



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
CURSO DE ENGENHARIA QUÍMICA

WESLEY MARQUES GONÇALVES

**DETECÇÃO E DIAGNÓSTICO DE FALHAS EM PROCESSOS
INDUSTRIALIS BASEADO EM MODELAGEM DATA-DRIVEN SOB A ÓTICA DA
INDÚSTRIA 4.0**

Recife
2022

WESLEY MARQUES GONÇALVES

**DETECÇÃO E DIAGNÓSTICO DE FALHAS EM PROCESSOS
INDUSTRIAIS BASEADO EM MODELAGEM DATA-DRIVEN SOB A
ÓTICA DA INDÚSTRIA 4.0**

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 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

Gonçalves, Wesley Marques.

Detecção e diagnóstico de falhas em processos industriais baseado em
modelagem data-driven sob a ótica da indústria 4.0 / Wesley Marques Gonçalves.

- Recife, 2022.

64 : 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. Engenharia de processos. 2. Indústria 4.0. 3. Simulação de processos. 4.
Machine learning. 5. Ciência da computação. I. Lucena, Sérgio. (Orientação). II.
Título.

660 CDD (22.ed.)

WESLEY MARQUES GONÇALVES

**DETECÇÃO E DIAGNÓSTICO DE FALHAS EM PROCESSOS INDUSTRIAS
BASEADO EM MODELAGEM DATA-DRIVEN SOB A ÓTICA DA INDÚSTRIA 4.0**

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

Aprovado em: 27/10/2022

BANCA EXAMINADORA

Documento assinado digitalmente



SÉRGIO LUCENA

Data: 11/11/2022 11:57:47-0300

Verifique em <https://verificador.iti.br>

Prof. Dr. Sérgio Lucena (Orientador)

Universidade Federal de Pernambuco

Documento assinado digitalmente



ELIANE BEZERRA DE MORAES MEDEIROS

Data: 11/11/2022 12:09:49-0300

Verifique em <https://verificador.iti.br>

Profa. Dra. Eliane Bezerra De Moraes Medeiros (Examinadora)

Universidade Federal de Pernambuco

Documento assinado digitalmente



RAFAEL ARAUJO MELO

Data: 11/11/2022 11:49:33-0300

Verifique em <https://verificador.iti.br>

Dr. Rafael Araújo De Melo (Examinador)

Universidade Federal de Pernambuco

AGRADECIMENTOS

Agradeço, primeiramente, a Deus por me dar forças e graças para chegar até aqui.

Agradeço ao meu pai, Werico Gonçalves da Silva, por ter me dado apoio e carinho por todo este caminho trilhado.

Agradeço à minha mãe, Sara Marques Cavalcante Gonçalves, por todo o suporte e por ter me tornado a pessoa que sou hoje.

A minha família, por todo apoio emocional durante a minha graduação.

Ao meu orientador Sérgio Lucena pela dedicação e paciência em me ajudar a concluir este trabalho.

Aos meus amigos de faculdade, Maria Karolina, Joyce Amorim, José Leonardo e tantos outros da minha graduação que me deram suporte e tornaram minha caminhada mais leve.

Aos docentes e funcionários do Departamento de Engenharia Química, do Centro de Tecnologia e Geociência e da Área 2 da UFPE, pela contribuição em minha formação.

À minha namorada Ana Clara Bione pelo grande suporte e companheirismo durante toda a construção deste trabalho.

RESUMO

O quantitativo de dados gerados no mundo digital vem crescendo exponencialmente nos últimos anos, tanto em relação à qualidade como à diversidade. A inteligência e a capacidade de utilizar a informação a partir dos dados é considerado por muitos como o “novo petróleo” da contemporaneidade. Nas indústrias, tal fato não ocorre de maneira distinta, as empresas que sabem utilizar os dados em seu benefício possuem uma vantagem competitiva de mercado imensa, tornando-se capazes de reduzir os custos de produção, aumentar a qualidade de seus produtos e garantir eficiência em relação a sua segurança operacional. Diante desse cenário, surge a indústria 4.0, baseada nas tecnologias de manufatura digital, tais como *IoT*, *Big Data*, *Cloud Computing*, além de outras tecnologias emergentes. Nessa revolução, as fábricas tornam-se cada vez mais inteligentes, flexíveis, dinâmicas, ágeis, integradas e capazes de prevenir acidentes de trabalho além de paradas não planejadas ao longo do seu processo de produção. Para que isso de fato ocorra, torna-se imprescindível detectar, diagnosticar e tratar satisfatoriamente as falhas que ocorrem ao longo do processo produtivo. Sabe-se que uma falha é definida como uma variabilidade não natural das variáveis de um processo em seu estado estacionário, as quais não foram devidamente corrigidas pelo controlador. Dessa forma, este trabalho propõe-se a realizar a detecção e o diagnóstico de falhas através de uma análise multivariada utilizando como estudo de caso simulações obtidas do *Tennessee Eastman Process*. Para a detecção, utilizou-se a técnica de PCA combinada com as estatísticas T^2 e Q e posteriormente comparou-se o desempenho do modelo com a análise univariada. E para realizar o diagnóstico das falhas, foram testado duas abordagens distintas, a primeira na qual foi treinada e testada um modelo de rede neural recorrente em LSTM a partir de todas as variáveis de processo normalizadas, e logo em seguida foi realizado o mesmo procedimento, porém com uma prévia transformação linear através da PCA, considerando uma variância acumulada de 90% das PCA's. Estes modelos foram otimizados através de uma variação aleatória dos seus hiperparâmetros e em seguida comparados entre si e entre outros algoritmos de *machine learning* disponíveis na biblioteca *scikit-time*. Para a detecção, o modelo baseado em PCA foi capaz de evidenciar, de maneira satisfatória, 13 das 20 falhas com taxas de detecção de falhas acima de 90%, e para o diagnóstico, o modelo combinado de PCA e LSTM foi capaz de diagnosticar, com mais de 85% de acurácia, 17 das 20 falhas, obtendo uma acurácia global de 94%.

Palavras-chave: Inteligência Artificial; Indústria 4.0; Machine Learn; Python; PCA; LSTM; Falhas; Indústria Petroquímica; Engenharia de processos.

ABSTRACT

The amount of estimated data in the digital world has been growing exponentially, both in quantity and diversity. Intelligence and the ability to use information from data is considered to be the “new oil”. In the industry this is not different, since companies can have competitive advantage. In this context, industry 4.0 emerges, based on digital production context technologies, such as IoT, Big Data, Cloud Computing and other emerging technologies. These changes have made those industries increasingly intelligent, flexible, dynamic, integrated and that must be designed to prevent work accidents and no planning process stops. For this, it is necessary to detect, diagnose and satisfactorily treat the failures that occur in the production process. A failure being defined as an unnatural variability in the variables of a stationary process that has not been corrected by the controller. Thus, this work proposes to carry out a detection and diagnosis of failures through a multivariate analysis using Tennessee Eastman Process simulations as case studies. For detection, the combined PCA technique and T^2 and Q statistics were used and the performance of the model was compared with the univariate one. In addition, to carry out the fault diagnosis, 2 approaches were tested, first a recurrent neural network model was trained and tested from the whole normalized process variables, and then this procedure was done, a linear transformation was previously performed through PCA, considering an accumulated variance of 90% of PCAs. Models were optimized by a series of random characters from these models compared to each other and to other machine learning models available in the scikit-time librarie. For detection, the PCA-based model was able to satisfactorily detect 13 out of 20 faults with fault detection above 90% and combined for diagnosis, the PCA and LSTM model was capable of diagnosis, with more than 85% accuracy, 17 out of 20 failures, with an overall accuracy of 94%.

Keywords: Artificial intelligence; Industry 4.0; Machine Learn; Python; PCA; LSTM; faults; Petrochemical industry; Process engineering.

LISTA DE FIGURAS

Figura 1 – Fluxograma de um modelo supervisionado.	17
Figura 2 – Fluxograma de um modelo não supervisionado.	18
Figura 3 - Fluxograma do Tennessee Eastman Process (TEP).	19
Figura 4 - Ciclo de monitoramento de falhas em processos industriais.	21
Figura 5 - Classificação dos métodos de detecção e diagnóstico de falhas.	22
Figura 6 - Carta de controle para um processo controlado.	24
Figura 7 - Arquitetura de uma rede neural convencional.	30
Figura 8 - Arquitetura de uma rede neural recorrente.	31
Figura 9 - Representação de uma célula LSTM.	32
Figura 10 – Fluxograma do trabalho.	38
Figura 11 – Visualização de parte dos dados não submetidos a falhas.	42
Figura 12 – Visualização de parte dos dados submetidos a falhas.	43
Figura 13 - Distribuição das variáveis de processo sob a condição normal.	43
Figura 14 - Correlação entre os dados de treino.	44
Figura 15 - Cartas de controle para o processo não submetido a falhas.	46
Figura 16 - Cartas de controle para o processo submetida a um erro do tipo randômico na composição das substâncias A, B e C no corrente 4.	47
Figura 17 – Variância explicada por cada componente.	49
Figura 18 – Estatísticas T e Q para o processo não submetido à falhas.	49
Figura 19 – Estatísticas T e Q para o processo submetido a uma falha do tipo perturbação degrau na razão de alimentação A/C na corrente 4, falha IDV(1).	50
Figura 20 – Contribuições de cada variável para a estatística Q submetidas a falha IDV(1).	50
Figura 21 – Estatísticas T e Q para o processo submetido a uma falha do tipo perturbação degrau na temperatura de entrada da água de resfriamento do reator, falha IDV(4).	51
Figura 22 – Contribuições de cada variável para a estatística Q submetidas a falha IDV(4).	52
Figura 23 – Taxas de detecção para todas as falhas do processo.	53
Figura 24 – Métricas dos modelos de <i>Machine Learning</i> importados da biblioteca <i>sktime</i> .	54
Figura 25 – Métricas dos modelos LSTM. Modelo LSTM com a entrada dos dados padrão (a), Modelo LSTM para os dados que sofreram uma transformação por meio da PCA (b).	56
Figura 26 – Previsões dos modelos LSTM. Modelo LSTM com a entrada dos dados padrão (a), Modelo LSTM para os dados que sofreram uma transformação por meio da PCA (b).	57
Figura 27 – Sistema proposto para monitoramento em tempo real de falhas na indústria.	58

LISTA DE TABELAS

Tabela 1 - Variáveis medidas de forma continuada no Tennessee Eastman.	34
Tabela 2 - Variáveis medidas da composição de cada substância no Tennessee Eastman.	35
Tabela 3 - Variáveis manipuladas do Tennessee Eastman.	36
Tabela 4 - Descrição das falhas a que o processo foi submetido.	37
Tabela 5 - Hiperparâmetros em que foi treinado a rede neural LSTM.	55

LISTA DE ABREVIATURAS E SIGLAS

BN	<i>Bayesian Network</i>
CPS	<i>Cyber Physical System</i>
IA	<i>Inteligência Artificial</i>
ICA	<i>Independent Component Analysis</i>
IDE	<i>Integrated Development Environment</i>
IoT	<i>Internet of Things</i> (Internet das Coisas)
IoS	<i>Internet of system</i>
KNN	<i>K-Nearest Neighbors</i>
LSTM	<i>Long Short Term Memory</i>
MLP	<i>Multilayer Layer Perceptron</i>
PCA	<i>Principal Component Analysis</i>
PLS	<i>Partial Least Squares</i>
RNN	<i>Recurrent neural network</i>
SVM	<i>Support Vector Machine</i>
TEP	<i>Tennessee Eastman Process</i>

SUMÁRIO

1 INTRODUÇÃO	11
2 FUNDAMENTAÇÃO TEÓRICA	14
2.1 INDÚSTRIA 4.0	14
2.1.1 Implementação da indústria 4.0 e seus desafios	15
2.2 MACHINE LEARNING	16
2.2.1 Modelo Supervisionado	17
2.2.2 Modelo não Supervisionado	17
2.2.3 Modelo Semi Supervisionado	18
2.3 TENNESSEE EASTMAN PROCESS	18
2.4 FALHAS EM PROCESSOS INDUSTRIAIS	20
2.5 MÉTODOS DE DETECÇÃO E DIAGNÓSTICO DE FALHAS EM PROCESSOS INDUSTRIAIS	22
2.6 MÉTODOS DE DETECÇÃO DE FALHAS APLICADAS AO TEP	23
2.6.1 Controle estatístico univariado do processo	23
2.6.2 Controle estatístico multivariado do processo	24
2.6.2.1 PCA	24
2.6.2.2 Controle estatístico multivariado com PCA	27
2.7 MÉTODOS DE DIAGNÓSTICO DE FALHAS APLICADAS AO TEP	29
2.7.1 Redes neurais	29
3 MATERIAIS E MÉTODOS	33
3.1 SOFTWARES	33
3.2 CONJUNTO DE DADOS	33
3.3 METODOLOGIA	38
3.3.1 Fluxograma do trabalho	38
3.3.2 Limpeza e tratamento dos dados	38
3.3.3 Análise dos dados	39
3.3.4 Cartas de controle	39
3.3.5 Modelagem	39
3.3.5.1 Detecção de Falhas	39
3.3.5.2 Diagnóstico de falhas	40
3.3.5.2.1 Demais modelos	40
3.3.5.2.2 Rede Neural LSTM	41
3.3.6 Sistema de monitoramento em tempo real de falhas na indústria	41
4 RESULTADOS E DISCUSSÃO	42
4.1 ANÁLISE E TRATAMENTO DE DADOS	42
4.2 CONTROLE ESTATÍSTICO UNIVARIADO DO PROCESSO	45
4.3 CONTROLE ESTATÍSTICO MULTIVARIADO DO PROCESSO A PARTIR DO MODELO DE DPCA	48
4.4 MODELOS PARA A PREVISÃO DE DIAGNÓSTICO DE FALHAS	53

4.5 SISTEMA PROPOSTO PARA MONITORAMENTO DE FALHAS EM TEMPO REAL	58
5 CONCLUSÕES	60
REFERÊNCIAS	61

1 INTRODUÇÃO

A Primeira Revolução Industrial iniciou-se na Inglaterra, ainda na segunda metade do século XVIII, e obteve como seu grande marco a mecanização dos processos, possibilitando o início da produção em larga escala. Nas fábricas, as máquinas a vapor eram alimentadas por carvão, tornando-se este o principal combustível utilizado na época. Já a Segunda Revolução Industrial, além de consolidar a produção em larga escala, acelerou o processo de desenvolvimento das indústrias químicas, elétricas, petróleo e aço, além de contribuir, de maneira significativa, para que houvesse uma maior disponibilidade de meios destinados a logística (CASTELLIS, 2007).

A Terceira Revolução Industrial teve o seu início datado em meados do século XX, sendo caracterizada pela substituição da mecânica do estilo analógico pelo digital, além de contar com um grande avanço no que tange a área de telecomunicações, possibilitado, sobretudo, devido à criação da internet. Para além disso, com o fim da Segunda Guerra Mundial e início da Guerra Fria, houve um interesse cada vez maior no ramo das pesquisas científicas, o que possibilitou a ampliação de novas fontes de energia, a exemplo da energia solar, eólica e nuclear (CASTELLIS, 2007).

Todo esse avanço tecnológico tornou-se essencial para o surgimento da Quarta Revolução Industrial, que teve o seu início na Alemanha no ano de 2011, cujo objetivo principal do país era o de fortalecer a sua indústria nacional. Pode-se afirmar que a Quarta Revolução Industrial vem transformando, de modo significativo, a maneira como as empresas fabricam, controlam e distribuem os seus produtos. Tecnologias como *Big Data and Data Analytics*, Inteligência artificial, Internet das coisas, *Machine Learning* além de outras emergentes, estão cada vez mais integradas ao chão de fábricas e na estratégia das empresas, não apenas em relação ao seu processo produtivo, mas também na importante pauta de relacionamento com os seus clientes. É a partir dessa ferramenta estratégica que as indústrias tornam-se cada vez mais ágeis, inteligentes, flexíveis e dinâmicas na contemporaneidade (ANNANTH; ABINASH;RAO, 2021).

O mundo tem se tornado cada vez mais digital, o comportamento dos indivíduos e a forma como estes interagem com os diferentes produtos e serviços vêm sofrendo fortes alterações nas últimas décadas (ANNANTH; ABINASH;RAO, 2021). É comum, por exemplo, pequenos comércios possuírem um site próprio ou um perfil nas redes sociais visando ampliar as suas oportunidades de vendas. Este fenômeno indica que estar conectado na atualidade não é mais uma questão de opção, mas sim uma questão de necessidade.

Dessa forma, é possível constatar que a ciência dos dados vem recebendo muita notoriedade nos últimos anos devido a capacidade de processar esses dados e transformá-los em informação. Trata-se de um campo na ciência capaz de combinar três grandes áreas do conhecimento, como a matemática e estatística, computação e conhecimento de negócio, além de apresentar diversas metodologias no que tange a mineração de dados, processamento, análise e até mesmo criação de modelos para a predição de dados (DA SILVA; SIQUEIRA, 2019).

A partir desse contexto, é possível perceber que a quantidade de dados gerados no mundo vem crescendo exponencialmente. Estima-se que os dados gerados entre os anos de 2018 e 2015, foram maiores do que os gerados entre 1995 e 2015. Para além disso, há ainda a previsão de que essa quantidade triplique de valor até 2025. Sendo assim, as empresas que conseguirem gerar informação a partir desses dados, seja sobre comportamento dos seus clientes, ou sobre o processo produtivo, terão uma imensa vantagem competitiva de mercado (RYDNING, 2018).

Nas indústrias, especialmente as químicas, houve um grande processo de automatização dos processos devido a redução de custos com analisadores, controladores e com armazenamento de dados. O que por sua vez possibilitou o aumento da quantidade de dados históricos armazenados, gerando uma maior complexidade durante a sua análise (SOARES, 2017).

No Brasil, em 2016, 48% das indústrias utilizaram a tecnologia digital em seus processos, já em 2021, esse número subiu para 69%, evidenciando, assim, que há uma tendência à digitalização (PORTAL DA INDÚSTRIA, 2022). Apesar disso, essa mesma pesquisa analisa que as empresas utilizam uma baixa quantidade de tecnologias digitais, constatando que ainda encontram-se na fase inicial da digitalização. Torna-se possível perceber que, a maior parte das indústrias brasileiras medem e armazenam dados sobre o processo, mas não os utilizam de modo eficiente, ou seja, não usufruem totalmente dos benefícios possibilitados pela indústria 4.0 (PORTAL DA INDÚSTRIA, 2022).

