no modelo (temperatura da corrente 3 (°C), pressão da corrente 3 (kPa), vazão da corrente 6 (kmol/h), temperatura da corrente 9 (°C), pressão da corrente 9 (kPa), temperatura da corrente 7 (°C), pressão da corrente 7 (kPa), conversão do reator 1, conversão do reator 2 e vazão da corrente 20 (kmol/h) - produção do Estireno). Por conta disso, observamos pontos sobrepondo nos gráficos, isso porque, tais pontos apresentam as variações dos outros parâmetros correlacionados, enquanto as variáveis que estão plotadas nos gráficos estão constantes.

Sendo assim, são apresentados abaixo alguns dos gráficos que apresentaram maior correlação do produto final desejado, o Estireno, e algum outro parâmetro avaliado:

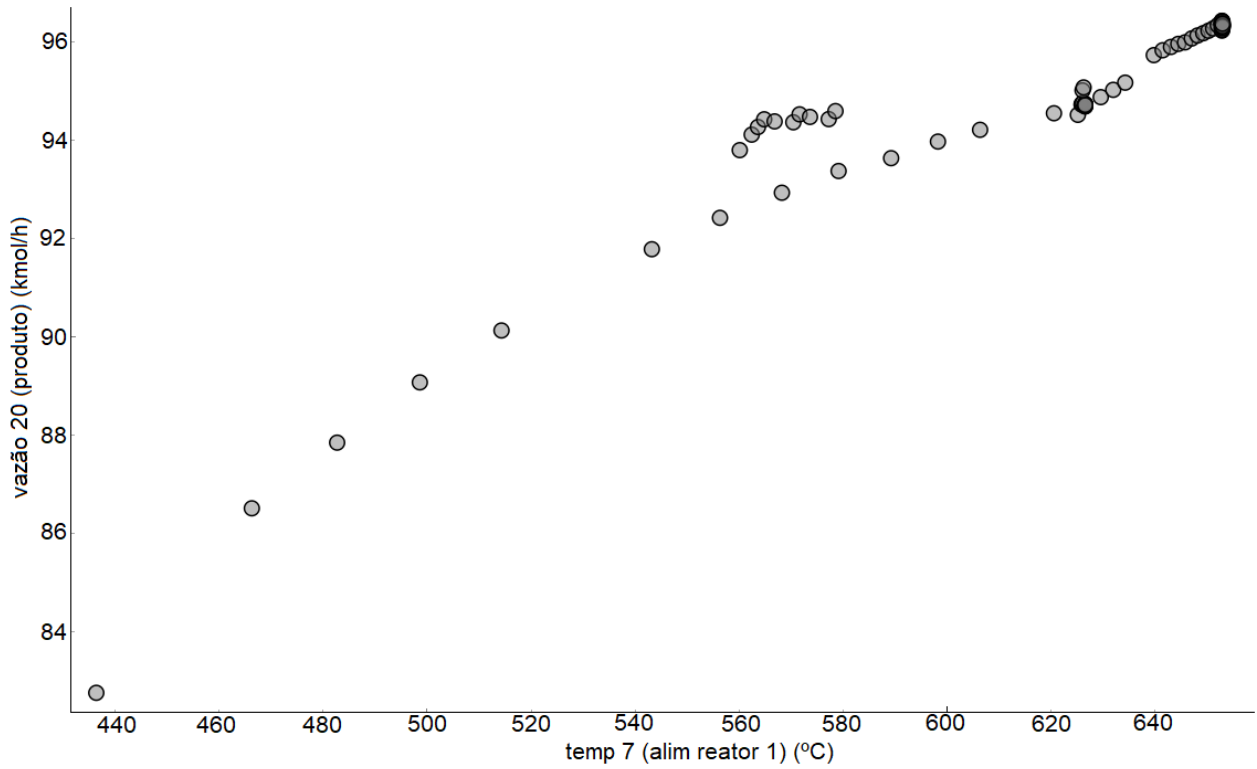
Na Figura 19, é possível observar a correlação entre a vazão 6 (kmol/h), que representa a corrente de vapor superaquecido na entrada do processo, e a vazão 20 (produto). Enquanto há variação da vazão da corrente do vapor superaquecido, é possível observar o aumento, também, da vazão do Estireno. Por outro lado, quando a vazão do vapor superaquecido permanece constante em 3.000 kmol/h, há pouca variação da vazão de Estireno.

Figura 19 – Gráfico vazão 6 (kmol/h) x vazão 20 (produto) (kmol/h)



Fonte: *print screen* do gráfico vazão 6 (kmol/h) x vazão 20 (produto) (kmol/h) no simulador *Orange*

Já na Figura 20, é possível observar a correlação entre a temperatura 7 (°C), que representa a temperatura de alimentação do reator 1, e a vazão 20 (produto). Durante todo o aumento de temperatura foi possível observar que a vazão do Estireno também aumentou de forma proporcional.

Figura 20 – Gráfico temp 7 (alim reator 1) (°C) x vazão 20 (produto) (kmol/h)

Fonte: *print screen* do gráfico temp 7 (alim reator 1) (°C) x vazão 20 (produto) (kmol/h) no simulador *Orange*

5. CONCLUSÃO

É possível observar, hoje, que muitas indústrias e empresas geram uma grande e significativa quantidade de dados que são deixados de lado. A fim de trazer otimizações, a indústria 4.0 com o *Big Data* e o *Machine Learning* apresentam soluções para o tratamento e processamento desses dados, com o intuito de transformá-los em informações relevantes para o auxílio nas tomadas de decisões.

