

PRISCILA BAYER DE OLIVEIRA SIMÕES

**Impactos de *Internet of Things* e *Data Streaming* na Manutenção 4.0**

São Paulo  
2022

PRISCILA BAYER DE OLIVEIRA SIMÕES

**Impactos de *Internet of Things* e *Data Streaming* na Manutenção 4.0**

**Versão Original**

**Monografia apresentada ao PECE –  
Programa de Educação Continuada em  
Engenharia da Escola Politécnica da  
Universidade de São Paulo como parte  
dos requisitos para a conclusão do  
curso de MBA em *Internet of Things*.**

**Orientador: Prof. Dr. Rogério Rossi**

São Paulo  
2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

## FICHA CATALOGRÁFICA

### Catalogação na publicação

Simões, Priscila

Impactos de Internet of Things e Data Streaming na Manutenção 4.0 / P.  
Simões -- São Paulo, 2022.  
60 p.

Monografia (MBA em MBA - Internet of Things (IoT)) - Escola Politécnica  
da Universidade de São Paulo. PECE – Programa de Educação Continuada em  
Engenharia.

1.Internet das Coisas 2.Indústria 4.0 3.Manutenção 4.Data streaming  
I.Universidade de São Paulo. Escola Politécnica. PECE – Programa de  
Educação Continuada em Engenharia II.t.

Nome: SIMÕES, Priscila Bayer de Oliveira

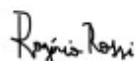
Título: Impactos de *Internet of Things* e *Data Streaming* na Manutenção 4.0

Monografia apresentada ao PECE – Programa de Educação Continuada em Engenharia da Escola Politécnica da Universidade de São Paulo como parte dos requisitos para a conclusão do curso de MBA em *Internet of Things*.

Aprovado em: **08/02/2022**

Banca Examinadora

Prof(a). Dr(a). Rogério Rossi



Instituição: USP-PECE

Julgamento: Aprovado

Prof(a). Dr(a). Carlos Eduardo Cugnasca



Instituição: PCS/EPUSP

Julgamento: Aprovado

Prof(a). Dr(a). Sérgio Roberto de Mello Canovas



Instituição: PECE

Julgamento: Aprovado

## **DEDICATÓRIA**

*Dedico este trabalho a meus pais e  
companheiro, que apoiaram e  
incentivaram minha jornada de  
estudos desde o início.*

## **AGRADECIMENTOS**

Agradeço aos meus pais que, desde cedo, apoiaram e incentivaram a minha curiosidade e gosto pelos estudos.

À Universidade de São Paulo – USP que, através de sua reputação consolidada e ensino respeitado internacionalmente, estimula o desenvolvimento do conhecimento e pesquisa essenciais à sociedade.

À Escola Politécnica da Universidade de São Paulo – EPUSP que proporcionou-me a oportunidade de obter valiosos conhecimentos e explorar novos caminhos.

Ao PECE – Programa de Educação Continuada em Engenharia que, através do programa de Extensão, possibilitou que vários de nós, alunos, tivéssemos o primeiro contato com a EPUSP.

Agradeço também ao Prof. Dr. Kichi Hirama pela organização do curso de MBA em IoT e por manter a excelente grade de disciplinas, apesar da pandemia do COVID-19 que está presente desde 2020.

Meus agradecimentos ao Prof. Dr. Rogério Rossi pela oportunidade de ter sido sua orientada. Agradeço pela dedicação, incentivo, orientações e por ter me apresentado a um fascinante universo de pesquisa relacionado a dados.

## RESUMO

SIMÕES, Priscila B. O. **Impactos de Internet of Things e Data Streaming na Manutenção 4.0.** 2022. P. 60. Monografia (MBA em *Internet of Things*). Programa de Educação Continuada em Engenharia da Escola Politécnica da Universidade de São Paulo. São Paulo. 2020.

Historicamente, em organizações tradicionais, ações e pesquisas relacionadas à manutenção não recebiam grande atenção, investimentos ou esforços. Costumava-se trazer o assunto à tona apenas na ocorrência de falhas, o que causava grande aumento nos custos operacionais em médio e longo prazo. Porém nota-se, nos últimos anos, maiores intenções de investimentos financeiros e iniciativas de pesquisa no assunto por conta das exigências de mercado relacionadas ao tempo e à qualidade dos produtos ofertados. Assim, a manutenção tem despertado maior interesse das organizações e das instituições de pesquisa. A maior complexidade dos equipamentos atuais e consequente elevação dos custos de manutenção é um dos fatores responsáveis pela mudança de foco da importância da manutenção nos mais diversos processos produtivos. Com o advento da Indústria 4.0 e a consequente utilização ubíqua de dados no funcionamento e gestão dos processos, desenvolveu-se a Manutenção 4.0. Esta possui como objetivo estabelecer ações de prevenção com base na comunicação entre dispositivos e análise de dados coletados em diversas fontes, a fim de evitar paradas em equipamentos devido a falhas e a queda na qualidade dos produtos derivados do processo de produção. Assim, o objetivo deste trabalho é analisar a utilização de *data streaming* em sistemas de Manutenção 4.0 apoiados pela IoT. Utilizou-se como método a busca de referências em bases de artigos científicos, livros especializados, sites de fornecedores de tecnologia e portais de notícias. Elaborou-se um comparativo entre estudos de caso de aplicação de *data streaming* na manutenção, em diferentes áreas. Conclui-se que é possível aplicar os princípios relacionados à Manutenção 4.0 além da manufatura, baseados em dados em tempo real.