Desse modo, uma das tecnologias que encontra-se ainda não muito explorada pelas indústrias brasileiras é a análise de detecção e diagnóstico de falhas que utilizam dados históricos do processo. Trata-se de uma tarefa extremamente importante em uma indústria, tendo em vista que esta falha pode vir a ocasionar paradas inesperadas durante o processo de execução, refletindo em perdas na produção, redução do tempo de vida útil dos equipamentos incluindo até mesmo acidentes de trabalho. Depender apenas do operador humano para lidar com eventos e emergências anormais torna-se um trabalho difícil, devido não apenas ao

tamanho e complexidade dos processos, como também ao amplo escopo da análise de diagnóstico, além da dependência de conhecimento prévio especializado sobre o assunto (GERMANO, 2017).

Nesse contexto, a análise de detecção e diagnóstico de falhas na indústria através da construção de modelos do tipo *data-driven* (orientado a dados do processo) vem ganhando bastante notoriedade ao longo dos últimos anos, uma vez que é capaz de realizar a detecção de falhas em tempo real, além de estabelecer o diagnóstico (indicar as causas raízes) de maneira simplificada baseada apenas em dados, ou seja, não torna-se necessário um conhecimento prévio sobre o processo (CAPACI *et al.*, 2019).

Para criar um modelo de detecção e diagnóstico de falhas em processos industriais é necessário que haja um banco de dados de variáveis do processo ao longo do tempo, bem como a resposta dessas variáveis quando submetidas a diversos tipos de falhas. Levando-se em consideração que os dados históricos reais de indústrias não costumam ser públicos, Downs e Vogel (1993), propuseram o problema de *Tennessee Eastman Process*, que trata-se de uma planta petroquímica que simula um processo real passando a ser considerada um *benchmark* em estudo de controle e simulação de processos.

Sendo assim, o objetivo deste presente trabalho é detectar e diagnosticar falhas em processos industriais a partir das variáveis de processo, por meio de técnicas estatísticas em conjunto a modelos de *machine learning*. Para isso, serão utilizados dados de simulação adaptados do Tennessee Eastman Process (TEP) processo obtidos por Rieth *et al.* (2017), e ferramentas computacionais, além da linguagem de programação Python e dos seus módulos para *machine learning* e *data science*, como as bibliotecas *Pandas*, *Numpy*, *Scikit-Learn* e *pytorch*.

Como objetivos secundários, deseja-se comparar o desempenho entre os diversos modelos analisados. Utilizar técnicas de pré-processamento nos dados para aumentar o desempenho dos modelos, verificar a correlação entre as variáveis de entrada e as variáveis alvo e, por fim, propor um sistema em tempo real de detecção e diagnóstico de falhas na indústria utilizando as tecnologias apresentadas neste trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 INDÚSTRIA 4.0

O termo “indústria 4.0” foi primeiramente mencionado na Alemanha durante a feira de Hannover no ano de 2011 como uma iniciativa estratégica para aumentar a competitividade das fábricas sendo definida como a transformação completa de toda a esfera da produção industrial através da fusão da tecnologia digital e da internet com a indústria convencional. Apesar do termo ser abordado inicialmente na Alemanha, espalhou-se pelo mundo todo como uma nova tendência estratégica dos governos para aumentar a produtividade de suas indústrias (BORTOLINI *et al.*, 2017).

Nesta revolução, as atuais tecnologias de fabricação são transformadas através de sistemas inteligentes que integram o meio físico e o virtual, tais como a internet das coisas, computação em nuvem, machine learning e inteligência artificial. O conjunto dessas tecnologias dá origem ao termo conhecido como *Cyber Physical Systems (CPS)* que referem-se a sistemas conectados por meio da internet das coisas (*IoT*) e que interagem entre si, analisam dados e adaptam-se às mudanças (PERES, 2020). Dessa forma, percebe-se que a indústria 4.0 combina a inteligência artificial, comunicação instantânea, capacidade de analisar um grande volume de dados e automação dos processos industriais para aumentar a performance do sistema produtivo, tendo em vista a possibilidade de tomar decisões inteligentes em tempo real (ANNANTH; ABINASH; RAO, 2021).

A indústria 4.0 não se limita, apenas, aos processos produtivos, mas também é capaz de integrar a logística com a manufatura. Conforme citado pela revista mundo logística (2017), o foco das empresas será dedicado a atender as necessidades dos clientes de forma a produzir bens cada vez mais personalizados, com o menor estoque possível e com o menor tempo de fabricação. Dessa forma, é necessário que haja uma conectividade e alinhamento entre toda a cadeia de suprimentos, o que tornou-se possível através da interconectividade oferecida pela indústria 4.0. É a partir disso que surge o termo Smart Factory, que refere-se a uma fábrica que produz produtos inteligentes, utilizando toda uma cadeia de suprimentos inteligentes, através de equipamentos que atendem às necessidades do consumidor (MABKHOT, 2018).

Mohamed (2018) destaca os principais componentes necessários para que haja a indústria 4.0:

- Mapeamento do processo: É necessário que haja o mapeamento do processo para planejar a implantação;

- Sistemas Cyber Físicos (CPS): Esse sistema integra o mundo físico e o digital, ou seja, cada etapa de produção é acompanhada por dispositivos que enviam dados para uma unidade central. Esta unidade processa os dados de todas as etapas além de disponibilizar e enviar essas informações;
- Internet das coisas (IoT): IoT faz parte dos CPS e permite a comunicação entre os CPSs e entre CPSs e usuários;
- *Big Data e Analytics*: Tratam-se dos avanços da eletrônica no que diz respeito a sensores e a capacidade de armazenamento de dados, em conjunto com os conceitos da logística 4.0 onde cada vez mais a produção encontra-se de maneira dinâmica e orientada para as necessidades dos clientes, são responsáveis também por gerar um aumento expressivo no número de dados colhidos. Tais dados, quando bem analisados, geram inteligência para o processo, e por consequência, vantagem ainda mais competitivas para o mercado;
- *Internet of Service* (IoS): É a internet das coisas (IoT) aplicada aos prestadores de serviços, no qual as empresas podem atrelar um serviços aos seus produtos, agregando ainda mais valor para o usuário final.

2.1.1 Implementação da indústria 4.0 e seus desafios

Obitko e Jirkovsky (2015) e Silveira (2019) citam os 7 princípios que devem ser respeitados para que haja a correta implementação da indústria 4.0. São eles:

- Capacidade de operação em tempo real: Aquisição e tratamento dos dados em tempo real. Assim, os dados devem reagir instantaneamente a qualquer alteração no ambiente, como ruídos e falhas, possibilitando a tomada de decisão em tempo real;
- Virtualização: A indústria 4.0 deve fornecer uma cópia virtual da planta industrial, possibilitando a rastreabilidade e o seu monitoramento em tempo real, promovendo também o uso de simulações;
- Descentralização: A tomada de decisão poderá ser realizada pelo sistema cyber-físico (CPS) de acordo com as necessidades da produção, excluindo a necessidade de indivíduos destinados para esta tarefa. Assim, cada módulo da fábrica poderá trabalhar de forma descentralizada tornando a indústria ainda mais eficiente;
- Orientação à serviços: Utilização de arquiteturas de software orientadas a serviços que estão aliados ao conceito de Internet of Services (IoS);

- Modularidade: A indústria é orientada para a demanda e necessidades dos clientes, proporcionando uma maior flexibilidade para as máquinas;
- Segurança da informação: É a capacidade de proteger todas as informações que serão trocadas durante o processo.

Apesar dos benefícios da indústria 4.0, a sua implementação envolve muitos desafios, tais como, tecnológicos, econômicos, problemas sociais e questões políticas, podendo levar cerca de 10 anos para que possa ser efetivamente concluída (MOHAMED, 2018).

Dennis Kü sters *et al.* (2017) citam como os principais desafios para a implementação da indústria 4.0 a falta de profissionais realmente qualificados e as poucas matérias destinadas aos cursos de engenharia a respeito do tema, além das incertezas acerca dos benefícios financeiros da implantação da indústria 4.0 devido a falta de negócios e também preocupações sobre a segurança dos dados através de provedores terceirizados. Já T. Stock e G. Seliger (2016) afirmam que os principais desafios são a tecnologia, a readaptação da estrutura da empresa face à quarta revolução industrial bem como a segurança da informação.

2.2 MACHINE LEARNING

Machine Learn é definida como a ciência que dá aos computadores a capacidade de aprendizado sem necessariamente serem programados, ou seja, as máquinas são capazes de lidar com dados de maneira eficiente e fazer excelentes previsões a partir deles. Para realizar esta operação, os modelos matemáticos e estatísticos são implementados a partir de uma linguagem de programação, e recebem como entrada as variáveis do sistema, que são previamente escolhidas a partir de técnicas estatísticas, retornando as variáveis de saída ou agrupando as variáveis, que são a previsão do modelo. Sendo assim, o objetivo da machine learning é proporcionar à máquina o aprendizado através dos dados (MAHESH, 2020).

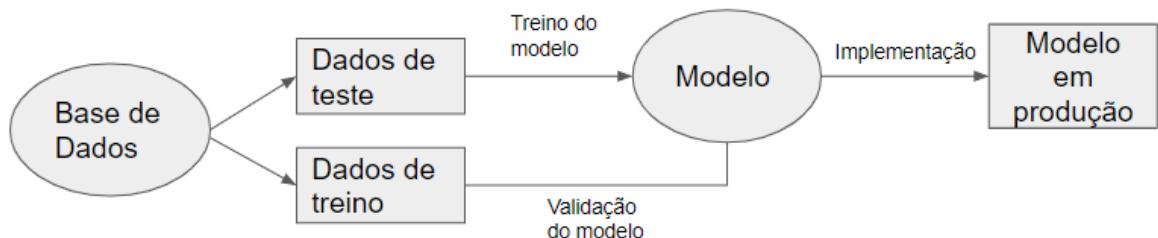
Para aplicar o aprendizado de máquina, existem diversos tipos de modelo, não sendo possível identificar um único modelo que se aplique de modo melhor para todos os problemas. Sendo assim, os modelos de machine learning são divididos em 3 grande classes: Modelos de aprendizagem supervisionada, não supervisionados e semi supervisionados (SANCHES, 2003).

2.2.1 Modelo Supervisionado

Os modelos supervisionados são baseados nos pares de entradas-saídas onde se busca uma função capaz de relacionar as variáveis de entrada e saída. Essa função é construída a partir de modelos matemáticos e estatísticos, em que os parâmetros desses modelos são ajustados de modo a reduzir o erro relativo entre o valor previsto e o valor real (ALPAYDIN, 2021).

Os dados históricos são divididos, geralmente, de maneira aleatória em dois tipos: os de treino e os de teste. Os dados de treino possuem variáveis de saída que devem ser classificadas ou previstas a partir das variáveis de entrada, enquanto que os dados de teste são utilizados para validar o modelo a partir de parâmetros estatísticos (ALPAYDIN, 2021). O fluxograma de funcionamento de um modelo supervisionado é evidenciado na Figura 1.

Figura 1 – Fluxograma de um modelo supervisionado.



Fonte – Adaptado de Mahesh (2020).

2.2.2 Modelo não Supervisionado

Os modelos não supervisionados não possuem uma variável de saída correta, dessa forma, não é possível treinar o modelo a partir do ajuste dos parâmetros. Logo, os modelos tentam agrupar os dados a partir de padrões de similaridades, ou seja, características em comum. Assim, os modelos não supervisionados aprendem as características dos dados e os agrupam dessa maneira, o que é conhecido como *clustering*. Quando um novo dado é introduzido, o modelo percebe suas características e as utiliza para agrupá-las pelos atributos já aprendidos (ALPAYDIN, 2021).

Esse tipo de modelo é bastante utilizado para agrupamento de atributos e também para reduzir variáveis. Possuem aplicações em sistemas de recomendação de produtos ou para detectar *spams* em e-mail (ALPAYDIN, 2021). O fluxograma de um modelo não supervisionado é apresentado na Figura 2 a seguir.

Figura 2 – Fluxograma de um modelo não supervisionado.



Fonte – Adaptado de Mahesh (2020).

2.2.3 Modelo Semi Supervisionado

Nesse tipo de modelo, há um conjunto de dados de treinamento e dados não rotulados que também estão disponíveis para treinamento. Dessa forma, admite-se o treinamento do modelo com dados não rotulados que está disponível, junto aos já rotulados. A motivação da existência desse modelo é que para muitos casos há menos dados rotulados ou disponíveis em comparação aos não rotulados. Um outro motivo é que os dados não rotulados são mais fáceis de serem adquiridos, pois não necessitam da intervenção humana ou de outros modelos de classificação para a sua rotulação (SANCHES, 2003).

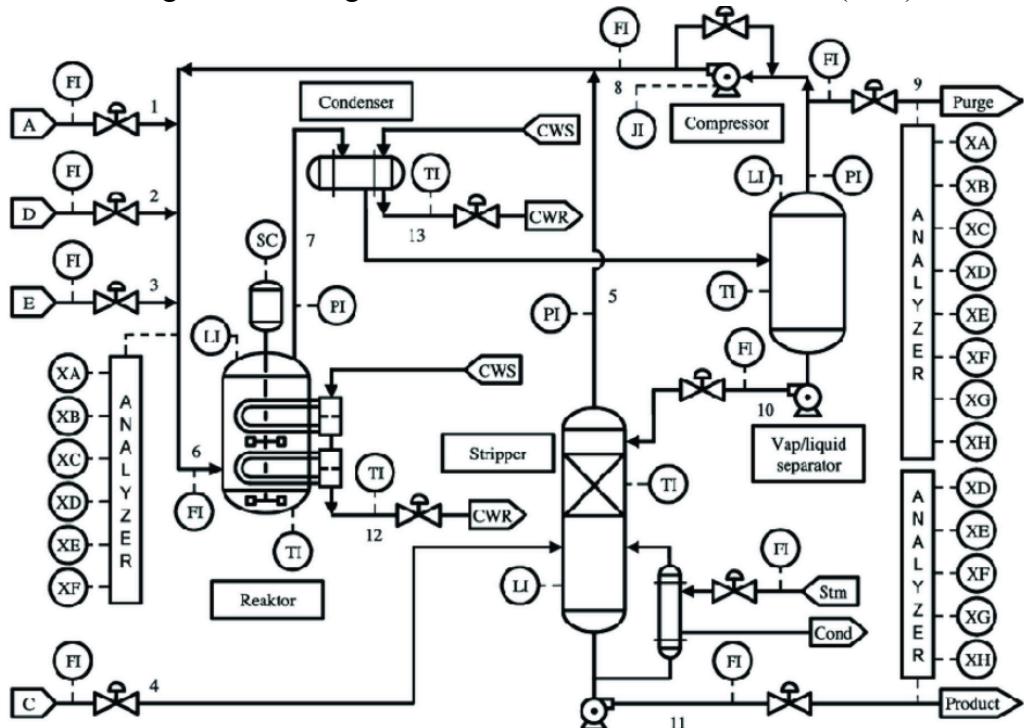
2.3 TENNESSEE EASTMAN PROCESS

A melhor forma de aplicar as principais técnicas de detecção e diagnóstico de falhas é através dos dados específicos relacionados a diversos tipos de falhas em uma indústria real, porém esses dados não costumam estar disponíveis de maneira pública (DOS SANTOS, 2018). É nesse contexto que o *Tennessee Eastman Processs* (TEP) foi desenvolvido pela Eastman Chemical Company com o intuito de simular falhas em processos industriais (DOWNS; VOGEL, 1993). Mesmo sendo criado em 1993, esse modelo se mostra extremamente importante para comparar modelos de detecção de falhas, uma vez que apresentam características de dados da indústria, como sistemas multicomponentes complexos, variáveis multicolineares e não lineares.

O TEP consiste em um processo contínuo com 5 principais componentes: um reator, um separador líquido-vapor, uma coluna de destilação, um compressor de reciclo e um condensador de produtos, conforme mostrado na Figura 3, possuindo 4 reagentes gasosos como entrada, A, C, D e E e um inerte identificado na entrada B, formando dois produtos líquidos na entrada G e H e um subproduto na F, como evidenciado nas reações que vai da

etapa 1 até a 4 (DOWNS; VOGEL, 1993):

Figura 3 - Fluxograma do *Tennessee Eastman Process* (TEP).



Fonte: Downs e Vogel (1993).



Em suma, os reagentes A,C, D, E e o inerte B formam uma corrente de entrada gasosa no reator, que em seguida reagem no reator, com auxílio de um catalisador não volátil e a reação, por ser exotérmica, é resfriada por meio de uma serpentina que circula água. Em seguida, o produto dessa reação, ainda na forma gasosa, segue para o condensador em que a corrente é resfriada e parte desta é condensada. Essa corrente por sua vez é transportada para um separador líquido-gasoso que retira o produto gasoso não condensado e o envia como refluxo, através de um compressor, em direção ao reator. Parte dessa corrente ainda sofre purga, justamente para evitar o acúmulo de subproduto inerte e para que se evite também a

diminuição da eficiência do processo. O produto condensado é enviado para a coluna de destilação, responsável por retirar o reagente que ainda permanece na corrente por meio da adição do reagente C. Finalmente, os produtos G e H são retirados da coluna de destilação e seguem para um processo de separação, que não estão incluídos nesta planta (DOWNS; VOGEL, 1993).

Quanto às reações que ocorrem no processo, estas são consideradas irreversíveis, exotérmicas e aproximadamente de primeira ordem em relação à concentração dos reagentes. Além disso, as constantes cinéticas se relacionam com a temperatura por meio da função de Arrhenius, em que a reação de formação de G tem uma energia de ativação superior a de H, resultando em uma maior sensibilidade à temperatura (DOWNS; VOGEL, 1993).

Os dados de simulação do TEP permitem analisar o comportamento do processo quando submetido a diversos tipos de falhas. Ao todo, são analisadas 20 tipos de falhas através da manipulação de 12 variáveis e da observação de 41 variáveis de processos (DOWNS; VOGEL, 1993).

2.4 FALHAS EM PROCESSOS INDUSTRIAIS

Devido a competitividade no mercado, a necessidade de aumento de segurança operacional, regras ambientais, aumento da qualidade dos produtos e redução dos custos com manutenção, os processos industriais se tornaram cada vez mais complexos e difíceis de serem analisados manualmente. Diante deste cenário, para manter a competitividade, muitas empresas recorrem às tecnologias da indústria 4.0, o que tem aumentado显著mente a quantidade de dados coletados, os quais são analisados em tempo real por operadores, e a maior parte encontram-se armazenados em uma base de dados (LOMOV *et al.*, 2021).

Contudo, com o nível de complexidade do processo e a quantidade enorme de dados sendo gerados, torna-se cada vez mais difícil para o operador, analisar todas as variáveis do processo e agir da maneira mais eficiente frente às anomalias (LOMOV *et al.*, 2021).