Sendo assim, foi apresentado, neste trabalho, um experimento com uso de algoritmos de *Machine Learning*, uma Rede Neural e uma Regressão Linear Múltipla, a fim de encontrar um modelo que possa prever e otimizar uma planta de Estireno com o aumento de sua produção.

Como conclusão deste trabalho, é possível afirmar que o uso de um algoritmo de *Machine Learning* de uma Rede Neural é satisfatório e pode encontrar um modelo que prediz e gere otimizações do produto final para uma indústria, como a de Estireno apresentada neste trabalho. A partir do qual é possível estimar uma produção, por meio de parâmetros pré-definidos.

Por fim, recomenda-se para trabalhos futuros que queiram abranger melhor o tema uma busca econômica, também, para que o modelo atenda tanto uma maior produção do que se deseja, quanto seja economicamente viável, visto que foi abordado apenas o aspecto da produção neste trabalho. Além disso, é importante observar que há uma limitação na variação dos parâmetros dentro da planta. Dito isso, recomenda-se o cuidado ao tratar novos dados nos modelos de *Machine Learning*, já que eles podem ser válidos nos algoritmos, porém, serem inviáveis no uso real da planta.

6. REFERÊNCIAS

ALVES, Priscila Mello. **Inteligência Artificial e Redes Neurais**. 11 jun. 2020. Disponível em: <<https://www.ipea.gov.br/cts/pt/central-de-conteudo/artigos/artigos/106-inteligencia-artificial-e-redes-neurais>>. Acesso em: 12 mai. 2021.

BAHRIN, Mohd Aiman Kamarul; OTHMAN, Fauzi; AZLI, Nor Hayati Nor; TALIB, Muhamad Farihin. **Industry 4.0: A review on industrial automation and robotic**. Journal Teknologi, 2016. p. 137 – 143.

CETESB. **Estireno**. São Paulo, 2012. Disponível em: <<https://cetesb.sp.gov.br/laboratorios/wp-content/uploads/sites/24/2013/11/Estireno.pdf>>. Acesso em: 12 mai. 2021.

DEMCHENKO, Yuri; MEMBREY, Peter; GROSSO, Paola; DE LAAT, Cees. **Addressing Big Data Issues in Scientific Data Infrastructure**. Research Gate, 2013. Disponível em: <https://www.researchgate.net/publication/256082290_Addressig_Big_Data_Issues_in_Scientific_Data_Infrastructure>. Acesso em: 12 mai. 2021.

DOMINGOS, Pedro. **O Algoritmo Mestre: como a busca pelo algoritmo de machine learning definitivo recriará nosso mundo**. São Paulo (SP): Novatec, 2017. 344 p.

GUJARATI, Damodar N.; Porter, Dawn C. **Econometria Básica**. 5. ed. Porto Alegre (RS): AMGH Editora Ltda., 2011. 924 p.

GRÜBLER, Murillo. **Entendendo o funcionamento de uma Rede Neural Artificial**. 11 jun. 2018. Disponível em: <<https://medium.com/brasil-ai/entendendo-o-funcionamento-de-uma-rede-neural-artificial-4463fcf44dd0>>. Acesso em: 12 mai. 2021.

HONDA, Hugo; FACURE, Matheus; YAOHAO, Peng. **Os Três Tipos de Aprendizado de Máquina**. 27 jul. 2017. Disponível em: <<https://lamfo-unb.github.io/2017/07/27/tres-tipos-am/>>. Acesso em: 12 mai. 2021.

HONEYWELL. Disponível em: <<https://www.honeywellprocess.com/en-US/explore/products/advanced-applications/unisim/pages/default.aspx>>. Acesso em: 12 mai, 2021.

HURWITZ, Judith; KIRSCH, Daniel. **Machine Learning for Dummies**. Hoboken (NJ): John Wiley & Sons, Inc., 2018. 68 p.

LEGATES, David; MCCABE, Gregory J. **Evaluating the Use Of “Goodness-of-Fit” Measures in Hydrologic and Hydroclimatic Model Validation**. Water Resources Research, 1999. p. 233 - 241.

MITCHELL, Tom M. **Machine Learning**. 1. ed. Nova York (NY). McGraw-Hill, 1997. 432 p.

ORANGE. Disponível em: <<https://orangedatamining.com/>>. Acesso em: 27 mai, 2021.

Russel, Stuart J.; Norvig, Peter. **Inteligência Artificial**. 3. ed. Rio de Janeiro (RJ): Elsevier, 2013. 1016 p.

SACOMANO, José Benedito; GONÇALVES, Rodrigo Franco; BONILLA, Sílvia Helena; SILVA, Márcia Terra da; SÁTYRO, Walter Cardoso. **Indústria 4.0: conceitos e fundamentos**. São Paulo (SP): Edgard Blucher, 2018. 169 p.

TURTON, Richard; BAILIE, Richard C., SHAEIWITZ, Joseph A.; BHATTACHARYYA, Debansu; WHITING, Wallace B. **Analysis, Synthesis, and Design of Chemical Processes**. 4. ed. Prentice Hall (NJ), 2012. 1104 p.