Palavras-chave: Internet das Coisas, Indústria 4.0, manutenção, *data streaming*.

## ABSTRACT

SIMÕES, Priscila B. O. **Impactos de Internet of Things e Data Streaming na Manutenção 4.0.** 2022. P. 60. Monografia (MBA em *Internet of Things*). Programa de Educação Continuada em Engenharia da Escola Politécnica da Universidade de São Paulo. São Paulo. 2020.

Historically, in traditional organizations, actions and research related to maintenance did not receive much attention, investment or effort. It used to be brought up only in the event of failures, which caused a large increase in operating costs in medium and long terms. However, in recent years, more intentions of financial investments and research initiatives on the subject have been noted, due to market requirements related to time and the quality of the products offered. Thus, maintenance has got interest from organizations and research institutions. The greater complexity of current equipment and the consequent increase in maintenance costs is one of the factors responsible for the change in focus of the importance of maintenance in the most diverse production processes. With the advent of Industry 4.0 and the consequent ubiquitous use of data in the operation and management of processes, Maintenance 4.0 was developed. This aims to establish preventive actions based on communication between devices and data analysis, collected from different sources, in order to avoid equipment downtime due to failures and the drop in the quality of products derived from the production process. Thus, the objective of this work is to analyze the use of data streaming in Maintenance 4.0 systems supported by IoT. The method used was to search for references in scientific articles databases, specialized books, technology supplier and news websites. A comparison between case studies of data streaming application in maintenance, in different areas, was made. It is concluded that it is possible to apply the principles related to Maintenance 4.0 beyond manufacturing, based on real-time data.

**Keywords:** Internet of Things, Industry 4.0, maintenance, data streaming

## LISTA DE ILUSTRAÇÕES

<b>Figura 1 – Ciclo de vida dos dados .....</b>	<b>18</b>
<b>Figura 2 – Modelos de bancos de dados NoSQL.....</b>	<b>23</b>
<b>Figura 3 – Modelo proposto de ciclo de vida de dados para <i>Big Data</i>.....</b>	<b>28</b>
<b>Figura 4 – Evolução da manutenção .....</b>	<b>33</b>
<b>Figura 5 – Taxonomia das abordagens de manutenção preditiva .....</b>	<b>41</b>
<b>Figura 6 – Tipos de desvio .....</b>	<b>47</b>

## LISTA DE TABELAS

<b>Tabela 1 – Resultados da pesquisa bibliográfica.....</b>	<b>14</b>
<b>Tabela 2 – Comportamento do número de referências na linha do tempo .....</b>	<b>15</b>
<b>Tabela 3 – Evolução dos modelos de armazenamento de dados.....</b>	<b>21</b>
<b>Tabela 4 – Ciclo de vida de dados em IoT.....</b>	<b>27</b>
<b>Tabela 5 – Comparação entre processamentos em lote e <i>streaming</i>.....</b>	<b>29</b>
<b>Tabela 6 – Algoritmos de aprendizado supervisionado .....</b>	<b>45</b>
<b>Tabela 7 – Algoritmos de aprendizado não supervisionado .....</b>	<b>46</b>
<b>Tabela 8 – Exemplos de aplicação de <i>data streaming</i> em Manutenção 4.0 .....</b>	<b>49</b>

## SUMÁRIO

1.	INTRODUÇÃO .....	12
1.1	Motivações.....	13
1.2	Objetivo.....	13
1.3	Justificativas.....	14
1.4	Método de Pesquisa .....	16
1.5	Estrutura do Trabalho .....	16
2	GESTÃO DE DADOS E <i>INTERNET OF THINGS</i> .....	18
2.1	Ciclo de vida dos dados genérico .....	18
2.1.1	Produção .....	19
2.1.2	Armazenamento .....	20
2.1.3	Transformação .....	24
2.1.4	Análise.....	25
2.1.5	Descarte .....	26
2.2	Ciclo de vida dos dados aplicado a <i>Big Data</i> .....	27
2.3	Formas de processamento de dados.....	28
2.3.1	Processamento <i>batch</i> (em lote).....	28
2.3.2	<i>Streaming</i> .....	29
2.4	Considerações sobre o capítulo.....	30
3	MANUTENÇÃO: DO TRATAMENTO DE FALHAS À PREDIÇÃO .....	31
3.1	Manutenção tradicional: evolução histórica .....	31
3.2	Manutenção 4.0 .....	34
3.2.1	A função de <i>Internet of Things</i> na Manutenção 4.0 .....	35
3.3	Considerações sobre o capítulo.....	36
4	DATA STREAMING E A MANUTENÇÃO 4.0.....	38
4.1	A importância do <i>data streaming</i> na Manutenção 4.0 .....	38
4.2	Técnicas e ferramentas para extração de conhecimento do <i>data streaming</i>	
4.2	43	
4.3	Utilização de <i>data streaming</i> em aplicações de Manutenção 4.0 .....	48
4.4	Considerações sobre o capítulo.....	50
5	CONSIDERAÇÕES FINAIS.....	51
	REFERÊNCIAS .....	54