Sendo assim, com o intuito de controlar as variáveis, os processos industriais operam em malha fechada, ou seja, uma série de conjunto de controladores são utilizados para manter as variáveis do processo em níveis aceitáveis, compensando erros considerados aleatórios durante o processo. Porém, naturalmente ocorrem mudanças onde os controladores não conseguem agir da maneira mais adequada, as quais são denominadas de falhas de processos (GERMANO, 2017).

A ocorrência de falhas gera uma grande quantidade de riscos operacionais, que podem

ocasionar perdas econômicas, danos à infraestrutura e sanções legais para uma empresa, além da possibilidade de incidentes ou acidentes catastróficos, riscos para trabalhadores, clientes e também para o meio-ambiente (FRANK, 2009). Dessa forma, nota-se a importância de evitar-se as falhas internas em uma empresa, porém, quando esta ocorre, torna-se imprescindível a rápida detecção de sua origem.

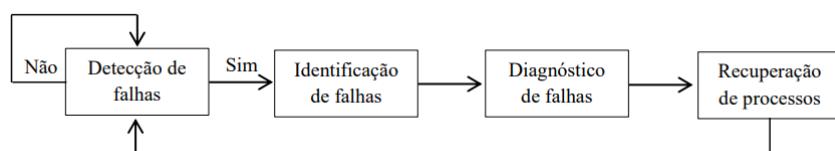
As falhas de processo contínuo podem ser classificadas quanto a sua forma de ocorrência em três tipos, são elas: falhas abruptas, incipientes e intermitentes. As falhas abruptas ocorrem em um pequeno intervalo de tempo e são responsáveis por um grande impacto durante o processo produtivo. Já as falhas incipientes ocorrem de maneira gradual e são bastante difíceis de serem detectadas pois normalmente são mascaradas pela ação de controladores. Por fim, as falhas intermitentes ocorrem de maneira repetida ao longo do processo, sendo caracterizadas por perturbações periódicas (SIMONI; FANTUZZI; ATTON, 2003).

Quanto às abordagens realizadas para que haja o devido monitoramento das falhas durante o processo produtivo, Chiang, *et al.*, (2000) citam 4 etapas basilares para solucioná-las:

- Detecção de falhas: Determina a ocorrência de uma falha a partir de alguma variabilidade anormal acerca de uma variável de processo;
- Identificação de falhas: Determinação do tamanho e do comportamento, no tempo correto, de uma falha. Essa etapa ocorre após a isolação (ou isolamento) da falha;
- Diagnóstico da falha: Determina quando, onde, qual a intensidade e qual a falha ocorreu no processo;
- Recuperação do processo: Refere-se ao tratamento propriamente dito dos efeitos da falha.

A Figura 4, por sua vez, esquematiza de maneira resumida as importantes etapas desse ciclo:

Figura 4 - Ciclo de monitoramento de falhas em processos industriais.

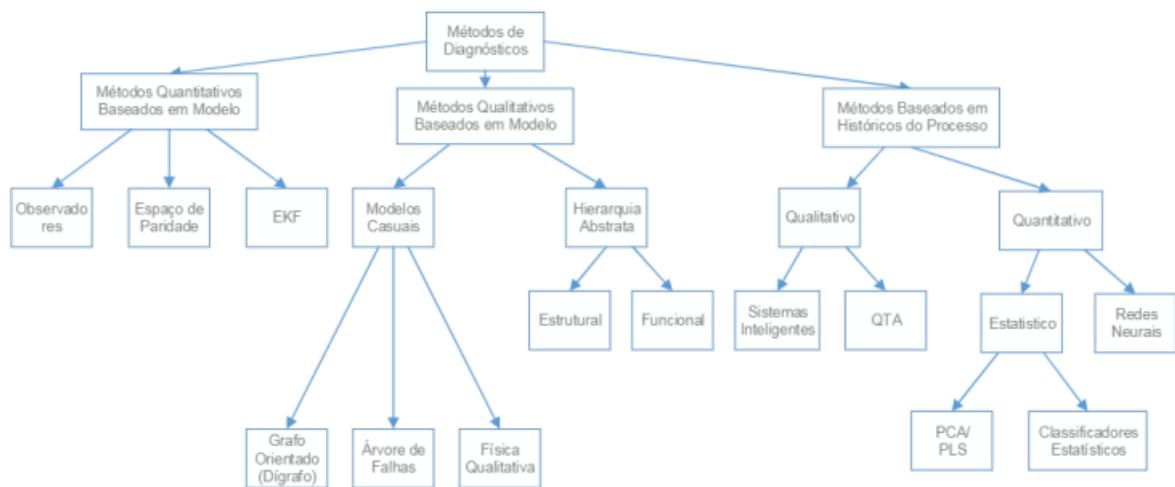


Fonte: Adaptado de CHIANG, *et al.* (2000).

2.5 MÉTODOS DE DETECÇÃO E DIAGNÓSTICO DE FALHAS EM PROCESSOS INDUSTRIAIS

Foram desenvolvidas várias formas de detectar e diagnosticar falhas durante o processo industrial. Venkatasubramanian *et al.* (2003), classificaram os métodos de detecção de falhas em: Métodos quantitativos baseados em modelos físicos, métodos qualitativos baseado em modelos e métodos baseados em histórico do processo, cuja classificação encontra-se esquematizada na Figura 5.

Figura 5 - Classificação dos métodos de detecção e diagnóstico de falhas.



Fonte: Germano (2017).

Os métodos baseados em modelos físicos requerem um conhecimento prévio acerca do processo a ser analisado, podendo ser expressado através de equações matemáticas que representam as relações de entrada e saída do processo. Essas relações podem ser tanto a nível quantitativo como também expressas através das funções qualitativas. Já os métodos que se baseiam no histórico do processo, necessitam apenas da análise e estudo histórico das variáveis de processo, dispensando um conhecimento prévio aprofundado acerca do tema (VENKATASUBRAMANIAN *et al.*, 2003).

Em processos ditos mais complexos, que são mais comuns em processos industriais, a detecção de falhas baseadas em modelos físicos tornam-se cada vez menos viáveis, devido à quantidade de variáveis, a multicolinearidade das variáveis e imprevisibilidade do sistema. Sendo assim, modelos baseados em dados históricos vêm se destacando cada vez mais, pois utiliza-se das técnicas de estatística, matemática e *machine learning* tornando-se capaz de gerar modelos para sistemas extremamente complexos e com inúmeras variáveis (CAPACI, *et*

al., 2019).

Modelos *Data Driven* para a detecção de falhas podem ser divididos em dois tipos: aprendizado supervisionado e não supervisionado. Modelos de aprendizado supervisionado incluem as redes neurais artificiais (SUN *et al.*, 2020);(SHENFIELD E HOWARTH, 2020), *support vector machine* (ONEL, CHRIS A; PISTIKOPOULOS, 2019) e redes bayesianas (BN) (CHEN; KHER; SOMANI, 2006). E modelos não supervisionados, incluem a análise de componentes principais (PCA) (AMIN; IMTIAZ; KHAN, 2018), análise dos mínimos quadrados (PLS) (BOTRE, 2017) e análise de componentes independentes (ICA) (CHEN, 2016).

Os modelos do tipo *Data Driven* para diagnóstico de falhas são utilizados em uma etapa posterior à análise de detecção de falhas. Sendo assim, uma vez detectada a falha, esses modelos são acionados para indicar o tipo exato da falha que está ocorrendo. São catalogados como modelos do tipo de aprendizagem supervisionada nos quais os algoritmos são treinados a partir de séries temporais multivariadas do processo. Dentro os algoritmos mais utilizados, incluem-se SVM, Árvores de decisão e KNN (TAQVI *et al.*, 2021).

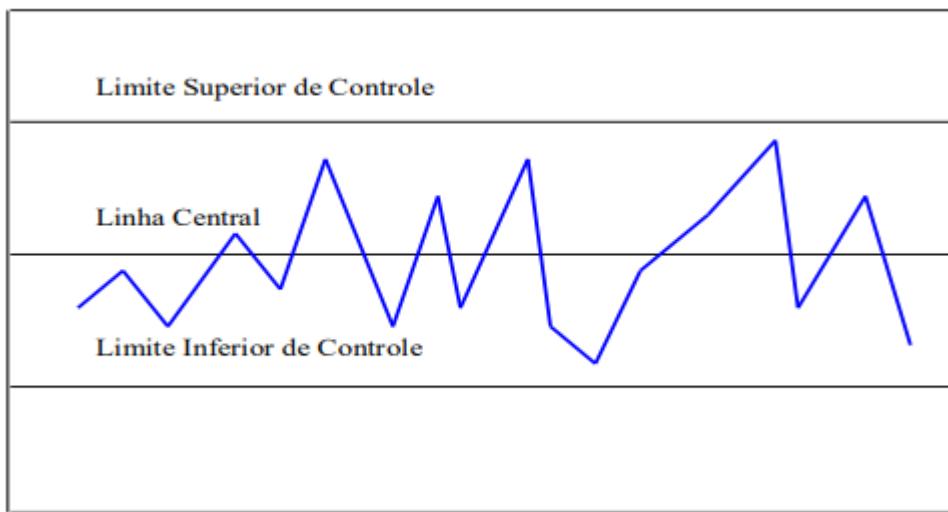
2.6 MÉTODOS DE DETECÇÃO DE FALHAS APLICADAS AO TEP

2.6.1 Controle estatístico univariado do processo

O controle estatístico de processo ou CEP surgiu para monitorar a variabilidade do processo, de modo a torná-lo estável e repetitivo. Logo, um dos principais objetivos da CEP é detectar a ocorrência de anomalias que causam uma alteração considerável nas variáveis de processo. Para que ocorra essa ação são utilizadas as cartas de controle, que tratam-se de gráficos que exibem a evolução de um processo ao longo do tempo, onde é determinado um limite para o qual se define uma faixa considerada aceitável (SUMAN; PRAJAPATI, 2018).

A Figura 6 evidencia uma típica carta de controle em que a linha central representa a média da variável analisada quando o processo está controlado, ou seja, dentro dos limites aceitáveis. Já as outras 2 linhas horizontais representam o limite superior e inferior de controle. Medições acima ou abaixo desses limites indicam que determinada amostra está fora da normalidade e deve ser verificada as possíveis causas dessa variabilidade (OLIVEIRA; ANUNCIAÇÃO; LOPES, 2018).

Figura 6 - Carta de controle para um processo controlado.



Fonte: Oliveira, Anunciação e Lopes (2018).

Na carta de controle padrão de Shewhart admite-se que uma variável encontra-se devidamente controlada quando ela está dentro do limite de $\mu \pm 3\sigma$, sendo μ , a média das observações e σ , o desvio padrão das observações. Dessa forma, admite-se um risco de 0,27% para um alarme falso, ou seja, quando uma amostra dentro dos limites é considerada fora de controle, considera-se também que 99,73% das amostras encontram-se em torno da variação da média (OLIVEIRA; ANUNCIAÇÃO; LOPES, 2018).

As cartas de controle são úteis quando o processo é mais simples ou deseja-se avaliar poucas características do processo. Porém, como é gerado uma carta de controle para cada variável, quando o número de variáveis a serem monitoradas aumenta, fica ainda mais complicado monitorar e gerir o processo. Além disso, as cartas de controle monitoram as variáveis de modo individual, portanto, não avaliam a relação entre as próprias variáveis. Para contornar essas problemáticas opta-se pela análise multivariada do processo (MEHMOOD, 2022).

2.6.2 Controle estatístico multivariado do processo

2.6.2.1 PCA

A técnica de PCA (Análise de Componentes Principais) é uma técnica de análise exploratória de dados que se aplica às mais diversas áreas de pesquisa e busca identificar tendências entre os dados através da redução do número de variáveis, sem que haja uma grande perda de informação, ou seja, de variância. Para além disso, uma característica muito importante da técnica de PCA é que ela ajuda a visualizar agrupamentos entre as amostras,

visto que reduzem a dimensionalidade do conjunto original de variáveis, tornando-se mais fácil plotar os gráficos com duas ou três dimensões sem perda significativa de informação (FERREIRA *et al.*, 2017).

A técnica de PCA aplica uma transformação linear no conjunto original dos dados, conforme evidenciado pela Equação 5 (SHLENS, 2014).

$$T = YP \quad (5)$$

Assim, para um conjunto de dados $Y \in \mathbb{R}^{n \times m}$, sendo n , o número de linhas e m , a quantidade de colunas, aplica-se uma mudança de base a partir da matriz $P \in \mathbb{R}^{m \times m}$ conhecida como matriz de componentes principais, sendo ainda caracterizada como uma matriz de variância e covariância, na qual a diagonal principal representa a variância e o restante da matriz, e a covariância entre cada variável, conforme mostrado pela Equação 6 (SHLENS, 2014).

$$\begin{bmatrix} cov(a_1a_1) & cov(a_1, a_2) & \dots & cov(a_1, a_m) \\ cov(a_2a_1) & cov(a_2, a_2) & \dots & cov(a_2, a_m) \\ \vdots & \vdots & \ddots & \vdots \\ cov(a_ma_1) & cov(a_m, a_2) & \dots & cov(a_m, a_m) \end{bmatrix} \quad (6)$$

As colunas de $T \in \mathbb{R}^{n \times m}$ representam as variáveis latentes e são projeções das colunas de Y na base P , ou seja, nos eixos de componentes principais que são ortogonais entre si e normalizadas, logo linearmente independentes e possuem o mesmo módulo (SHLENS, 2014).

Dado que hajam, “ m ” variáveis no conjunto de dados originais, serão produzidas “ m ” PCAs, porém, essas PCAs são ordenadas de modo a traçar a maior variabilidade entre os dados. Logo, a primeira PCA possui a direção de maior variância dos dados, a segunda PCA apresenta a segunda maior direção com maior variância entre os dados e assim por diante. Sendo assim, é possível escolher um número de PCAs que represente uma porcentagem, em geral determinada como 90%, de variabilidade acumulada dos dados, possibilitando uma simplificação no conjunto de dados (SHLENS, 2014).

Dessa forma, com o intuito de reduzir a dimensionalidade, é selecionada uma

quantidade “a” de componentes principais que dá origem a uma matriz P reduzida denominada de P_a que aplicada ao conjunto original de dados forma a matriz de variáveis latentes T_a (SHLENS, 2014), conforme mostra a Equação 7 abaixo:

$$Ta = yPa \quad (7)$$

Como as variáveis mais importantes são aquelas que mais variam, a matriz T_a representa de maneira mais eficiente e adequada a relação entre os dados, pois não contém as variáveis com as menores variâncias, que correspondem em geral ao ruído da medição. Essa matriz pode ser projetada na matriz original gerando uma matriz reconstruída \hat{Y} com uma quantidade de a variáveis (SHLENS, 2014). Essa reconstrução é mostrada a seguir na Equação 8.

$$\hat{Y} = TaPa' = Y(PaPa)' \quad (8)$$

A diferença entre Y e \hat{Y} é a matriz residual E , dada por:

$$E = Y - \hat{Y} \quad (9)$$

Enquanto a matriz \hat{Y} contém a variabilidade escolhida pelo modelo PCA, E contém a variabilidade não capturada do modelo. Assim o subespaço vetorial gerado por \hat{Y} é considerado o espaço principal e o gerado por E , é considerado o espaço residual (SHLENS, 2014).

Porém, para para o uso da técnica de PCA, é necessário que os dados originais cumpram 4 pré-requisitos básicos, são eles: (SHLENS, 2014).

- Linearidade: Os dados devem possuir relações lineares de modo que as variáveis latentes possam ser geradas a partir da transformação linear da matriz original;
- Redundância ou multicolinearidade: Devem estar presentes relações de redundância entre as variáveis, pois na ausência, a transformação linear não faria efeito;
- Alto razão de sinal/ruído: Essa hipótese deve ser respeitada para que se considere que variâncias grandes representam informações úteis e não apenas ruídos;

- Conjunto de dados centralizados: De forma que a média de Y e T sejam nulas, quando aplicada a normalização.

2.6.2.2 Controle estatístico multivariado com PCA

Utiliza-se a técnica de PCA para análise multivariada dinâmica do processo através de estatísticas, é também utilizada para analisar séries temporais, como a exemplo dos casos que ocorrem em processos industriais. Ainda assim, é necessário duas condições adicionais além das hipóteses enumeradas para o uso da técnica de PCA, são elas: que o processo esteja na condição estacionária e que siga uma distribuição normal (PARK; FAN; HSU, 2020).

Com tais pré requisitos sendo atendidos, o monitoramento do processo industrial se dá em duas etapas: (PARK; FAN; HSU, 2020)

1. Na etapa de treino, utiliza-se os dados em operação normal da fábrica, o que implica em calcular as matrizes de projeção de Y em P_a . Assim, supõe-se que o modelo é capaz de explicar a variabilidade nas condições normais de operação;
2. Na etapa de teste, os novos registros são avaliados se estão conforme o modelo. Assim, caso as observações não se enquadrem como o modelo prever para as condições normais de processo, o processo sai do comportamento considerado normal e pode ser classificado como uma falha.

Dessa forma, de acordo com a dinâmica do modelo, ele é classificado como não supervisionado, pois sua função é detectar *clusters* de variáveis latentes que não estão de acordo com a normalidade do processo.

Como o PCA separa o espaço original dos dados em 2 subespaços, o principal, gerado por \hat{Y} e o residual, gerado por E , torna-se útil para monitorar mudanças em ambos. O modo de monitoramento desses subespaços é a partir das estatísticas T^2 e Q , e quando essas estatísticas ultrapassam o limite de detecção é acionado um alarme, que indica uma possível falha no processo (PARK; FAN; HSU, 2020).

A estatística T^2 monitora o subespaço vetorial principal e são aplicadas às variáveis latentes, que são as projeções de Y nos primeiros a componentes principais, de acordo com equação abaixo:

$$T^2 = t_a' \Lambda_a^{-1} t_a \quad (10)$$

Em que t , representa a variável latente, Λ_a a variância explicada.

Para Kundu e Damarla (2017) a estatística T^2 indica a variância global explicada pelo modelo no espaço principal para as a variáveis, associadas ao processo. O aumento do valor de T^2 sugere algum tipo de falha da mesma natureza que as variáveis escolhidas pelo modelo PCA. Desse modo, delimitar um limite de controle para T^2 significa demarcar uma região de hiper elipse no subespaço vetorial das variáveis latentes.