## 1. INTRODUÇÃO

*Internet of Things* (IoT) (do inglês, Internet das Coisas) é uma realidade em crescente expansão. De acordo com o relatório publicado em setembro de 2021 pelo IOT ANALYTICS (2021) (instituto alemão que fornece dados de mercado ou da indústria referentes à IoT), em 2020 existiam 11,3 bilhões de dispositivos conectados e a previsão para 2025 é que este número atinja 27,1 bilhões de dispositivos. O número real é ainda maior, já que a pesquisa não leva em conta a quantidade de sensores individuais existentes. Sinha (2021) descreve que o estudo se refere à quantidade de dispositivos que concentram sensores conectados via redes cabeadas, celulares, *Low Power Wide Area Network* (LPWAN) (do inglês, rede de baixa energia e longo alcance), *Wireless Personal Area Network* (WPAN) (do inglês, rede de área pessoal sem fio) e satélite.

Apesar de ser uma atividade que não costuma receber a devida atenção em indústrias baseadas em modelos tradicionais, a gestão de manutenção é responsável por processos que visam corrigir e evitar efeitos de falhas. De acordo com Nepomuceno (2014), historicamente, a manutenção assumia formatos meramente corretivos; ou seja, ações eram tomadas apenas em caso de ocorrência de falha ou incidente. Posteriormente foram introduzidos métodos proativos, tais como a programação periódica de revisão de equipamentos e substituição de componentes. Na atualidade, o foco da manutenção gradualmente progride para métodos preditivos e utiliza tecnologias avançadas como apoio.

Segundo relatório divulgado por Lueth (2020), a indústria era o campo que mais possuía projetos de IoT em 2020. A manutenção industrial compõe este campo; adotar processos baseados em IoT permite monitorar os equipamentos, coletar e armazenar dados e identificar oportunidades de melhorias a fim de evitar indisponibilidades no processo de produção. Interrupções podem ocorrer por diversos motivos, desde a substituição de peças em casos de falha até a programação periódica de atividades de manutenção. Conforme Stevan Jr *et al.* (2018), este monitoramento pode ser apoiado por processos baseados em dados e trazer benefícios para a operação e para o negócio.

De acordo com os números divulgados por IOT ANALYTICS (2021a), a manutenção preditiva passa por um processo de expansão nos últimos cinco anos. Em 2016 era adotada em pequenos nichos, poucos dispositivos eram conectados e o retorno sobre o investimento era desconhecido. O potencial crescimento previsto para 2025 é que a manutenção preditiva seja ampliada para modelos mais complexos, acessíveis a diferentes perfis de usuários e formada por conexões mais complexas entre dispositivos. Além disso, Bruegge (2021) afirma que o investimento financeiro passará de US\$ 1,5 bilhões a US\$ 28 bilhões no intervalo de 2016 a 2025.

## 1.1 Motivações

Percebe-se, através dos números previamente citados, os desafios e oportunidades que a área de Manutenção 4.0 apresenta para os próximos anos. Por estar em transição entre o estado de domínio de aplicações restritas e atingir grande alcance e complexidade técnica, mais pesquisas e investimentos serão necessários. Tal evolução é baseada na ubiquidade e compartilhamento de dados em tempo real por toda a extensão da organização. Além disso, ao explorar bases da literatura científica, percebe-se que o interesse de pesquisa das relações entre *data streaming* (fluxo de dados gerados em tempo real) e Manutenção 4.0 apresenta tendência de crescimento nos últimos anos.

## 1.2 Objetivo

Este trabalho tem como objetivo apresentar, a partir de pesquisa na literatura, o uso de dados em tempo real, originados de IoT, como insumo para a Manutenção 4.0. A Manutenção 4.0 é conhecida também como manutenção inteligente ou digital. Desta forma, a pesquisa relaciona-se à investigação da função desempenhada pelo *data streaming* em sistemas de manutenção inteligente, a fim de detectar comportamentos anormais, viabilizar a atuação no tempo apropriado e minimizar a ocorrência de falhas ou interrupções do processo produtivo.