Um limite de controle muito utilizado é evidenciado através da Equação 11 (GARCIA-ALVAREZ, *et al.*, 2009). Sendo T^2_{lim} o limite de controle, e $F_{\alpha(a,n-a)}$ a distribuição de Fisher-Snedecor com graus de liberdade α e $n - a$.

$$T^2_{lim} = \frac{a(n^2-1)}{n(n-a)} F_{\alpha(a,n-a)} \quad (11)$$

Já a estatística Q monitora o subespaço residual e são definidas como a distância euclidiana de um registro em Y e sua respectiva projeção no espaço vetorial \hat{Y} , conforme mostrado abaixo na Equação 12:

$$Q = |y - \hat{y}|^2 \quad (12)$$

Sendo assim, por monitorar o subespaço residual, a estatística Q traduz a variabilidade global não detectada pelo modelo PCA associadas a variações não explicadas pelo processo. Portanto, ao monitorar Q , quando seu valor aumenta em novas observações, sugere-se uma mudança nas relações entre as variáveis que foram treinadas no modelo de PCA, ou seja, o modelo descrito pela PCA não inclui a direção dessa falha (KUNDU; KUNDU; DAMARLA , 2017). Para Li e Qui (2018), delimitar um limite de controle para a estatística Q significa demarcar uma hiperesfera no subconjunto residual. Um limite de controle muito utilizado para a estatística Q é mostrado na Equação 13 (KOTZ; JOHNSON; BOYD, 1967).

$$Q_{lim} = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{-\frac{1}{h_0}} \quad (13)$$

Sendo:

- $\theta_1 = \sum m_j = a + 1 \lambda_{ij}$;
- $h_0 = 1 - \left(\frac{2\theta_1 \theta_3}{3\theta_2^2} \right)$;
- c_α é o valor crítico do percentil $1 - \alpha$ da distribuição gaussiana.

2.7 MÉTODOS DE DIAGNÓSTICO DE FALHAS APLICADAS AO TEP

Diagnosticar uma falha significa indicar qual falha ocorreu no processo, além da sua extensão e localização, pois tratam-se de informações cruciais para agir corretivamente e propor métodos de prevenção efetivos. Dessa forma, percebe-se que o diagnóstico de falhas é considerado um problema de multi classificação. Para os métodos de diagnóstico de falhas em plantas industriais, a maioria das aplicações encontra-se baseada em séries históricas do processo (LAMEDA, 2015).

Sendo assim, os principais métodos de diagnóstico de falhas são multi classificadores que aplicam aprendizagem supervisionada baseadas em séries históricas. Dentre esses algoritmos, as redes neurais vêm recebendo um enorme destaque, devido a sua capacidade de aprendizagem e adaptação, até mesmo para sistemas dinâmicos e complexos (TAQVI *et al.*, 2021).

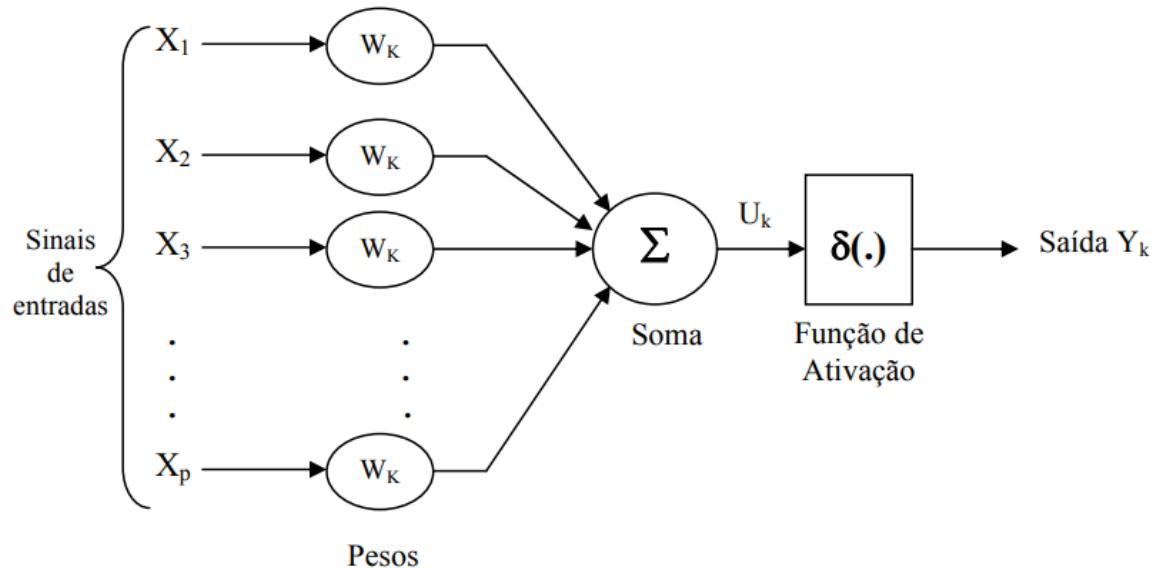
2.7.1 Redes neurais

Ao longo do tempo, os cientistas vêm estudando o cérebro humano para descobrir como funciona o raciocínio, as sensações e a nossa capacidade de se relacionar com outros indivíduos. A partir desse contexto, passou-se a analisar os neurônios e como eles transmitem informações entre si. A replicação desse estudo para os computadores deu origem ao modelo conhecido como rede neural. Sendo assim, essa tecnologia, inspirada no cérebro humano, é capaz de analisar uma grande quantidade de dados para descobrir padrões sem que haja uma interferência humana (BINI, 2018).

A arquitetura de uma rede neural é formada por neurônios que se distribuem em uma camada de entrada, uma camada de saída e várias camadas intermediárias. A camada de

entrada é responsável pela captação dos sinais ou das variáveis de entrada, as camadas intermediárias são chamadas de ocultas. A camada oculta é responsável por atribuir pesos aos sinais de entrada, assim, cada neurônio pode tratar o impulso de maneira diferente, podendo esses pesos ser negativos ou positivos. Além disso, há uma função de ativação que soma todos os pesos e que tem como objetivo limitar a saída e atribuir uma não linearidade ao modelo. Por fim, há uma ou mais saídas que podem classificar ou prever algum resultado. Um exemplo da arquitetura de uma rede neural convencional é mostrado a seguir na Figura 7 (ABIODUN, 2018).

Figura 7 - Arquitetura de uma rede neural convencional.



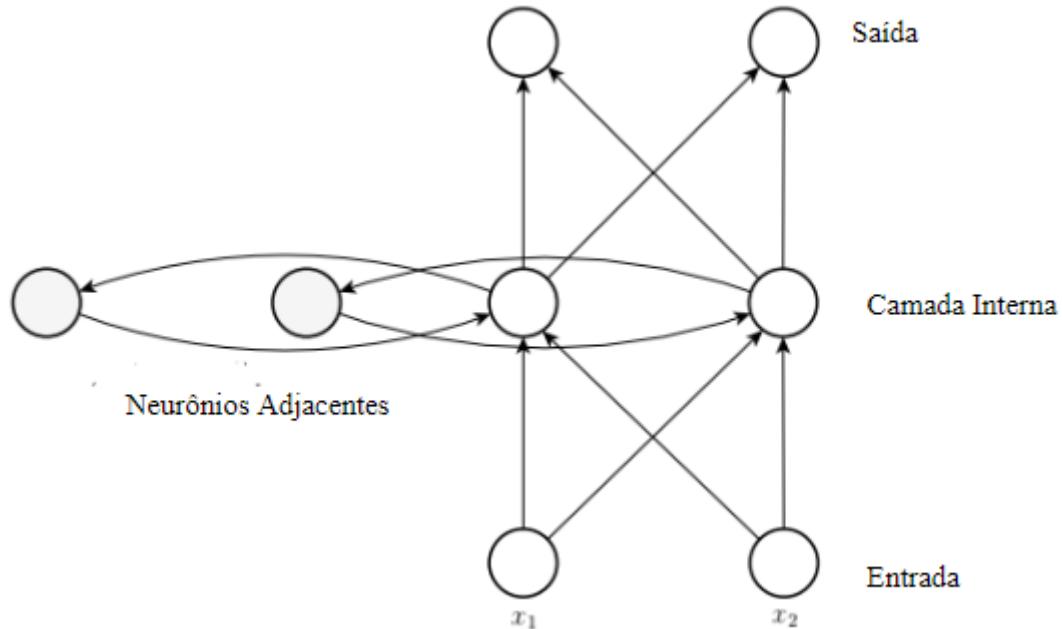
Fonte – Bianchini (2001).

Quando uma rede neural possui mais de uma camada oculta, podendo chegar à casa dos milhões, o treinamento do modelo tende a ficar mais custoso, atribuindo-se o nome de *deep learning*. Assim, pode-se afirmar que o *deep learning* é uma ferramenta bastante poderosa que complementa modelos tradicionais de *machine learning* em alguns casos. Possuem aplicações tanto nas áreas de processamento de linguagem natural como visão computacional, análise de *big data*, detecção de anomalias entre outros (CHEN; RAN, 2019).

Treinar um modelo de *deep learning* é computacionalmente custoso e requer uma grande memória do computador, tendo em vista os milhões de parâmetros que são iterados até estarem refinados o suficiente. Apesar disso, modelos de *deep learning* conseguem acurárias bastante altas e já superaram métodos tradicionais no que diz respeito ao reconhecimento de imagem (CHEN; RAN, 2019).

Uma evolução do modelo clássico de redes neurais artificiais para a aplicação em séries temporais são as redes neurais recorrentes (RNN), visto que, diferentemente das redes neurais convencionais (MLP) que transportam a informação em apenas uma direção, as RNNs trocam informações através da realimentação das saídas de um ou mais neurônios. Desta forma, os pesos serão alterados tanto por novas entradas, como também ponderando seus valores através dos estados passados, configurando uma espécie de memória na rede neural (STAUDEMAYER; MORRIS, 2019). Um exemplo de rede neural recorrente simples é mostrado na Figura 8. Cada neurônio da camada interna possui seu sinal de entrada e saída externa e, para além disso, recebem como entrada outros neurônios adjacentes.

Figura 8 - Arquitetura de uma rede neural recorrente.



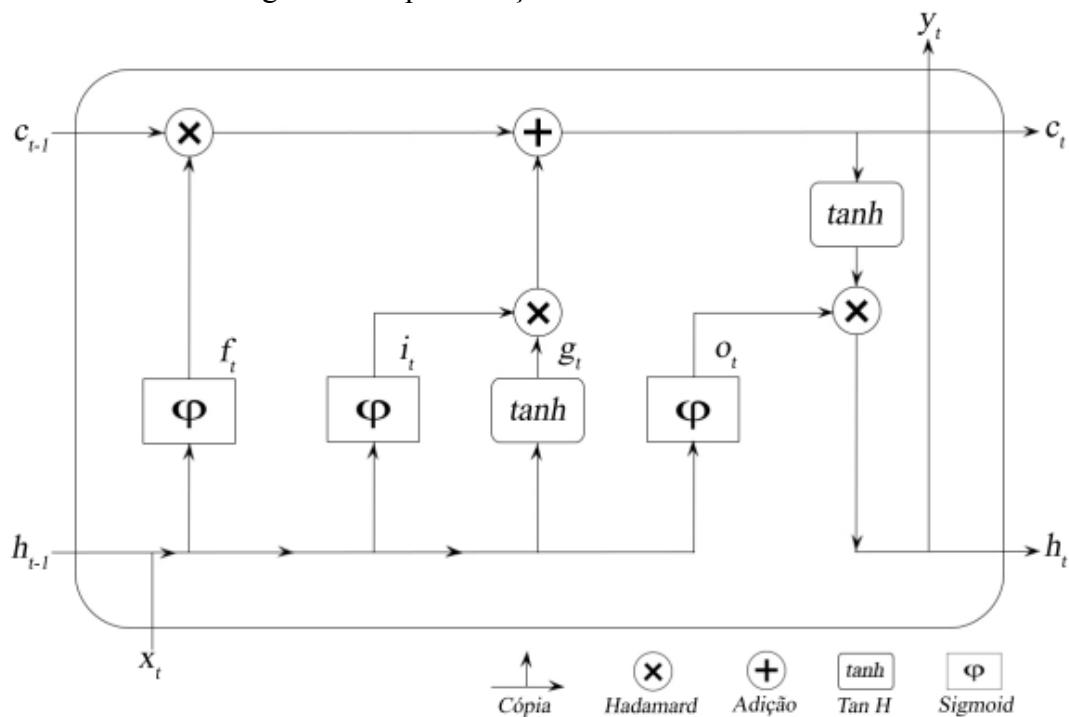
Fonte – Adaptado de Staudemeyer e Morris (2019).

Assim, as redes neurais recorrentes armazenam eventos recentes da entrada de outros neurônios, o que recebe o nome de memória de curto prazo. Para aplicações que não necessitam de uma influência muito grande de eventos passados essa técnica é muito útil, todavia, existem aplicações que necessitam de uma memória mais extensa, e para suprir essa necessidade, foram desenvolvidos algoritmos mais complexos capazes de armazenar uma maior quantidade de eventos. Entretanto, é importante citar que esses algoritmos mais complexos tornam o aprendizado mais lento (STAUDEMAYER; MORRIS, 2019). Existe,

portanto, o *trade-off* entre a capacidade de memorização da rede neural recorrente e a maior necessidade de capacidade computacional.

Para solucionar esse problema, foi proposta a LSTM (*Long Short-Term Memory*) que retém as informações de uma maior quantidade de entradas passadas, enquanto mantém a relevância de eventos recentes como maior. Para fazer isso, a LSTM faz uso dos estados denominados *cell state* e *hidden state* que são responsáveis pelo transporte de informações entre os neurônios das redes, como mostrado na Figura 9 (OLIVEIRA, 2020).

Figura 9 - Representação de uma célula LSTM.



Fonte – Oliveira (2020).

3 MATERIAIS E MÉTODOS

3.1 SOFTWARES

O ambiente utilizado para desenvolver o modelo foi o *jupyter notebook*, pois é uma *IDE open source*, ferramenta grátis que possui as principais bibliotecas utilizadas na ciência dos dados, a partir dele também é possível compilar trechos de código e inserir textos, facilitando, assim, a compreensão do código. Com essa ferramenta torna-se possível também plotar e visualizar tanto gráficos como tabelas.

Já a linguagem de programação utilizada foi o Python na versão 3.10.2, devido a sua facilidade de tratar, analisar e modelar os dados. As principais bibliotecas utilizadas neste trabalho foram a biblioteca *Pandas* na versão 1.2.4 e *Numpy* na versão 1.20.1 foram utilizadas para importar, manipular, tratar e realizar análise de dados. A biblioteca *Matplotlib* na versão 3.3.4 e *Seaborn* na versão 0.11.1 foram utilizadas para visualizar os dados através de gráficos e tabelas. Já a biblioteca *Scikit-Learn* na versão 0.24.1, *Scikit-Time* na versão 0.13.4 e *pytorch* na versão 1.7.7 serviram para criar, treinar e validar os modelos.

3.2 CONJUNTO DE DADOS

Os dados aplicados neste trabalho foram retirados de Rieth *et al.* (2017) e representam 500 simulações em diversas condições operacionais de uma planta petroquímica, chamada de *Tennessee Eastman Process* utilizados para produzir dois produtos, denotados genericamente por G e H.

Os dados foram divididos em treino e teste pelos autores Rieth *et al.* (2017). Sendo os dados de treino, o conjunto de amostras em operação normal da planta industrial, ou seja, não submetido a falhas, e os dados de teste, o conjunto de amostras submetidas a algum tipo de falha.

Ao todo, o processo apresenta 41 variáveis medidas, denotadas por XMEAS e mostradas na Tabelas 1 e 2 as quais trazem as variáveis contínuas e as medidas por amostragem, respectivamente, e 12 manipuladas, denotadas por XMV e apresentadas na Tabela 3. Os tempos de amostragem variam entre 3, 6 e 15 minutos.

Tabela 1 - Variáveis medidas de forma continuada no Tennessee Eastman.

Variável	Descrição	Unidade
XMEAS(1)	Entrada de A (corrente 1)	kscmh
XMEAS(2)	Entrada de D (corrente 2)	kg.h ⁻¹
XMEAS(3)	Entrada de E (corrente 3)	kg.h ⁻¹
XMEAS(4)	Entrada de total (corrente 4)	kscmh
XMEAS(5)	Corrente de reciclo (corrente 8)	kscmh
XMEAS(6)	Entrada do reator (corrente 6)	kscmh
XMEAS(7)	Pressão do reator	kpa
XMEAS(8)	Nível do reator %	%
XMEAS(9)	Temperatura do reator	°C
XMEAS(10)	Vazão de purga (corrente 9)	kscmh
XMEAS(11)	Temperatura do separador 0C	°C
XMEAS(12)	Nível do separador %	%
XMEAS(13)	Pressão do separador kPa	kPa
XMEAS(14)	Corrente de fundo do separador (corrente 10)	kg.h ⁻¹
XMEAS(15)	Nível do stripper	%
XMEAS(16)	Pressão do stripper	kPa
XMEAS(17)	Vazão de fundo do stripper	m ³ .h ⁻¹
XMEAS(18)	Temparatura do stripper	°C
XMEAS(19)	Vazão de vapor do stripper	kg.h ⁻¹
XMEAS(20)	Trabalho do compressor	kw
	Temperatura da saída da água de resfriamento do reator	
XMEAS(21)		°C
	Temperatura da saída da água de resfriamento do condensador	
XMEAS(22)		°C

Fonte: Adaptado de Rieth *et al.* (2017).

Tabela 2 - Variáveis medidas da composição de cada substância no Tennessee Eastman

Variável	Descrição	Corrente	Intervalo de amostragem (min)	Unidade
XMEAS(23)	Componente A	6	6	% molar
XMEAS(24)	Componente B	6	6	% molar
XMEAS(25)	Componente C	6	6	% molar
XMEAS(26)	Componente D	6	6	% molar
XMEAS(27)	Componente E	6	6	% molar
XMEAS(28)	Componente F	6	6	% molar
XMEAS(29)	Componente A	9	6	% molar
XMEAS(30)	Componente B	9	6	% molar
XMEAS(31)	Componente C	9	6	% molar
XMEAS(32)	Componente D	9	6	% molar
XMEAS(33)	Componente E	9	6	% molar
XMEAS(34)	Componente F	9	6	% molar
XMEAS(35)	Componente G	9	6	% molar
XMEAS(36)	Componente H	9	6	% molar
XMEAS(37)	Componente D	11	15	% molar
XMEAS(38)	Componente E	11	15	% molar
XMEAS(39)	Componente F	11	15	% molar
XMEAS(40)	Componente G	11	15	% molar
XMEAS(41)	Componente H	11	15	% molar

Fonte: Adaptado de Rieth *et al.* (2017).