### 1.3 Justificativas

É possível, por meio da busca de referências nas bases de artigos acadêmicos *Scopus* e *Web of Science*, verificar o crescente interesse em pesquisas relacionadas a *data streaming*. Além disso percebe-se, nos últimos anos, uma tendência de aumento nos esforços de pesquisa relacionados ao uso de *data streaming* na manutenção industrial. Tais resultados são apresentados nas Tabelas 1 e 2. Conforme já descrito por Bruegge (2021), a Manutenção 4.0 receberá investimentos maiores nos próximos anos, o que a torna um objeto de pesquisa apropriado e atual.

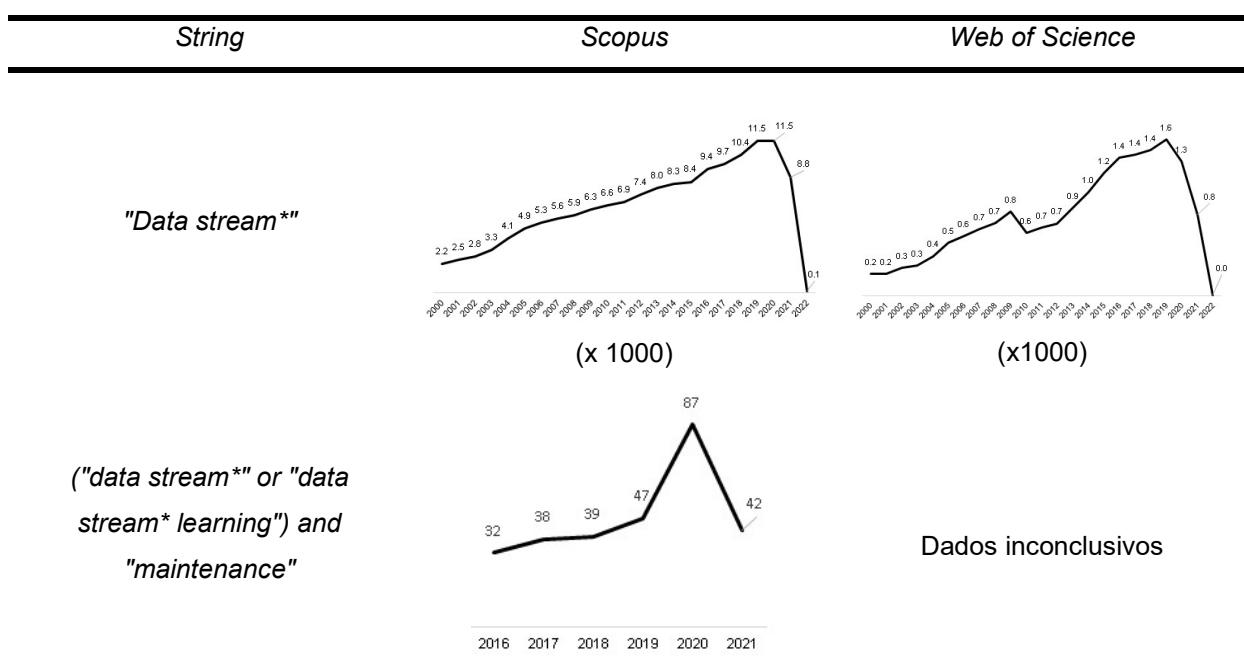
Tabela 1 – Resultados da pesquisa bibliográfica

<i>String</i>	Intervalo	<i>Scopus</i>		<i>Web of Science</i>	
		Resultados	Autores relevantes	Resultados	Autores relevantes
<i>"Data stream*" AND ("data stream*" AND "mainten*") AND ("Internet of things" OR "iot")</i>	2000 a 2022	149799	Gama, J.	17881	Bifet, A.
			Bifet, A.		Zhang, Y.
			Yu, P.S.		Gama, J.
			Soulsby, C.		Wang, Y
			Singh, V.P.		Wang, W
<i>Cap 4:</i> <i>("digital maintenance" OR "maintenance 4.0") AND ("stream*")</i>	2016 a 2021	2	Alexandru, A.M.	1	Alexandru, A. M.
			Apostolou, D.		Apostolou, D.
			Billaudelle, S.		Bousdekis, A.
			Bouguelia, M.R.		Fasanotti, L.
			Bousdekis, A.		Fiasche, M.
<i>Cap. 4:</i> <i>("data stream*" or "stream processing") and ("maintenance" or "abnormal") and ("4.0" or "intelligent" or "digital")</i>	2016 a 2021	103	Akkari, A.C.S.	35	Akkari, A.C.S.
			Metso, L.		Valamede, L.S.
			Thenent, N.E.		
			Valamede, L.S.		
			Borangiu, T.		
<i>Cap. 4:</i> <i>("data stream*" or "data stream* learning") and "maintenance"</i>	2016 a 2021	285	Anton, F.	147	Lewandowski, M.
			Anton, S.		Morariu, O.
			Morariu, O.		Raileanu, S.
			Răileanu, S.		Adhikari, B.
			Affenzeller, M.		Borangiu, T.
<i>Cap. 4:</i> <i>Zahmatkesh, S.</i>	2016 a 2021	285	Zahmatkesh, S.	147	Dey, M.
			Zenisek, J.		Dudley, S.
			Borangiu, T.		Kitagawa, H.
			Dey, M.		Liu, C.