Tabela 3 - Variáveis manipuladas do Tennessee Eastman

Variável	Descrição	Unidade
XMV(1)	Vazão de entrada de D (corrente 2)	kg.h ⁻¹
XMV(2)	Vazão de entrada de E (corrente 3)	kg.h ⁻¹
XMV(3)	Vazão de entrada de A (corrente 1)	kscmh
XMV(4)	Vazão de entrada de total (corrente 4)	kscmh
XMV(5)	Válvula de reciclo do compressor	%
XMV(6)	Válvula de purga (corrente 9)	%
	Vazão de saída de líquido do separador	
XMV(7)	(corrente 10)	m ³ .h ⁻¹
XMV(8)	Saída de líquido do stripper (corrente 11)	m ³ .h ⁻¹
XMV(9)	Válvula do vapor do stripper	%
XMV(10)	Vazão da água de resfriamento do reator	m ³ .h ⁻¹
	Vazão da água de resfriamento do	
XMV(11)	condensador	m ³ .h ⁻¹
XMV(12)	Velocidade do agitador	rpm

Fonte: Adaptado de Rieth *et al.* (2017).

Além disso, o processo foi submetido a 20 tipos de falhas que se diferenciam entre aumento de variabilidade das variáveis de processo por meio de perturbação do tipo degrau e randômica, agarramento de válvulas, variação na cinética do processo incluindo algumas falhas desconhecidas. As variáveis, a descrição da falha e o tipo da falha foram evidenciadas na Tabela 4.

Tabela 4 - Descrição das falhas a que o processo foi submetido.

Falha	Descrição	Tipos
IDV(1)	Razão de entrada A/C composição de B constante (corrente 4)	Degrau
IDV(2)	Composição de B, razão A/C constante (corrente 4)	Degrau
IDV(3)	Temperatura de entrada de D (corrente 2)	Degrau
IDV(4)	Temperatura de entrada da água de resfriamento do reator	Degrau
IDV(5)	Temperatura de entrada da água de resfriamento do condensador	Degrau
IDV(6)	Perda de entrada de A (corrente 1)	Degrau
IDV(7)	Queda de pressão de C, disponibilidade reduzida (corrente 4)	Degrau
IDV(8)	Composição da entrada de A, B e C (corrente 4)	Variação randômica
IDV(9)	Temperatura de entrada de D (corrente 2)	Variação randômica
IDV(10)	Temperatura de entrada de C (corrente 4)	Variação randômica
IDV(11)	Temperatura de entrada da água de resfriamento do reator	Variação randômica
IDV(12)	Temperatura de entrada da água de resfriamento do condensador	Variação randômica
IDV(13)	Cinética da reação	Desvio lento
IDV(14)	Válvula de água de resfriamento do reator	Agarramento
IDV(15)	Válvula de água de resfriamento do condensador	Agarramento
IDV(16)	Desconhecido	
IDV(17)	Desconhecido	
IDV(18)	Desconhecido	
IDV(19)	Desconhecido	
IDV(20)	Desconhecido	

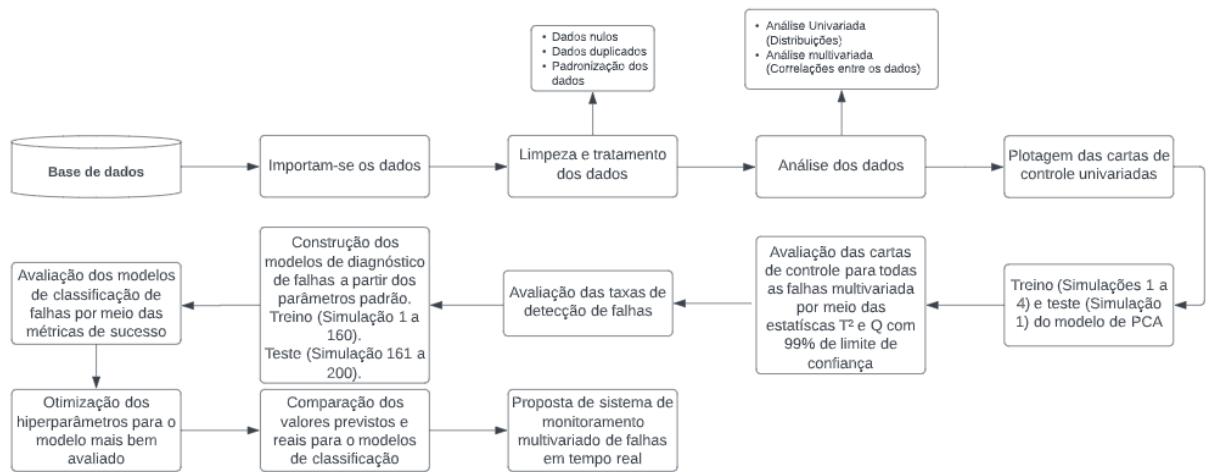
Fonte: Adaptado de Rieth *et al.* (2017).

3.3 METODOLOGIA

3.3.1 Fluxograma do trabalho

O fluxograma geral de como ocorreu o presente trabalho se encontra na Figura 10, no qual serão explicados com mais detalhes nos itens seguintes.

Figura 10 – Fluxograma do trabalho.



Fonte: o Autor.

3.3.2 Limpeza e tratamento dos dados

Nessa etapa, os dados foram importados por meio da biblioteca *pandas* e foram armazenados em um objeto chamado *DataFrame*, que é bastante similar a uma planilha do MS Excel com suas linhas e colunas.

Para evitar erros durante a modelagem, foi verificada a existência de dados duplicados, nulos ou faltantes. E uma vez que eles existam, estes serão apagados, pois não possuem informações relevantes para a análise do modelo.

Além disso, como os dados possuem dimensões diferentes, fez-se necessário a padronização dos dados para evitar efeitos de escala antes de inseri-los no modelo. Esse processo ocorreu conforme mostrado na equação abaixo.

$$z = \frac{X - \mu}{\sigma} \quad (14)$$

Sendo Z a medida padronizada, X um valor qualquer de determinada variável, μ e σ , a média e o desvio padrão do conjunto, respectivamente.

3.3.3 Análise dos dados

A fim de compreender melhor os dados, foram realizadas as suas análises descritivas de modo a determinar o tamanho do seu conjunto, como eles são classificados em relação ao seu tipo (número inteiro, número real, texto e etc.), e as principais medidas estatísticas como média, mediana, desvio padrão, quartis, mínimo e máximo. Além disso, foram realizadas outras 2 análises de modo a determinar o tipo de distribuição (isolada) do conjunto e como eles se relacionam (correlação entre os dados).

Na análise isolada, foi plotado a distribuição dos dados de modo a comparar os intervalos e a distribuição de grupos numéricos. Já para a análise relacional, foi determinada a correlação de Pearson entre as variáveis, isto é, um teste que mede a correlação linear estatística entre duas variáveis contínuas e varia de -1, para variáveis muito correlacionadas de forma inversamente proporcional, e +1, para variáveis muito correlacionadas diretamente proporcional. Essa análise foi visualizada a partir de um gráfico em estilo de mapa de calor.

3.3.4 Cartas de controle

Foram plotadas 53 cartas de controle, uma para cada variável, de modo se fazer a análise univariada do processo e verificar se o mesmo encontra-se devidamente controlado. Para isso, definiu-se a linha central como sendo a média μ do conjunto, com o processo controlado, ou seja, sem falhas e, também, definiu-se o limite superior de controle como sendo $\mu + 3\sigma$ e o limite inferior de controle como sendo $\mu - 3\sigma$, sendo μ a média e σ o desvio padrão.

Para observar o efeito das falhas no processo, foram plotadas, também, as 53 cartas de controle quando o processo foi submetido a uma falha do tipo perturbação do tipo randômica nas composições das substâncias A, B e C na corrente 4.

3.3.5 Modelagem

3.3.5.1 Detecção de Falhas

De início, foi verificado se as hipóteses para a aplicação da técnica de PCA foram cumpridas. Após isso, os dados foram divididos de operação normal da planta em treino e teste, de modo que foram usadas as simulações de 1 a 4 para treino do modelo PCA e as simulações 6 a 9 para o teste do modelo não submetido a falhas. Para o teste do modelo submetido a falhas de processo, foi filtrado a simulação de número 1 para as 20 falhas.

De posse desses dados, foi realizada uma transformação dos dados de treino, sob o comportamento natural do processo, ou seja, sem falhas, para o domínio das componentes principais, de modo a representar 90% da variância do conjunto de dados original.

Para a detecção de falhas, uma vez construído o modelo de estatístico multivariado, foi possível acompanhar o comportamento das amostras a partir das estatísticas T^2 e Q , em que o limite de detecção para T^2 e Q são dados pelas equações 11 e 13, respectivamente, considerando um limite de confiança de 99%. É importante salientar que pontos que ultrapassem o limite de controle são considerados como falhas do processo pelo modelo.

O monitoramento das falhas foi verificado no conjunto de dados em operação normal da planta, para examinar a taxa de falsos alarmes, assim como para a operação da planta industrial submetida a falhas, de modo a se examinar a taxa de alarmes positivos.

De modo a verificar quais variáveis sofreram mais alterações quando submetidos a uma falha, foi construído um gráfico do tipo “mapa de calor” para mostrar as contribuições de cada variável para o aumento da estatística Q .

Para acompanhar o desempenho do modelo como um todo na detecção das falhas, foi definida uma métrica chamada de taxa de detecção de falhas, que indica a quantidade de observações que são falhas e que de fato foram detectadas como falhas pelo modelo.

$$tx = \frac{\text{Falhas Previstas}}{\text{Falhas Totais}} \quad (15)$$

3.3.5.2 Diagnóstico de falhas

3.3.5.2.1 Demais modelos

Para a criação dos modelos de diagnóstico de falhas, foram utilizados os dados relacionados à operação industrial submetidos aos 20 tipos de falhas. Devido ao poder computacional limitado disponível para a construção dos modelos, foram considerados 200 simulações como um todo, sendo 140 simulações utilizadas para o treino e 60 para o teste, sendo estas divididas aleatoriamente, mas fixando o parâmetro `random_state = 20`, para manter a reproduzibilidade.

Antes de serem utilizados para treinar o modelo, os dados também foram submetidos a uma padronização, conforme mostrado na equação 14.

Além disso, por se tratar de modelos de classificação para séries históricas, os dados foram agrupados a partir de cada simulação, pois cada simulação representa uma série

histórica. Assim, de início, foram treinados e testados os seguintes modelos extraídos da biblioteca *Scikit-Time*, considerando seus hiperparâmetros padrão, os modelos: *RocketClassifier*, *Arsenal*, *BossEmsemble*, *ContractableBoss* e o *MUSE*.

As métricas de avaliação consideradas foram a acurácia média, precisão média, recall médio e f1-score médio. Foi considerada a média dessas métricas, pois o modelo é de multi classificação, ou seja, classifica 20 tipos de falhas e cada tipo de classificação possui sua própria métrica.

3.3.5.2.2 Rede Neural LSTM

A rede neural LSTM (*Long Short-Term Memory*) foi construída e testada utilizando o framework *pytorch*. Para a construção do modelo, foram utilizadas duas abordagens. A primeira sendo a entrada de dados considerada igual a escolhida para o desenvolvimento dos demais modelos, mas que encontrava-se convertida em tensores (estrutura de dados multidimensionais), que configura o formato de dado aceito pela biblioteca. Já na segunda os dados foram submetidos a uma transformação linear através da PCA, considerando uma quantidade de PCA's que representassem 90% da variância dos dados.

A otimização da rede neural foi realizada selecionando a combinação de diversos parâmetros para as duas abordagens, de modo a obter o valor máximo para a acurácia de teste, além disso, também foi monitorado as demais métricas, como já citadas no item 3.3.5.2.1. Os hiperparâmetros ajustados identificados foram:

- Número de camadas;
- Número de neurônios;
- Taxa de crescimento;
- Número de épocas.

Uma vez criado e otimizado o modelo, ele foi utilizado para predizer as falhas do conjunto de testes, sendo construída também uma matriz de confusão que compara o valor real da falha e o valor previsto.

3.3.6 Sistema de monitoramento em tempo real de falhas na indústria

Por fim, foi proposto um sistema de monitoramento de falha em tempo real baseado nas tecnologias utilizadas neste trabalho.

4 RESULTADOS E DISCUSSÃO

4.1 ANÁLISE E TRATAMENTO DE DADOS

A Figura 11 mostra uma visualização de parte dos dados de operação normal da fábrica. Esses dados foram armazenados em tipo específico de objeto da biblioteca pandas, chamado de *DataFrame*, que pode conter diversos tipos de dados e tem aspecto parecido com uma tabela de dados do MS Excel. Nota-se, também, que a primeira coluna é a numeração de dados por linha, facilitando o indexamento dos dados na tabela, e nas demais colunas encontram-se os dados principais.

Figura 11 – Visualização de parte dos dados não submetidos a falhas.

faultNumber	simulationRun	sample	xmeas_1	xmeas_2	xmeas_3	xmeas_4	xmeas_5	xmeas_6	xmeas_7	...	xmv_2	xmv_3	xmv_4	xmv_5	xmv_6	
0	0	1	1	0.25038	3674.0	4529.0	9.2320	26.889	42.402	2704.3	...	53.744	24.657	62.544	22.137	39.935
1	0	1	2	0.25109	3659.4	4556.6	9.4264	26.721	42.576	2705.0	...	53.414	24.588	59.259	22.084	40.176
2	0	1	3	0.25038	3660.3	4477.8	9.4426	26.875	42.070	2706.2	...	54.357	24.666	61.275	22.380	40.244
3	0	1	4	0.24977	3661.3	4512.1	9.4776	26.758	42.063	2707.2	...	53.946	24.725	59.856	22.277	40.257
4	0	1	5	0.29405	3679.0	4497.0	9.3381	26.889	42.650	2705.1	...	53.658	28.797	60.717	21.947	39.144
...	
249995	0	500	496	0.29325	3640.1	4473.0	9.1949	26.867	42.379	2700.2	...	53.429	29.249	60.773	21.532	40.451
249996	0	500	497	0.29134	3625.7	4506.2	9.2109	26.889	42.291	2700.6	...	53.830	28.975	61.517	21.750	42.762
249997	0	500	498	0.29438	3600.2	4478.3	9.1957	26.820	42.448	2700.3	...	54.163	28.676	61.656	21.487	42.109
249998	0	500	499	0.25269	3683.5	4486.4	9.2832	27.188	42.757	2697.4	...	53.453	24.889	61.564	21.392	39.334
249999	0	500	500	0.25214	3648.2	4467.8	9.1344	26.886	42.534	2695.1	...	53.676	24.943	61.254	21.208	38.991

250000 rows × 55 columns

Fonte: o Autor.

Já os dados de operação da planta submetidos a falhas foram obtidos a partir de 500 simulações para os 20 cenários de falhas, com 960 pontos de amostragens para as 53 variáveis de processo, totalizando 9.600.000 registros, conforme evidenciado na Figura 12.

Figura 12 – Visualização de parte dos dados submetido a falhas

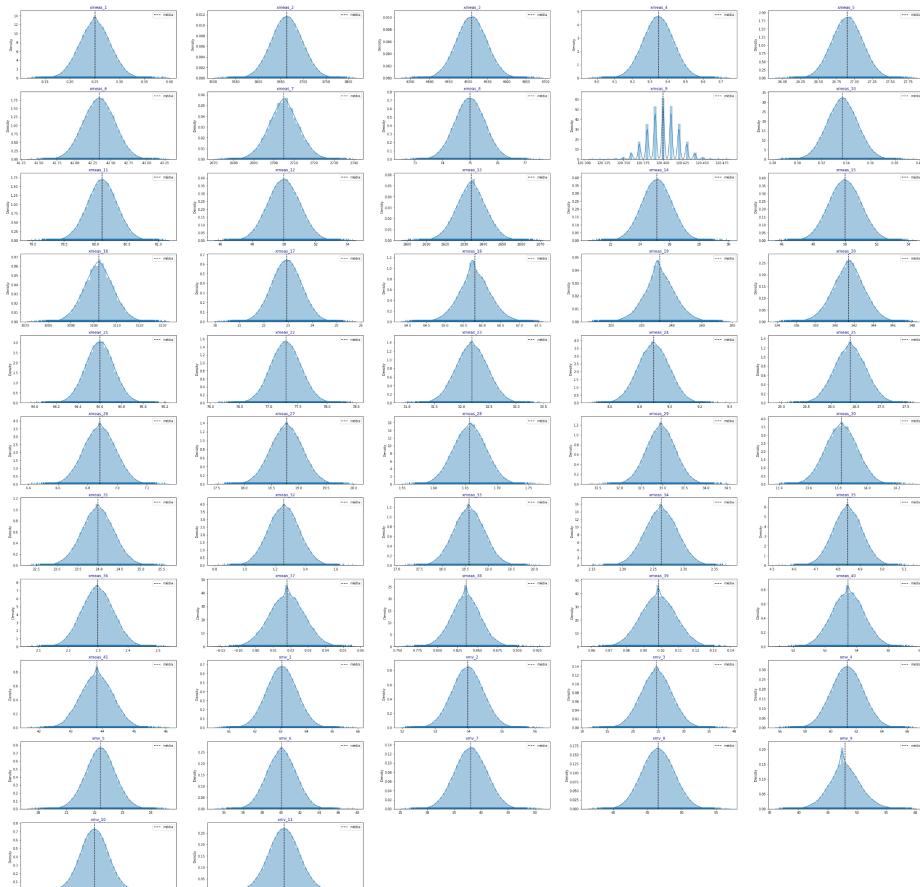
faultNumber	simulationRun	sample	xmeas_1	xmeas_2	xmeas_3	xmeas_4	xmeas_5	xmeas_6	xmeas_7	xmv_2	xmv_3	xmv_4	xmv_5
0	1	1	0.25171	3672.4	4466.3	9.5122	27.057	42.473	2705.6	54.494	24.527	59.710	22.357
1	1	1	0.25234	3642.2	4568.7	9.4145	26.999	42.586	2705.2	53.269	24.465	60.466	22.413
2	1	1	0.24840	3643.1	4507.5	9.2901	26.927	42.278	2703.5	54.000	24.860	60.642	22.199
3	1	1	0.25153	3628.3	4519.3	9.3347	26.999	42.330	2703.9	53.860	24.553	61.908	21.981
4	1	1	0.21763	3655.8	4571.0	9.3087	26.901	42.402	2707.7	53.307	21.775	61.891	22.412
...
9599995	20	500	956	0.26494	3719.6	4536.7	9.2265	26.875	42.401	2708.3	54.382	26.018	62.068
9599996	20	500	957	0.25252	3724.0	4494.3	9.1873	27.221	41.999	2706.0	54.236	25.098	61.557
9599997	20	500	958	0.25164	3700.8	4537.3	9.2514	26.659	42.180	2704.7	53.722	25.185	61.169
9599998	20	500	959	0.29097	3641.8	4525.2	9.3053	26.823	42.234	2705.0	54.185	28.771	61.140
9599999	20	500	960	0.29359	3622.2	4485.1	9.3070	26.922	42.610	2704.8	54.665	28.513	59.399

9600000 rows × 55 columns

Fonte: o Autor.