Fonte: Própria autora.

Ao analisar os resultados das pesquisas em ambas as bases, verifica-se que a produção acadêmica acerca do tema “*data streaming*” mantém comportamento crescente desde 2000, limite inicial para a busca. Executou-se, para compreender o contexto de *data streaming* e *stream learning* no contexto de manutenção, a última consulta (com número mais significativo de resultados) descrita na Tabela 2. Conclui-se, ao verificar a quantidade identificada de referências nos intervalos de tempo, que pesquisas relacionadas a *data streaming* aplicado à Manutenção 4.0 vêm recebendo atenção da comunidade acadêmica nos últimos anos. Este comportamento é ilustrado na Tabela 2.

Tabela 2 – Comportamento do número de referências na linha do tempo



Fonte: Própria autora.

De acordo com Angelov e Kordon (2010), os processos produtivos enfrentam, atualmente, o desafio de desenvolver sensores inferenciais que sejam adaptativos e autocalibráveis a fim de reduzir os custos de manutenção e manter alta qualidade e precisão na produção. Zhou *et al.* (2021) apresentam um modelo inteligente de detecção de anomalias a partir de um conjunto de dados coletados pelo laboratório de uma empresa de segurança de redes na Austrália. Assim, propõe-se apresentar

variadas possibilidades de utilização de *data streaming* no campo da manutenção inteligente.

#### 1.4 Método de Pesquisa

Este trabalho foi construído através de consultas às bases de literatura científica *Scopus* e *Web of Science*. Aplicaram-se *strings* de busca (limitadas a *abstract*, título e palavras-chave no idioma inglês) a fim de refinar os resultados e obter dados a respeito do tema, tais como número de publicações por ano e autores mais relevantes. Os resultados das buscas estão representados na Tabela 1.

A partir das buscas realizadas, foram selecionados artigos que contribuem com a elaboração desta monografia. Além da bibliografia acadêmica, foram utilizadas outras referências na construção deste trabalho: livros que tratam sobre Indústria 4.0, manutenção, gestão de dados e *data streaming* aplicado à manutenção preditiva baseada em IoT; páginas de um fornecedor de tecnologias estruturantes (Amazon) e institutos de pesquisa de IoT no mundo dos negócios.

#### 1.5 Estrutura do Trabalho

O Capítulo 1 INTRODUÇÃO apresenta motivações, objetivo, justificativas, método de pesquisa e estrutura do trabalho.

O Capítulo 2 GESTÃO DE DADOS E INTERNET OF THINGS apresenta o ciclo de vida dos dados genérico, no contexto de IoT e voltado a *Big Data*, a serem aplicados na geração de insumos para a Manutenção 4.0.

O Capítulo 3 MANUTENÇÃO: DO TRATAMENTO DE FALHAS À PREDIÇÃO apresenta a evolução histórica das atividades de manutenção, desde a atuação para a correção de falhas até atuações preditivas que operam com maior independência.

O Capítulo 4 INTERAÇÕES ENTRE MANUTENÇÃO 4.0 E *DATA STREAMING* descreve o processo de aplicação de *data streaming* a ações inteligentes de manutenção (não apenas industrial).

O Capítulo 5 CONSIDERAÇÕES FINAIS apresenta a conclusão, contribuição e trabalhos futuros.

REFERÊNCIAS relaciona as fontes de pesquisa consultadas para a construção desta Monografia.

## 2 GESTÃO DE DADOS E *INTERNET OF THINGS*

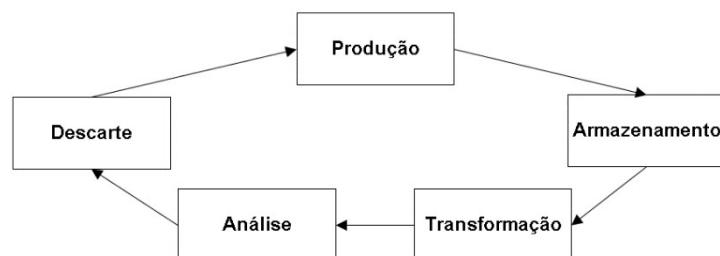
Este capítulo contextualiza a gestão de dados no domínio de IoT. O subcapítulo 2.1 apresenta o ciclo de vida dos dados de forma geral e no contexto de IoT; o subcapítulo 2.2 apresenta os modelos de processamento comumente utilizados neste contexto e o subcapítulo 2.3 discorre sobre a aplicação destes modelos em IoT.