O conjunto de dados também não possui amostras com valores nulos ou duplicados que precisam ser removidos do conjunto. Também foi avaliado a distribuição dos dados ao longo do conjunto de treino, conforme mostrado na Figura 13.

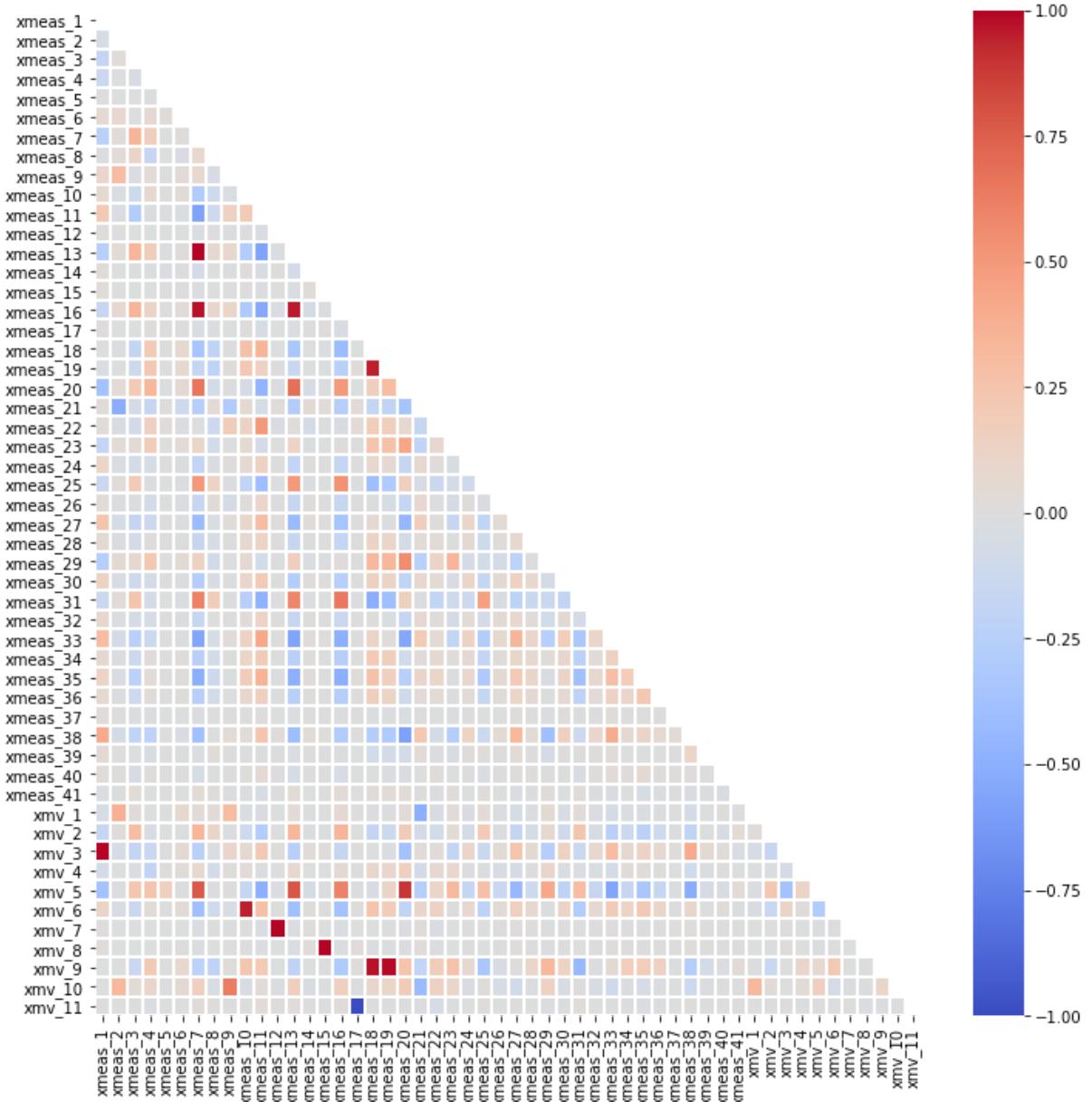
Figura 13 - Distribuição das variáveis de processo sob a condição normal.



Fonte: o Autor

Notou-se que os dados seguem uma distribuição normal, o que já era esperado, pois, segundo o teorema do limite central, quando o tamanho da amostra aumenta, a distribuição amostral de sua média aproxima-se cada vez mais de uma distribuição normal. Além disso, foi verificado um baixo desvio padrão na distribuição dos dados, o que também é esperado visto que o processo atua em uma malha fechada em um processo estacionário.

Figura 14 - Correlação entre os dados de treino.



Fonte: o Autor.

De acordo com o gráfico acima, nota-se:

- As variáveis de processo XM39, XM40 e XM41 não possuem correlação linear nenhuma com as outras;
- A maior parte das variáveis possuem uma correlação, em módulo, alta entre si;

Tais informações evidenciam que as variáveis têm impacto significativo nas demais.

Portanto, uma análise univariada, em que há o desprezo das relações entre as variáveis, não será a mais adequada, sendo necessário métodos de análise multivariada para monitorar essas variáveis.

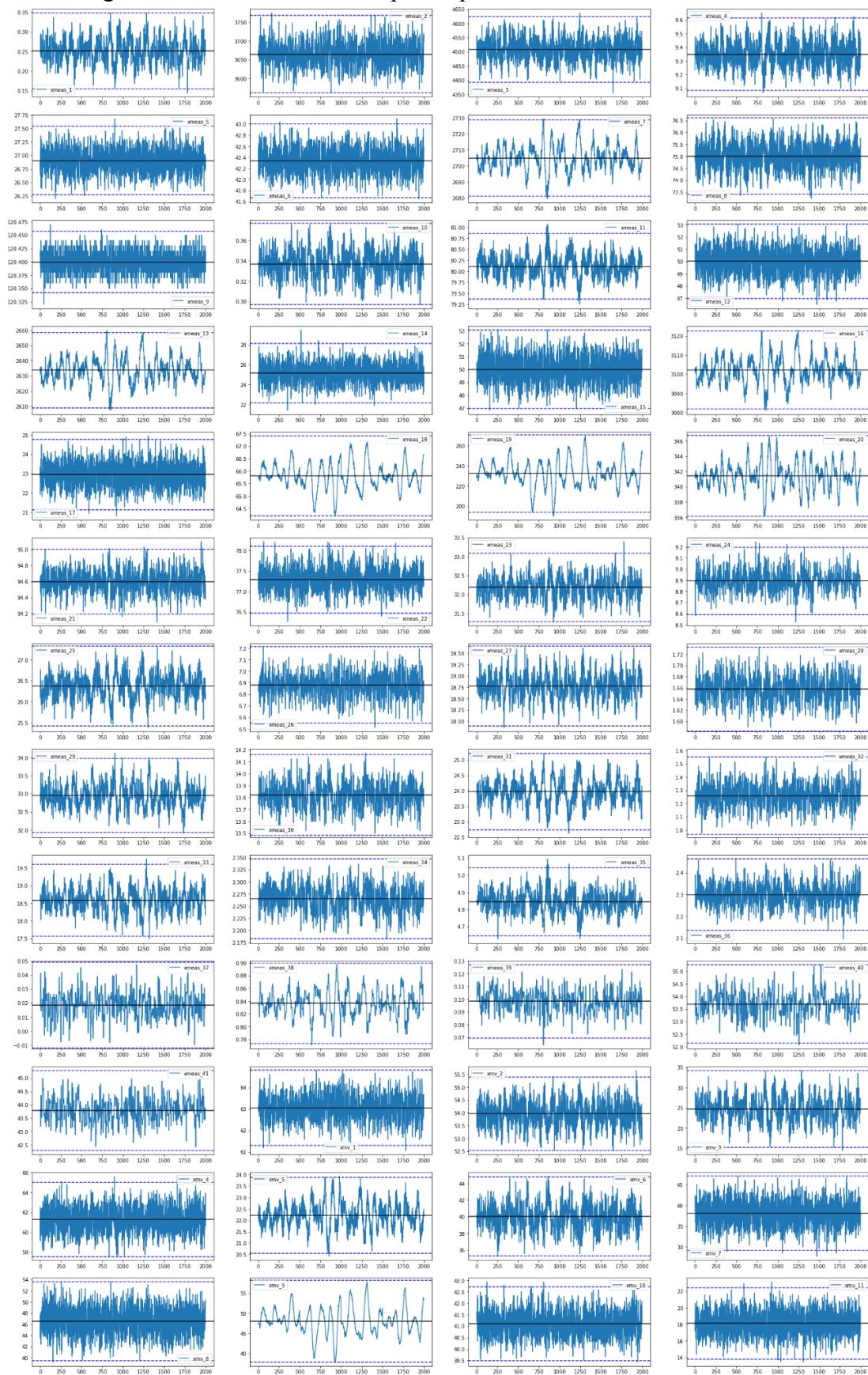
4.2 CONTROLE ESTATÍSTICO UNIVARIADO DO PROCESSO

A partir da análise de correlação, notamos que a análise univariada do processo não é o método mais adequado, porém, ele será analisado visto a simplicidade do método e por ser ainda bastante utilizado nas indústrias (Miranda *et al.*, 2019).

De início, foram analisadas as cartas de controle para todas as variáveis de processo. De acordo com a Figura 15, nota-se que a maior parte das variáveis se mantém no mesmo nível, havendo variações com pouca amplitude que pode ser explicada devido ao ruído do equipamento de medição. Entretanto, algumas outras, possuem um caráter dinâmico, havendo uma oscilação em torno da média, como em XMEAS_18 e XMEAS_19.

No geral, nota-se que o processo está bem controlado, pois a maior parte dos dados estão concentrados entre os limites inferiores e superiores.

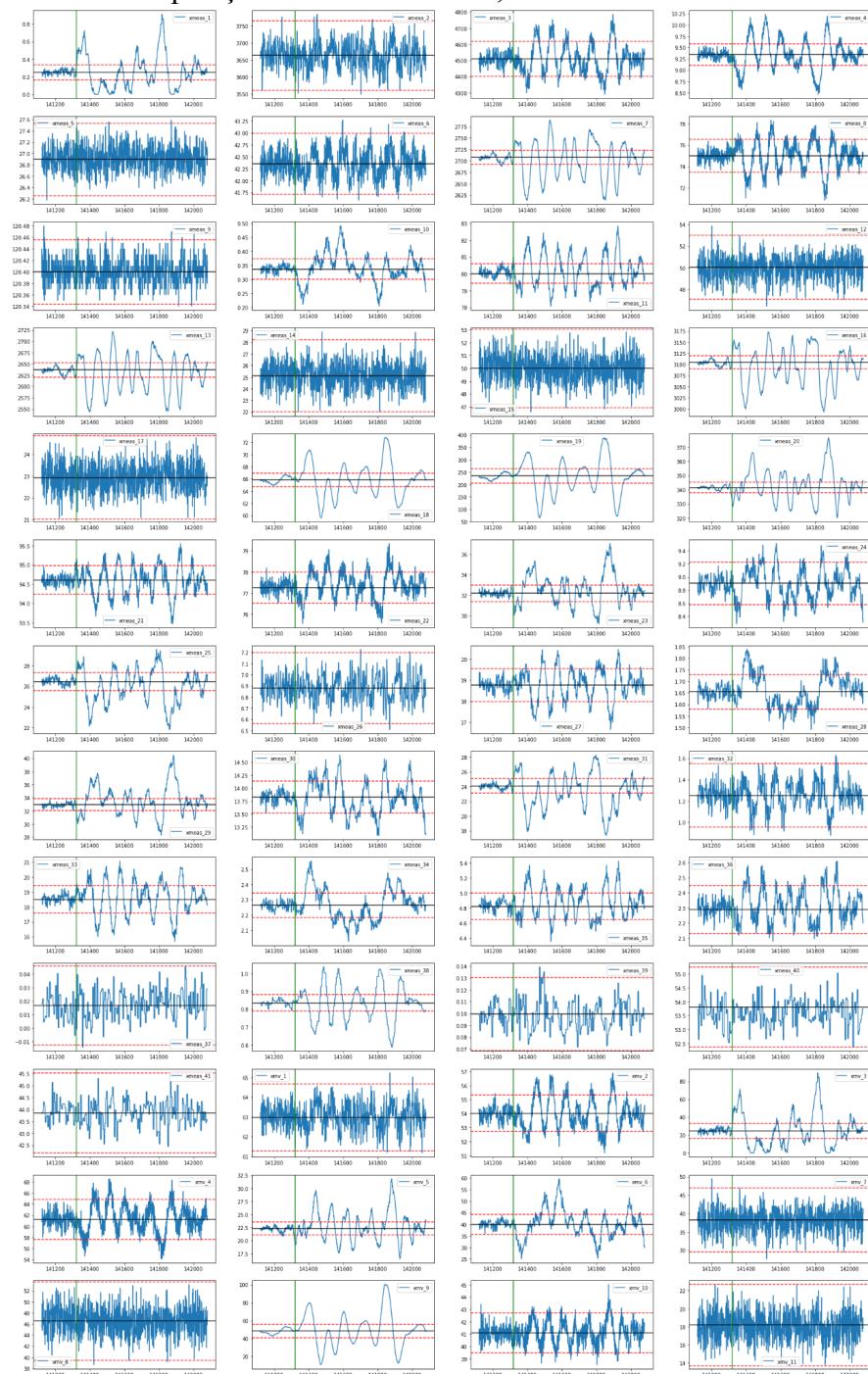
Figura 15 - Cartas de controle para o processo não submetido a falhas.



Fonte: o Autor.

No entanto, quando ocorre uma falha no processo, espera-se que haja uma perturbação na estabilidade do processo. Para ilustrar, nesse processo, foi realizada a simulação com uma perturbação do tipo randômica nas composições das substâncias A, B e C na corrente 4, falha IDV (8). Esse tipo de falha no processo o altera como um todo, pois modifica o seu balanço de massa. Tal fato fica evidente nas cartas de controle evidenciadas abaixo.

Figura 16 - Cartas de controle para o processo submetida a um erro do tipo randômico na composição das substâncias A, B e C no corrente 4.



Fonte: o Autor.

Embora as cartas de controle gerassem um alarme de modo satisfatório, a quantidade de gráficos a serem analisados é bastante grande, tornando inviável gerir o processo dessa maneira, principalmente quando o número de variáveis a serem analisadas aumentam com o tempo. Além disso, as cartas de controle não são capazes de analisar as relações entre as variáveis, evidenciando de maneira clara os alarmes falsos negativos.

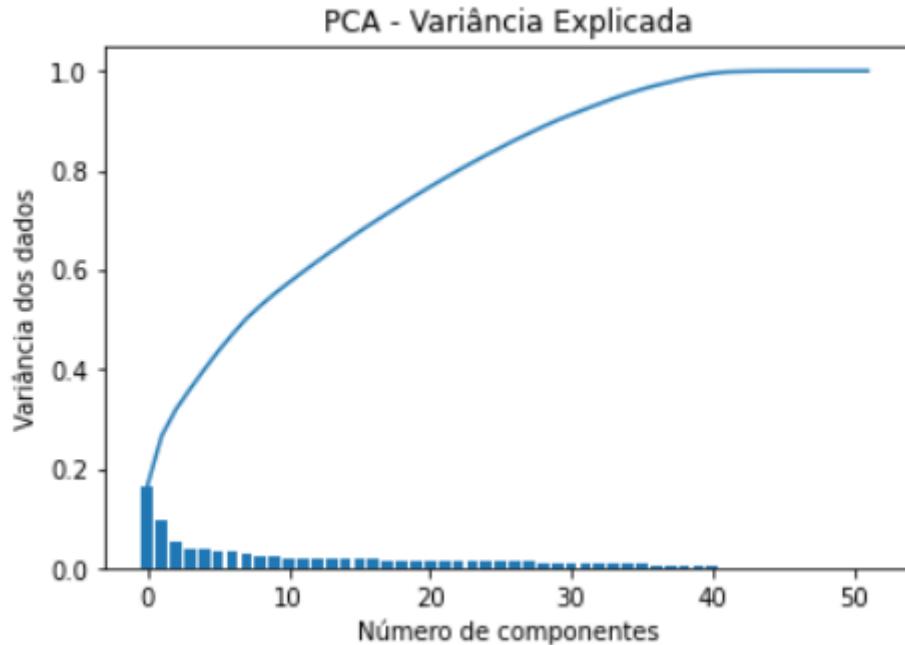
4.3 CONTROLE ESTATÍSTICO MULTIVARIADO DO PROCESSO A PARTIR DO MODELO DE DPCA

De início, foi verificado se as hipóteses para a aplicação do modelo DPCA foram atendidas:

- Linearidade: É possível observar relações lineares a partir da análise de correlação de Poisson, mostrado na Figura 14;
- Redundância ou multicolinearidade: Também presente a multicolinearidade, conforme mostrado na Figura 14;
- Alto razão de sinal/ruído: Para uma parte das variáveis, a razão sinal/ruído é alta, porém para outras, o ruído é predominantemente. Isso pode ser observado na Figura 15. Assim, isso pode vir a ser um problema na modelagem.
- Conjunto de dados centralizados: Todas as variáveis seguem uma distribuição normal, com a maior parte dos dados se concentrando na média. É possível observar isso analisando a Figura 13;
- Processo em estado estacionário: O processo se encontra no estado estacionário;
- Dados seguem uma distribuição normal: Também seguem uma distribuição normal, conforme mostrado na Figura 13.

Com as hipóteses sendo atendidas, inicia-se a construção do modelo. Assim, Para realizar o controle estatístico do processo, os dados da planta sob operação normal, ou seja, não submetido a falhas, foram submetidos a uma transformação linear por meio da PCA. Os componentes principais são ordenados a partir da variância explicada. A porcentagem de variância explicada por cada PCA, bem como a variância acumulada das PCAs são mostradas na Figura 17.

Figura 17 – Variância explicada por cada componente.



Fonte: o Autor.

De acordo com o gráfico acima, nota-se que poucas combinações lineares das variáveis são suficientes para explicar boa parte da variabilidade do processo. Por exemplo, as 10 primeiras PCAs já explicam mais de 50% de toda a variância. Isso já era esperado, uma vez que boa parte dos dados estão correlacionados entre si, como pode ser observado na Figura 14.

Para capturar 90% da variância dos dados, serão necessários 30 componentes. Assim, das 53 variáveis originais, 30 serão utilizadas para modelar o subespaço principal e 23, para modelar o subespaço residual.

Para verificar a taxa de falsos alarmes no processo, foi plotado as cartas de controle para a operação normal da planta. O resultado é mostrado na Figura 18.

Figura 18 – Estatísticas T^2 e Q para o processo não submetido à falhas.

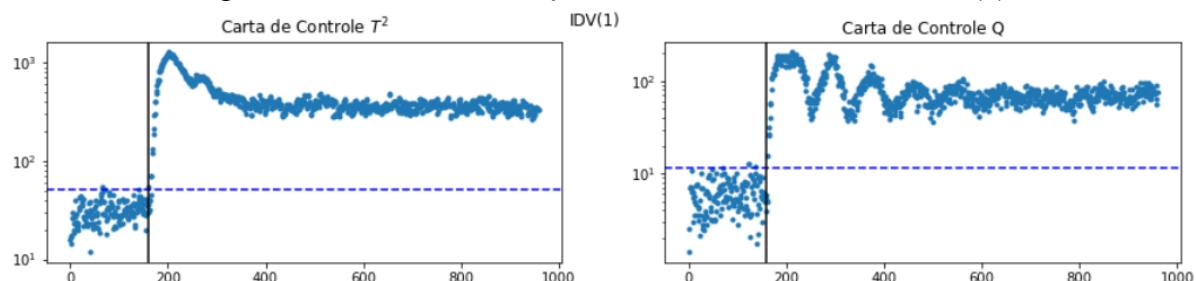


Fonte: o Autor.

A taxa de alarmes falsos foi de 0,85% monitorando a estatística T^2 e 2,7% para a estatística Q. O limite de confiança nos dois casos é de 99% para o limite de controle, ou seja, espera-se uma taxa de alarmes falsos de no máximo 1%. A taxa de alarmes falsos ficou abaixo do esperado para a estatística T^2 e um pouco acima do esperado para a estatística Q.

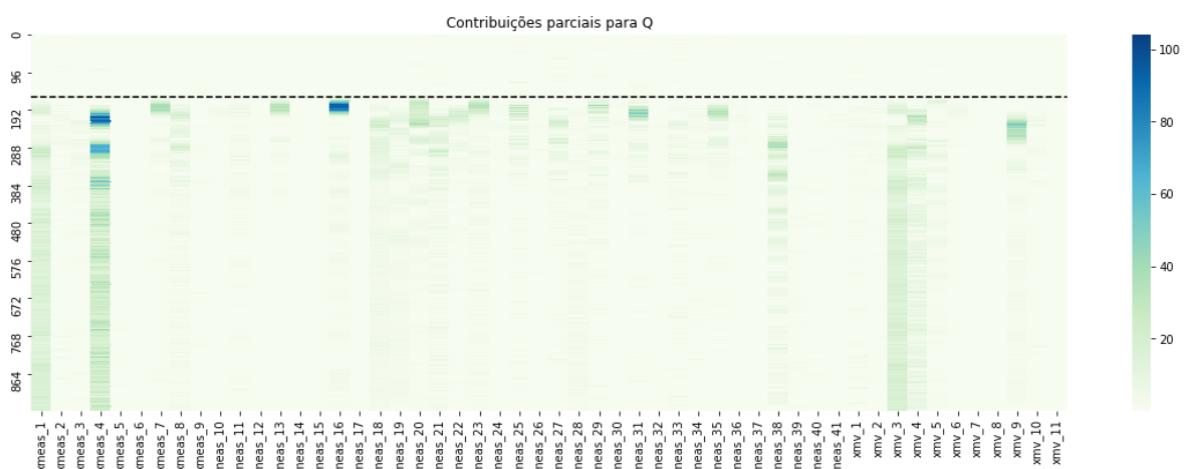
Para verificar se o modelo irá alarmar corretamente, foi verificado as cartas de controle multivariada, tanto para a estatística T^2 , como para a estatística Q, quando o sistema foi afetado pela falha IDV(1), perturbação degrau na razão de alimentação A/C na corrente 4. Além disso, foi plotado os gráficos de contribuições parciais para a estatística Q, de modo a identificar as variáveis mais afetadas pelas falhas.

Figura 19 – Estatísticas T^2 e Q para o processo submetido a uma falha do tipo perturbação degrau na razão de alimentação A/C na corrente 4, falha IDV(1).



Fonte: o Autor.

Figura 20 – Contribuições de cada variável para a estatística Q submetidas a falha IDV(1).



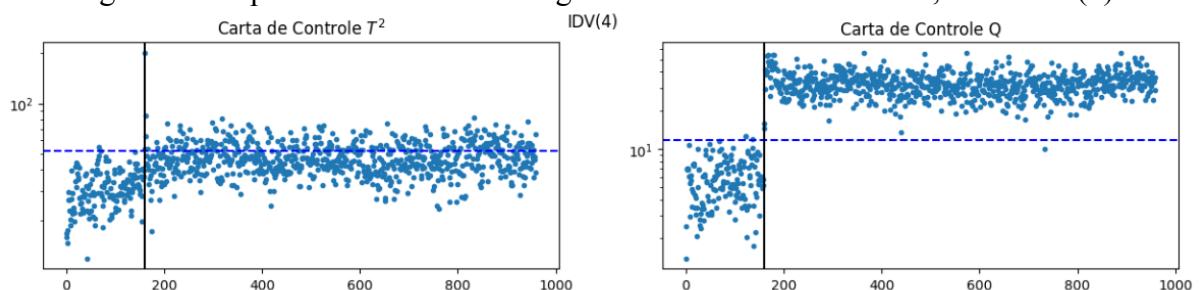
Fonte: o Autor.

De início, percebe-se uma vantagem muito positiva em analisar as cartas de controle multivariadas em relação às cartas de controle univariadas, visto que houve a síntese das informações em apenas 2 cartas.

Para a falha IDV(1), espera-se que a alimentação de A, na corrente de reciclo 5, reduza, e a própria malha de controle ajuste a alimentação de reciclo. O que, em geral, irá afetar boa parte das variáveis do processo, como é possível verificar na Figura 20. Portanto, esta falha é considerada de fácil detecção pelo modelo.

No entanto, para a falha IDV(4), torna-se mais difícil a detecção da falha, pois ela representa uma perturbação degrau positiva na temperatura da água de resfriamento do reator, o que tem um alcance mais restrito ao longo do processo.

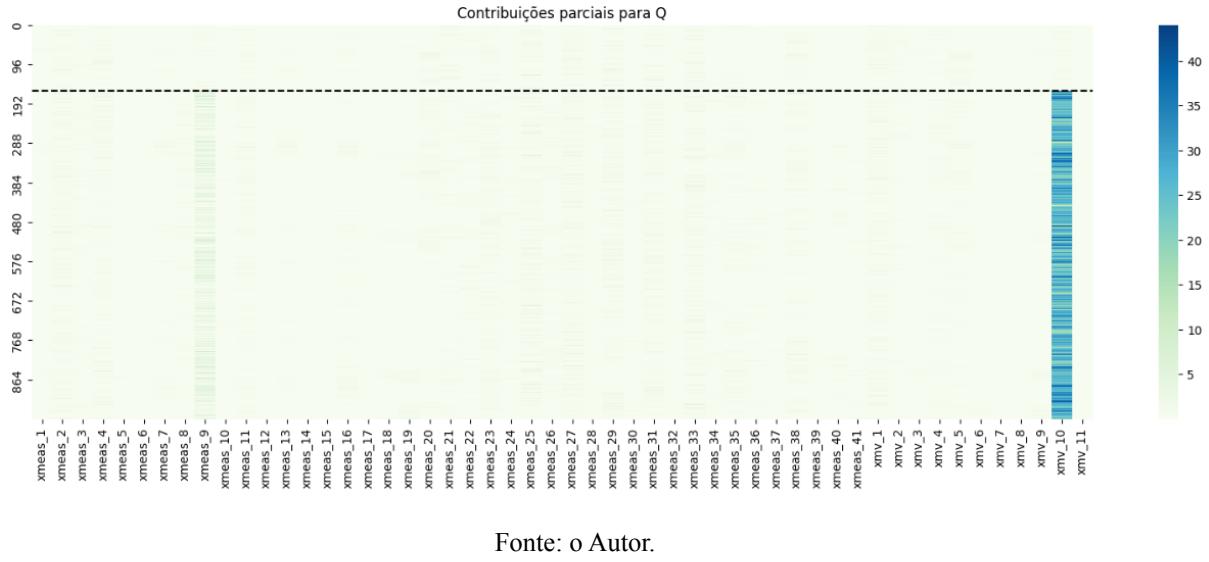
Figura 21 – Estatísticas T e Q para o processo submetido a uma falha do tipo perturbação degrau na temperatura de entrada da água de resfriamento do reator, falha IDV(4).



Fonte: o Autor.

Como é possível observar na Figura acima, apenas a estatística Q (Subespaço residual) foi capaz de detectar eficientemente a falha. Dessa forma, como a detecção da falha se encontrou no espaço residual, a falha não foi detectada pelas PCAs, mas sim por outras direções não identificadas pelo modelo PCA. Por isso, a importância de se avaliar o subespaço tanto principal como o residual.

Figura 22 – Contribuições de cada variável para a estatística Q submetidas a falha IDV(4).



Fonte: o Autor.

Como era esperado, a partir da análise da Figura 22, a extensão da falha encontrou-se restrita na variável de XMV_10 que representa a vazão de água de resfriamento do reator e foi utilizada pela malha fechada para amenizar os efeitos das falhas.

Por fim, foi verificada a taxa de alarmes quando o processo foi submetido a todas as falhas. Da análise da Figura 23, verifica-se que, quando combinada a análise a partir das cartas de controle estatísticas T^2 e Q, o modelo detectou satisfatoriamente 13 das 20 falhas apresentadas.

Figura 23 – Taxas de detecção para todas as falhas do processo.

Falha	Taxa de Detecção de Falha T2	Taxa de Detecção de Falha Q
IDV(1)	0.99375	0.99750
IDV(2)	0.98750	0.98375
IDV(3)	0.00875	0.02750
IDV(4)	0.31000	1.00000
IDV(5)	0.17000	0.14000
IDV(6)	0.99125	1.00000
IDV(7)	1.00000	1.00000
IDV(8)	0.92125	0.86750
IDV(9)	0.01375	0.03250
IDV(10)	0.07875	0.22875
IDV(11)	0.48375	0.71125
IDV(12)	0.98750	0.95750
IDV(13)	0.90000	0.91500
IDV(14)	0.99125	0.99750
IDV(15)	0.00250	0.02750
IDV(16)	0.05125	0.28125
IDV(17)	0.79375	0.96875
IDV(18)	0.95625	0.96125
IDV(19)	0.10250	0.11125
IDV(20)	0.30625	0.55875

Fonte: o Autor.

4.4 MODELOS PARA A PREVISÃO DE DIAGNÓSTICO DE FALHAS

Para obter um *benchmark*, com o modelo de rede neural, foram treinados e testados modelos disponíveis na biblioteca *sktime*, utilizando os hiperparâmetros padrões de cada algoritmo. É importante destacar que todos os modelos se caracterizam como sendo do tipo classificadores de séries históricas. Assim, a entrada dos modelos caracteriza simulações com 500 amostras das 53 variáveis de processo, e saída como do tipo de falha a ser detectada. As

métricas de avaliação dos modelos baseados nos dados de teste estão mostrados na Figura abaixo.

Figura 24 – Métricas dos modelos de *Machine Learning* importados da biblioteca *sktime*.

Modelos	Precisão	Acurácia Média	Recall Médio	f1-score Médio
RocketClassifier	0.59	0.57	0.53	0.54
Arsenal	0.60	0.60	0.55	0.56
BOSSEnsemble	0.64	0.59	0.54	0.57
ContractableBOSS	0.52	0.46	0.50	0.55
MUSE	0.60	0.57	0.56	0.58

Fonte: o Autor.

De acordo com a Figura 24 acima, os 5 tipos de modelos não se ajustaram tão bem aos dados de teste, pois a acurácia média variou entre 0.46 e 0.60. Sendo assim, foi necessário utilizar um modelo mais robusto e com melhores opções de ajuste de hiperparâmetros. Para isso, foi utilizado o modelo de rede neural recorrente LSTM.

Para esse modelo, utilizou-se duas abordagens. A primeira, o tipo de entrada e saída dos dados foram os mesmos utilizados para os modelos passados, visto que eles possuem as características de classificação para séries temporais. Já na segunda, esses dados foram submetidos a uma transformação linear através da técnica de PCA, considerando uma variância acumulada de 90%, totalizando PCA's. Os modelos LSTM foram treinados variando seus hiperparâmetros de modo a otimizar a acurácia para o teste, conforme evidencia a Tabela 5 abaixo:

Tabela 5 - Hiperparâmetros em que foi treinado a rede neural LSTM.

Parâmetro	Valores
N_epochs	10, 100, 150, 250,500
learning_rate	0.001, 0.0001
n_layers	3, 4
n_hidden	32, 64 ,128
batch_size	20
dropout	0.1

Fonte: o Autor.

O melhor conjunto de hiperparâmetros foi obtido para $n_layers = 3$, $n_hidden = 64$, $dropout = 0.1$, $batch_size = 20$, $learning_rate = 0.001$ e $N_epochs=500$ para ambos os cenários. As métricas de classificação para cada falha, bem como a média geral das métricas para ambos cenários estão mostradas na Figura 25 e a comparação entre a falha real e a falha prevista estão evidenciados na Figura 26.

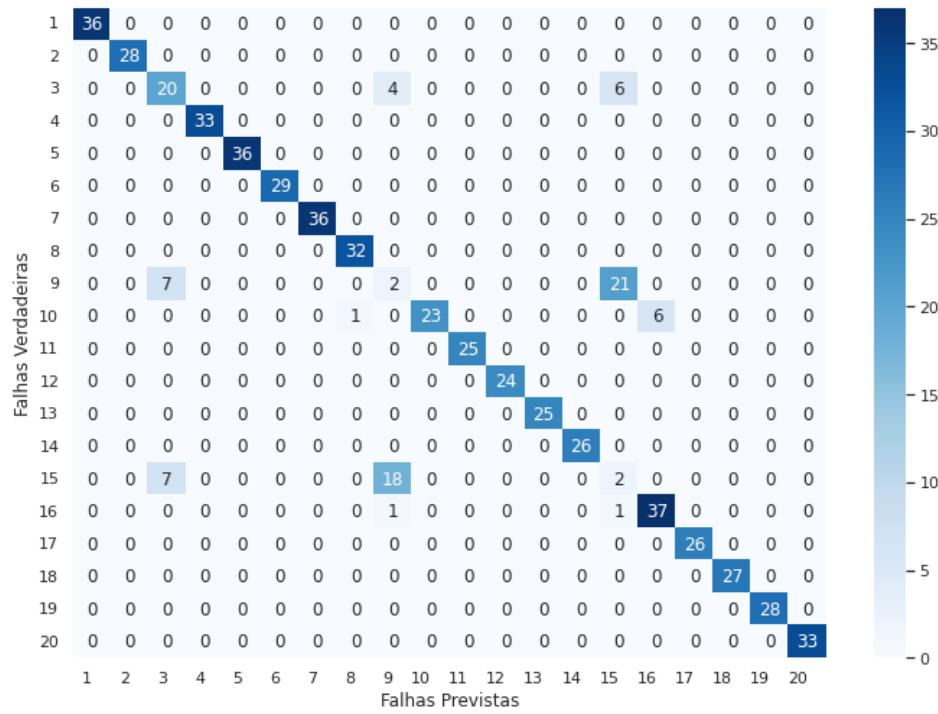
Figura 25 – Métricas dos modelos LSTM. Modelo LSTM com a entrada dos dados padrão (a), Modelo LSTM para os dados que sofreram uma transformação por meio da PCA (b).

	precision	recall	f1-score		precision	recall	f1-score	
1	1.00	1.00	1.00		1	1.00	1.00	1.00
2	1.00	1.00	1.00		2	1.00	1.00	1.00
3	0.59	0.67	0.62		3	0.98	0.84	0.91
4	1.00	1.00	1.00		4	1.00	1.00	1.00
5	1.00	1.00	1.00		5	1.00	1.00	1.00
6	1.00	1.00	1.00		6	1.00	1.00	1.00
7	1.00	1.00	1.00		7	1.00	1.00	1.00
8	0.97	1.00	0.98		8	1.00	1.00	1.00
9	0.08	0.07	0.07		9	0.51	0.74	0.60
10	1.00	0.77	0.87		10	0.89	0.82	0.85
11	1.00	1.00	1.00		11	1.00	1.00	1.00
12	1.00	1.00	1.00		12	1.00	1.00	1.00
13	1.00	1.00	1.00		13	1.00	1.00	1.00
14	1.00	1.00	1.00		14	1.00	1.00	1.00
15	0.07	0.07	0.07		15	0.50	0.39	0.44
16	0.86	0.95	0.90		16	0.84	0.86	0.85
17	1.00	1.00	1.00		17	1.00	1.00	1.00
18	1.00	1.00	1.00		18	1.00	1.00	1.00
19	1.00	1.00	1.00		19	1.00	1.00	1.00
20	1.00	1.00	1.00		20	1.00	1.00	1.00
accuracy	0.88	0.88	0.88		accuracy	0.94	0.94	0.94
macro avg	0.88	0.88	0.88		macro avg	0.94	0.93	0.93
weighted avg	0.88	0.88	0.88		weighted avg	0.94	0.94	0.94

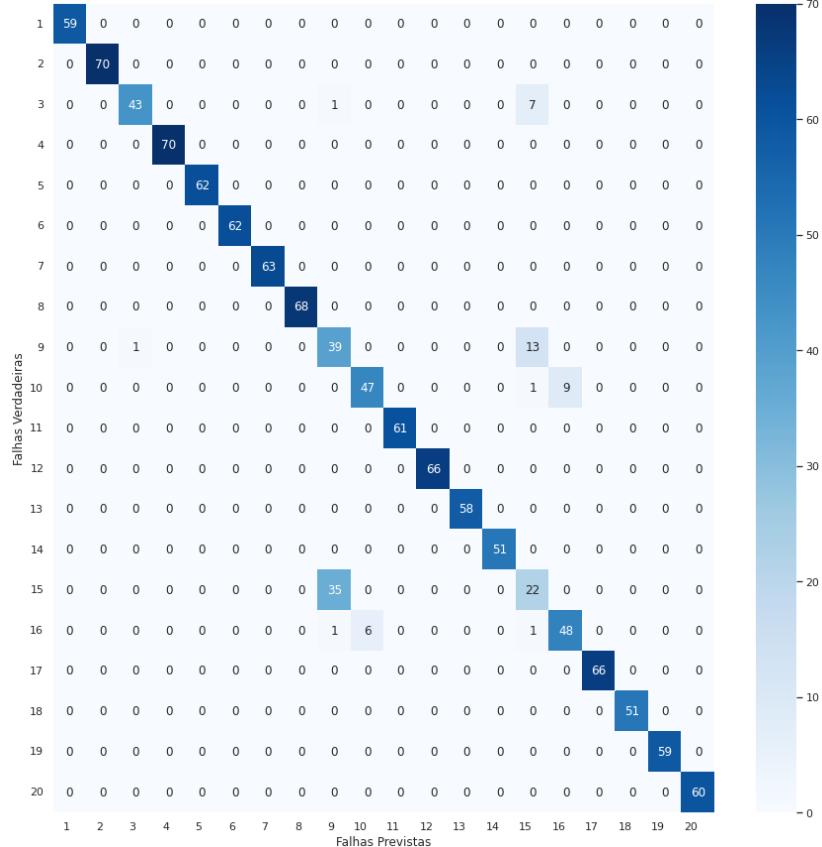
Fonte: o Autor.