De acordo com Helu *et al.* (2020), a fim de se extrair valor dos dados amplamente disponíveis, é necessária a construção de fluxos robustos, que suportem grande volume e variedade de dados. Dados brutos são coletados em diversas fontes e formatos e devem ser enriquecidos semanticamente a fim de fornecer contexto para análises sob diferentes pontos de vista.

### 2.1 Ciclo de vida dos dados genérico

Fluxos de dados são idealizados de forma a viabilizar todo o ciclo de vida dos dados, desde a coleta até o descarte. Amaral (2016) descreve o ciclo de vida de dados por meio de cinco etapas, que serão apresentadas nos próximos itens. São eles: 1. produção, 2. armazenamento, 3. transformação, 4. análise e 5. descarte. Os conceitos de cada etapa são apresentados e contextualizados de acordo com sua aplicação em IoT. A Figura 1 ilustra a estrutura do ciclo de vida dos dados.

Figura 1 – Ciclo de vida dos dados



Fonte: Amaral (2016).

### 2.1.1 Produção

De acordo com Amaral (2016), a produção de dados em processos tradicionais é viabilizada através dos seguintes recursos:

- Produção por processamento e análise: a saída de um processamento de dados atua como insumo para outros processos. A construção de modelos estatísticos e de Aprendizado de Máquina também podem ser enquadradas nesta classificação.
- Produção por transformação: alterar a estrutura dos dados de forma a adequá-los a um processo ou finalidade específicos, tais como processos de análise ou construção de *data warehouse* (do inglês, armazém de dados). Podem ocorrer transformações significativas nos dados.
- Sensores: telefones celulares e *smartphones* são equipados com sensores desde a sua concepção, que permitem a execução de diversas atividades. Alguns exemplos de sensores em *smartphones*: câmeras, telas sensíveis ao toque, acelerômetros, módulos de geolocalização, giroscópio e magnetômetro. Como passaram a ser mais eficientes e baratos, os sensores serão incorporados a mais dispositivos e sua quantidade aumentará. Desta forma, pode-se esperar que o volume de dados produzidos através de sensores aumente na mesma proporção nos próximos anos. Além de coletar dados, os sensores podem disparar atuadores e executar ações de acordo com o que se deseja obter.

Segundo Shimei *et al.* (2020), sob a ótica de IoT, dados podem originar-se de diversas fontes: manuais de operação, sistemas legados, registros de operação, sensores e Realidade Aumentada. Além dos recursos já conhecidos, a seguir são apresentadas particularidades específicas dos sensores inteligentes e Realidade Aumentada.

Sensores Inteligentes: no contexto de IoT, o sensoriamento envolve a coleta e devolução de dados por determinados objetos através da rede. Os dados coletados são analisados e devolvidos à rede a fim de disparar ações programadas. De acordo com Corral-Plaza *et al.* (2020), os dispositivos inteligentes podem servir como sensores ou atuadores e operam com dados de diferentes formatos, em grandes volumes.

Realidade Aumentada e Simulação: consiste em uma visão melhorada do mundo físico por meio da sobreposição de imagens, sons e interações por toques. É uma das tecnologias de ponta da Indústria 4.0 na geração de funcionalidades inteligentes. Previne erros que poderiam ser detectados em vários estágios da operação e colabora com melhorias de produtividade. Segundo Oztemel e Gursev (2020), a aplicação de Realidade Aumentada na manutenção e assistência remota reduz tempos de execução e erros humanos, além de disponibilizar análises de desempenho aos gerentes de manutenção.

### 2.1.2 Armazenamento

A segunda fase do ciclo de vida de dados apresentada por Amaral (2016) é o armazenamento. Após o processo de produção, os dados precisam ser persistidos para uso futuro. Esta operação depende de algumas premissas: segurança, integridade, minimização de redundância e concorrência e otimização do uso de espaço de armazenamento. Muitas vezes a utilização dos dados, em seu ciclo de vida, requer movimentação entre diferentes tipos de dispositivos, como é o caso das cópias de segurança. Neste caso, os dados são apenas replicados sem que ocorram quaisquer alterações. Já no caso dos armazéns de dados, há a alteração na estrutura dos dados a fim de proporcionar melhor qualidade e manter registros históricos. A depender da função a desempenhar, os dados podem ser armazenados em diferentes estruturas durante o ciclo de vida.

A Tabela 3 resume o contexto histórico dos modelos de armazenamento de dados e suas principais características.

Tabela 3 – Evolução dos modelos de armazenamento de dados

Modelo	Década	Característica principal
Pré-relacional	1960	Baseados em armazenamento e navegação entre vínculos
Relacional	1970	Normalização dos dados e recuperação de dados através de linguagem estruturada para consultas
Orientado a Objetos	1980	Suporte a dados orientados a objetos, provenientes das linguagens de programação com esta característica
NoSQL	2000	Suporte a dados não normalizados, poucas restrições de integridade e controle mínimo de transações

Fonte: Própria autora.

Amaral (2016) descreve cada um destes modelos de armazenamento de dados:

- Modelos pré-relacionais: desenvolvidos na década de 1960, os modelos pré-relacionais eram baseados em estruturas de navegação, em que os dados eram armazenados em registros vinculados. A fim de se recuperar um registro, era necessário navegar entre os vínculos. Modelos hierárquicos e em rede fazem parte desta classificação.
- Modelos relacionais: criados em 1970 com a finalidade de manter a integridade entre as transações e reduzir redundâncias de dados. Permitem inclusão, exclusão e alteração de registros sem renunciar à integridade. As operações em bancos de dados relacionais devem apresentar as características de atomicidade, consistência, isolamento e durabilidade. A consulta de dados é feita através de linguagem SQL (*structured query language*) (do inglês, linguagem estruturada de consulta).
- Bancos Orientados a Objetos: desenvolvidos a fim de fornecer suporte à Programação Orientada a Objetos. Porém, o uso dos bancos de dados relacionais já estava consolidado na época e as empresas não estavam dispostas à migração. Assim, foi desenvolvida a arquitetura conhecida como “Banco de Dados Objeto-Relacional”. Utiliza também a linguagem SQL.

- *NoSQL (Not Only SQL)* (do inglês, não somente SQL): tecnologia construída a fim de trabalhar com grandes volumes e diversos formatos de dados, o que requer uma estrutura mais flexível, escalável e de menor custo. Fatima e Wasnik (2016) descrevem o modelo *NoSQL* como não relacional, que conta com um esquema flexível, capaz de trabalhar com variedade de dados. Bancos de dados *NoSQL* são escaláveis e fornecem grande disponibilidade, o que os torna favoráveis à armazenagem de dados de *Big Data* e IoT. Apresenta o modelo *NewSQL*, que apresenta características dos modelos relacionais tradicionais e *NoSQL*. Os bancos de dados *NewSQL* são considerados promissores por serem mais adeptos ao processo de enriquecimento para IoT.

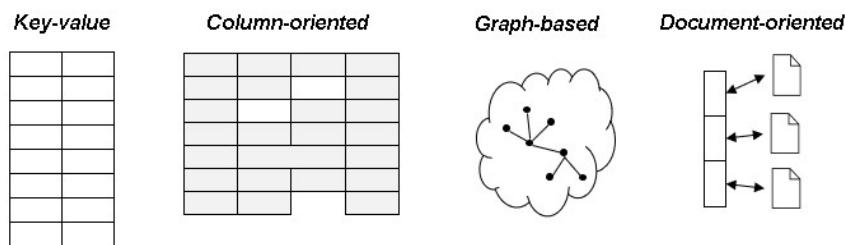
Os três grupos de bancos de dados *NoSQL* apresentados por Amaral (2016) são os *Key-Value* (Chave-Valor), *Column-Oriented* (Orientado a Colunas) e *Graph Based* (Baseados em Grafos).

- *Key-Value* (Chave-Valor): segundo a definição fornecida por Amazon (2021a), os bancos de dados chave-valor não possuem atributos fixos para descrever os dados, como ocorre nas estruturas relacionais tradicionais. Ao invés disso, armazena dados em um formato no qual os registros são construídos em uma estrutura simples, composta por uma chave e um valor. Esta chave é utilizada para identificar de forma única qualquer formato de dados, desde objetos simples aos mais complexos.
- *Column-Oriented*: estes bancos, ao invés da tradicional orientação por linhas, possuem dados orientados a colunas. A principal vantagem da adoção de bancos de dados orientados a colunas é a otimização de recursos para rápida recuperação de dados. As consultas analíticas apresentam melhor desempenho pois este modelo de banco de dados reduz expressivamente a quantidade necessária de entradas, saídas e recuperação de dados em disco, de acordo com definições fornecidas por Amazon (2021c).
- *Graph-Based*: de acordo com Amazon (2021b), bancos de dados baseados em grafos possuem como principal função o armazenamento e a navegação

entre relacionamentos, o principal item gerador de valor para este formato de banco de dados. Utilizam nós para armazenar entidades de dados e bordas para armazenar o relacionamento entre tais entidades. Assim, é possível existir uma infinidade de tipos de relacionamentos entre os nós, que são armazenados no próprio banco de dados e não se faz necessário calculá-los em todas as consultas. São utilizados em redes sociais, mecanismos de recomendação e detecção de fraudes.

Além dos três formatos citados, bancos de dados *NoSQL* podem ser *Document-Based* (Orientados a Documentos). Hashem e Ranc (2016) descrevem que tais documentos, geralmente associados a uma chave, são formados por um conjunto de campos encapsulados que podem ser individualmente indexados. São construídos em diversos formatos que utilizam estruturas de dados auto descritivos e em árvore hierárquica. Os documentos são agrupados em coleções; um conjunto de coleções compõe um banco de dados orientado a documentos. A Figura 2 ilustra os quatro formatos de bancos de dados *NoSQL* descritos.