Figura 26 – Previsões dos modelos LSTM. Modelo LSTM com a entrada dos dados padrão (a), Modelo LSTM para os dados que sofreram uma transformação por meio da PCA (b).

(a)



(b)



Fonte: o Autor.

Em comparação aos modelos mais simples, extraídos da biblioteca *sktime*, ambos os modelos de rede neural obtiveram um desempenho bastante superior. Isso pode ser explicado devido à capacidade de personalização do modelo, uma vez que devido à quantidade de dados e pela complexidade destes, é necessário testar um número grande de hiperparâmetros para se alcançar o modelo ideal.

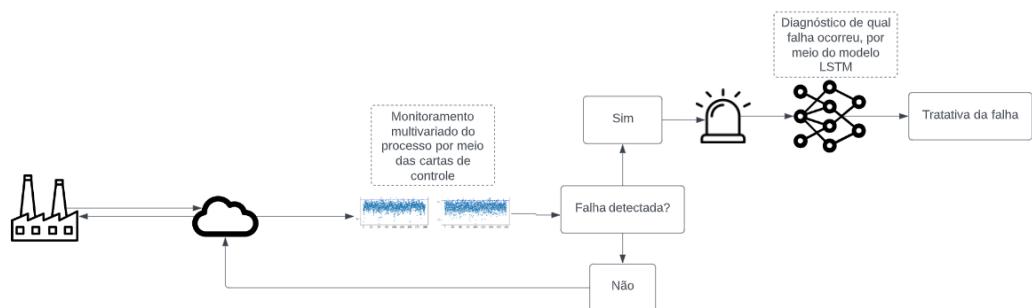
Quando comparado o desempenho dos modelos LSTM, com e sem PCA, a combinação dos modelos PCA e LSTM conseguiu uma acurácia global 6,8% maior, predizendo satisfatoriamente, com uma acurácia acima de 85%, 17 das 20 falhas. Inclusive, o modelo PCA combinado com LSTM aumentou a acurácia consideravelmente nas falhas mais difíceis de serem diagnosticadas apenas pelo modelo LSTM, que foram as falhas 9 e 15.

Tal fato pode ser explicado devido a técnica de PCA, que funcionou com uma espécie de *feature selection*, pois foi capaz de reduzir a dimensionalidade de 53 variáveis de processo para 30 componentes principais, o que tornou mais fácil o processo de aprendizagem da rede neural e aumentou a sua capacidade de generalização.

4.5 SISTEMA PROPOSTO PARA MONITORAMENTO DE FALHAS EM TEMPO REAL

No sistema proposto, os dados da indústria são adquiridos de diversas formas e são armazenados em tempo real na nuvem. Essas informações sofrem uma transformação linear para o domínio das PCAs e são monitoradas através das 2 cartas de controle multivariadas sob a ótica da estatística T^2 e Q . Uma vez detectada a falha, o sistema aciona um alarme e a série histórica desses dados é enviada para o modelo LSTM, que realiza as transformações necessárias e indica o tipo provável de falha que ocorreu. Todo esse processo é mostrado na Figura 27.

Figura 27 – Sistema proposto para monitoramento em tempo real de falhas na indústria.



Fonte: o Autor.

Uma vez que é detectada e diagnosticada essa falha atrativa e o desenvolvimento de métodos de prevenção se tornam mais eficientes, uma vez que todo o processo acontece em tempo real e ocorre a análise de todas as variáveis, ou seja, há uma análise global de toda a manufatura produtiva.

5 CONCLUSÕES

A indústria 4.0 vem transformando rapidamente o processo de automação, troca de informações, modelos de negócio e a produtividade das indústrias, uma vez que ela é responsável pelo aumento da eficiência no uso dos recursos no contexto industrial. Diante desse contexto, foram aplicados conceitos da indústria 4.0 para propor uma solução comum em engenharia de processos, a detecção e o diagnóstico de falhas na indústria.

Uma falha ocorre quando um controlador não consegue manter os níveis aceitáveis de uma variável, gerando uma anomalia no processo, que pode afetar as variáveis locais, bem como o processo como um todo. Inúmeras indústrias sofrem anualmente com gastos decorrentes de falhas que não foram detectadas no tempo correto.

Do ponto de vista operacional, a detecção e diagnóstico de falhas torna-se cada vez mais complexa nas indústrias, tendo em vista a quantidade de variáveis a serem monitoradas ao longo do processo. Tal fator torna inviável a análise univariada do processo pelo próprio indivíduo. Sendo assim, este trabalho recorreu-se a técnicas de *machine learning* para monitorar o processo, objetivando detectar e diagnosticar falhas de modo mais eficiente.

Para detectar as falhas, foi treinado um modelo de PCA que realizou uma transformação linear para um novo domínio ordenado pela variância dos dados. Desse modo, 90% da variância dos dados foi explicada por 30 componentes principais, reduzindo, assim, a complexidade da análise. Após isso, o processo foi monitorado através das cartas de controle utilizando as estatísticas T^2 e Q . Através desse mecanismo, 13 das 20 falhas foram monitoradas corretamente, com taxas de detecção positivas acima de 90%.

Para diagnosticar as falhas, observou-se que o modelo de rede neural LSTM, combinado à técnica de PCA, obteve o melhor desempenho, com uma acurácia média de 0.94, conseguindo diagnosticar satisfatoriamente 17 das 20 falhas apresentadas. Por fim, foi proposto um sistema de monitoramento em tempo real do processo, responsável por integrar as tecnologias de detecção e diagnóstico de falhas utilizadas neste trabalho.

REFERÊNCIAS

- ABIODUN, Oludare Isaac; JANTAN, Aman; OMOLARA, Abiodun Esther; DADA, Kemi Victoria; MOHAMED, Nachaat Abd Elatif; ARSHAD, Humaira. State-of-the-art in artificial neural network applications: A survey. **Heliyon**, v. 4, n. 11, p. e00938, 2018.
- ALPAYDIN, Ethem. **Machine learning**. MIT Press, 2021.259 p.
- AMIN, Md Tanjin; IMTIAZ, Syed; KHAN, Faisal. Process system fault detection and diagnosis using a hybrid technique. **Chemical Engineering Science**, v. 189, p. 191-211, 2018.
- ANNANTH, V. Kishorre; ABINASH, M.; RAO, Lokavarapu Bhaskara. Intelligent manufacturing in the context of industry 4.0: A case study of siemens industry. In: **Journal of Physics: Conference Series**. IOP Publishing, 2021. p. 012019.
- BATHELT, Andreas; RICKER, N. Lawrence; JELALI, Mohieddine. Revision of the Tennessee Eastman process model. **IFAC-PapersOnLine**, v. 48, n. 8, p. 309-314, 2015.
- BIANCHINI, Ângelo Rodrigo. Arquitetura de redes neurais para o reconhecimento facial baseado no neocognitron. 2001.
- BINI, Stefano A. Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?. **The Journal of arthroplasty**, v. 33, n. 8, p. 2358-2361, 2018.
- BORTOLINI, Marco; FERRARI, Emili; GAMBERI, Mauro; PILATI, Francesco; FACCIO, Maurizio. Assembly system design in the Industry 4.0 era: a general framework. **IFAC-PapersOnLine**, v. 50, n. 1, p. 5700-5705, 2017.
- BOTRE, Chiranjivi; MANSOURI, Majdi; KARIM, M. Nazmul; NOUNOU, Hazem; NOUNOU, Mohamed. Multiscale PLS-based GLRT for fault detection of chemical processes. **Journal of Loss Prevention in the Process Industries**, v. 46, p. 143-153, 2017.
- CAPACI, Francesca; VANHATALO, Erik; KULAHCI, Murat; BERGQUIST, Bjarne. The revised Tennessee Eastman process simulator as testbed for SPC and DoE methods. **Quality Engineering**, v. 31, n. 2, p. 212-229, 2019.
- CASTELLS, Manuel; ESPANHA, Rita. **A era da informação: economia, sociedade e cultura**. Fundação Calouste Gulbenkian. Serviço de Educação e Bolsas, 2007.
- CHEN, Jiasi; RAN, Xukan. Deep learning with edge computing: A review. **Proceedings of the IEEE**, v. 107, n. 8, p. 1655-1674, 2019.
- CHEN, Jinran; KHER, Shubha; SOMANI, Arun. Distributed fault detection of wireless sensor networks. In: **Proceedings of the 2006 workshop on Dependability issues in wireless ad hoc networks and sensor networks**. 2006. p. 65-72.

CHEN, Mu-Chen; HSU, Chun Chin; MALHOTRA, Bharat; TIWARI, Manoj Kumar. An efficient ICA-DW-SVDD fault detection and diagnosis method for non-Gaussian processes. **International Journal of Production Research**, v. 54, n. 17, p. 5208-5218, 2016.

CHIANG, Leo H.; RUSSELL, Evan L.; BRAATZ, Richard D. **Fault detection and diagnosis in industrial systems**. Springer Science & Business Media, 2000.

DA SILVA, Fernanda Aparecida Rocha; SIQUEIRA, Regiane Máximo. **Mineração de processos aplicada à gestão de negócios**: um mapeamento sistemático. Bauru: 08 nov. 2019.

DOWNS, James J.; VOGEL, Ernest F. A plant-wide industrial process control problem. **Computers & chemical engineering**, v. 17, n. 3, p. 245-255, 1993.

DOS SANTOS, Bruno Ferreira. **Uso de modelos de redes neurais artificiais para detecção de falhas no processo Tennessee Eastman**. 2018. Tese de Doutorado. PUC-Rio.

FERREIRA, K. B.; OLIVEIRA, A. G.G.; GONÇALVES, A. S.; GOMES, J. A.. Evaluation of hyperspectral imaging visible/near infrared spectroscopy as a forensic tool for automotive paint distinction. **Forensic Chemistry**, v. 5, p. 46-52, 2017.

FRANK, Paul M. Fault diagnosis for linear systems. **Control Systems, Robotics and AutomatioN–Volume XVI: Fault Analysis and Control**, p. 29, 2009.

GARCIA-ALVAREZ, Diego; FUENTE, M. J.; VEGA, P.; SAINZ, G.. Fault Detection and Diagnosis using Multivariate Statistical Techniques in a Wastewater Treatment Plant. **IFAC Proceedings Volumes**, v. 42, n. 11, p. 952-957, 2009.

GERMANO, Amanda Lucena. **Análise de desempenho de abordagens orientadas a fluxo de dados aplicadas à detecção de falhas de processos industriais**. 2017. Dissertação de Mestrado. Brasil.

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: 09 ago. 2022.

KOTZ, Samuel; JOHNSON, Norman L.; BOYD, D. W. Series representations of distributions of quadratic forms in normal variables. I. Central case. **The Annals of Mathematical Statistics**, v. 38, n. 3, p. 823-837, 1967.

KUNDU, Madhusree; KUNDU, Palash Kumar; DAMARLA, Seshu Kumar. **Chemometric Monitoring: Product Quality Assessment, Process Fault Detection, and Applications**. CRC Press, 2017.

LAMEDA, Carlos. Métodos relacionados con diagnósticos de fallas con síntomas imprecisos mediante comparación de casos. **Revista Digital de Investigación y Postgrado**, v. 5, n. 3, p. 3, 2015.

LI, Gang; QIN, S. Joe. Comparative study on monitoring schemes for non-Gaussian distributed processes. **Journal of Process Control**, v. 67, p. 69-82, 2018.

LOMOV, Ildar; LYUBIMOV, Mark; MAKAROV, Ilya; ZHUKOV, Leonid E.. Fault detection in Tennessee Eastman process with temporal deep learning models. **Journal of Industrial Information Integration**, v. 23, p. 100216, 2021.

MABKHOT, Mohammed M.; AL-AHMARI, Abdulrahman M.; SALAH, Bashir; ALKHALEFAH, Hisham. Requirements of the smart factory system: A survey and perspective. **Machines**, v. 6, n. 2, p. 23, 2018.

MAHESH, Batta. Machine learning algorithms-a review. **International Journal of Science and Research (IJSR). [Internet]**, v. 9, p. 381-386, 2020.

MEHMOOD, Rashid; RIAZ, Muhammad; LEE, Muhammad Hisyam; ALI, Iftikhar; GHARIB, Mona. Exact computational methods for univariate and multivariate control charts under runs rules. **Computers & Industrial Engineering**, v. 163, p. 107821, 2022.

MIRANDA, Ana Cláudia Leite; ALVES, Alessandro Ferreira; GOMES, Celso Augusto dos Santos; MOREIRA, Alessandro Messias; SILVA, Nancy Christiane Ferreira; MELO, Carina Adrielle Duarte de; CUNHA, Alberane Lúcio Thiago da . O controle estatístico de processos no monitoramento da fabricação em uma empresa no ramo colchoeiro. **Brazilian Journal of Development**, v. 5, n. 12, p. 29165-29185, 2019.

MOHAMED, Mamad. Challenges and benefits of Industry 4.0: an overview. **International Journal of Supply and Operations Management**, v. 5, n. 3, p. 256-265, 2018.

OLIVEIRA, Emerson Vilar de. **Análise de desempenho de método baseado em rede LSTM para classificação de falhas em um processo de controle de nível**. 2020. Dissertação de Mestrado. Universidade Federal do Rio Grande do Norte.

OLIVEIRA, Herbert Pereira de; ANUNCIAÇÃO, Wellington Ramos da; LOPES, Maria de Fátima Dos Santos. PROPOSTA DE MODELO SIMPLIFICADO PARA IMPLEMENTAÇÃO DO CONTROLE ESTATÍSTICO DE PROCESSO (CEP) NA INDÚSTRIA QUÍMICA/PETROQUÍMICA DE PROCESSO CONTÍNUO. **REVISTA DE TRABALHOS ACADÊMICOS-UNIVERSO SALVADOR**, v. 1, n. 5, 2018.

ONEL, Melis; KIESLICH, Chris A.; PISTIKOPOULOS, Efstratios N. A nonlinear support vector machine-based feature selection approach for fault detection and diagnosis: Application to the Tennessee Eastman process. **AIChE Journal**, v. 65, n. 3, p. 992-1005, 2019.

PARK, You-Jin; FAN, Shu-Kai S.; HSU, Chia-Yu. A review on fault detection and process diagnostics in industrial processes. **Processes**, v. 8, n. 9, p. 1123, 2020.

PERES, Ricardo Silva; JIA, Xiaodong; LEE, Jay; SUN, Keyi; COLOMBO, Armando Walter; BARATA, Jose. Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook. **IEEE Access**, v. 8, p. 220121-220139, 2020.

RIETH, Cory A.; Amsel, Ben D.; Tran, Randy; Cook, Maia B., 2017, "Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation", <https://doi.org/10.7910/DVN/6C3JR1>, **Harvard Dataverse**, V1.

RYDNING, David Reinsel-John Gantz-John. The digitization of the world from edge to core. **Framingham: International Data Corporation**, p. 16, 2018.

SANCHES, Marcelo Kaminski. **Aprendizado de máquina semi-supervisionado: proposta de um algoritmo para rotular exemplos a partir de poucos exemplos rotulados**. 2003. Tese de Doutorado. Universidade de São Paulo.

SHENFIELD, Alex; HOWARTH, Martin. A novel deep learning model for the detection and identification of rolling element-bearing faults. **Sensors**, v. 20, n. 18, p. 5112, 2020.

SHLENS, Jonathon. A tutorial on principal component analysis. **arXiv preprint arXiv:1404.1100**, 2014.

SIMANI, Silvio; FANTUZZI, Cesare; PATTON, Ronald Jon. Model-based fault diagnosis techniques. In: Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques. **Springer, London**, 2003. p. 19-60.

SOARES, Felipo Doval Rojas. **Técnicas de Machine Learning Aplicadas a Inferência e Detecção e Diagnóstico de Falhas de Processos Químicos Industriais em Contexto Big Data**. Universidade Federal do Rio de Janeiro, 2017. Tese (Mestrado).

STAUDEMAYER, Ralf C.; MORRIS, Eric Rothstein. Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. **arXiv preprint arXiv:1909.09586**, 2019.

SUMAN, Gaurav; PRAJAPATI, DeoRaj. Control chart applications in healthcare: a literature review. **International Journal of Metrology and Quality Engineering**, v. 9, p. 5, 2018.

SUN, Weike; PAIVA, Antonio R.C.; XU, Peng; SUNDARAM, Anantha; BRAATZ, Richard D.. Fault detection and identification using Bayesian recurrent neural networks. **Computers & Chemical Engineering**, v. 141, p. 106991, 2020.

TAQVI, Syed Ali Ammar; ZABIRI, Haslinda; TUFA, Lemma Dendena; UDDIN, Fahim; FATIMA, Syeda Anmol; MAULUD, Abdulhalim Shah. A review on data-driven learning approaches for fault detection and diagnosis in chemical processes. **ChemBioEng Reviews**, v. 8, n. 3, p. 239-259, 2021.

VENKATASUBRAMANIAN, Venkat; RENGASWAMY, Raghunathan; YIN, Kewen; KAVURI, Surya N.. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. **Computers & chemical engineering**, v. 27, n. 3, p. 293-311, 2003.