Figura 2 – Modelos de bancos de dados *NoSQL*



Fonte: adaptado de Hashem e Ranc (2016).

Os principais formatos de dados utilizados pelos bancos descritos acima são os estruturados, não estruturados e semiestruturados. Segue breve descrição, segundo Eberendu (2016):

- Dados estruturados: possuem formato definido, o que torna simples sua análise e armazenamento. Os dados são armazenados em estruturas identificáveis e permitem consultas para utilização organizacional. São

comumente armazenados em bancos de dados relacionais e permitem diferentes tipos de dados, tais como número, data, caractere e outros. O foco da análise de dados tradicional são os dados estruturados.

- Dados semiestruturados: dados de formato irregular, que podem ser incompletos e sua estrutura, que não está vinculada a um modelo fixo, pode ser alterada rapidamente. Permitem a carga de dados de diferentes origens com propriedades relacionadas e que podem ser agrupadas em um grande conjunto. Alguns exemplos são *e-mails* e arquivos de texto.
- Dados não estruturados: além de texto, os dados não estruturados incluem formatos que não fazem parte de um modelo tradicional. Algumas classificações apresentadas pela autora incluem estáticos (documentos digitalizados), dinâmicos (documentos que podem ser criados, editados, revistos e aprovados, tais como procedimentos e políticas), mídias digitais (áudio, vídeo e imagens) e documentos de comunicação (*e-mail*, conteúdos de mídias sociais e registros de comunicação via mensagens instantâneas).

Os dados não estruturados compõem mais de 80% do total de dados disponíveis atualmente, segundo Rizkallah (2017). Assim, é necessário estabelecer ações a fim de explorar tais dados, pois são ricas fontes de conhecimento. Os bancos de dados *NoSQL* são parte importante deste processo.

### 2.1.3 Transformação

De acordo com Amaral (2016), os dados, após a coleta, precisam passar por um processo de transformação. Isto ocorre por conta da diversidade e heterogeneidade das fontes; assim, faz-se necessário preparar e adequar os dados ao consumo. Os processos de transformação podem efetuar diversas operações com os dados, tais como sumarização, agregação, cálculos, qualidade e limpeza de dados, mudança de codificação e alteração de formatos.

Nas transformações tradicionais de dados são executados os processos de extração, transformação e carga. O objetivo é, conforme mencionado acima, adequar os dados a fim de servirem como insumo para análises. Os processos de extração conectam-se às diversas fontes de dados, copiam-nos a uma área temporária e, então, os dados são transformados e armazenados na base definitiva.

De acordo com Mehdi Pour *et al.* (2016), no contexto de *Big Data* e IoT, a etapa de transformação de dados é desdobrada em etapas. De forma semelhante aos processos de extração, transformação e carga tradicionais, após a aquisição ou coleta os dados precisam ser integrados devido a suas diversas fontes e formatos. O próximo passo é a limpeza dos dados, responsável por reduzir significativamente a volumetria a fim de otimizar tempo e esforço nas futuras análises. Os dados brutos geralmente são não estruturados; assim, não estão organizados em um modelo predefinido. Podem ser, caso necessário, transformados em dados estruturados e semiestruturados na última fase do processo.

Khare e Totaro (2019) definem a etapa de transformação de dados para IoT como o processo de limpeza a fim de torná-los corretos e consistentes tecnicamente. Os passos envolvidos na melhoria e refinamento de dados são codificação, decodificação, conversão, rotulagem e associação a metadados. Após tal processamento ainda podem existir inconsistências; assim, é necessário domínio de conhecimento do dispositivo em questão a fim de identificar e corrigir tais erros.

#### 2.1.4 Análise

De acordo com Amaral (2016), a análise de dados é definida como a busca de conhecimento nos dados. Tal processo pode ser exploratório, implícito e explícito:

- Processo exploratório: de acordo com Amaral (2016), antes de iniciar uma análise, é necessário explorar os dados em termos de distribuição, médias, desvios padrões, relacionamentos e verificação de valores anormais. As análises exploratórias podem utilizar técnicas quantitativas e qualitativas. Técnicas quantitativas incluem medidas de dispersão e posição, tais como

média, mediana, amplitude e desvio padrão. Nestas análises é possível, além da análise numérica, utilizar os diagramas de dispersão, diagramas de caixa, histogramas e nuvens de palavras.

- Processo explícito: Amaral (2016) descreve que o processo explícito possui como objetivo destacar informações já existentes nos dados analisados. A principal diferença entre o processo explícito e o exploratório é que, enquanto o primeiro procura atingir um objetivo claro e específico, o segundo é utilizado para conhecer os dados. Alguns exemplos de processos explícitos são as junções, resumos, estratificação e identificação de distorções.
- Processo implícito: segundo Amaral (2016), processos implícitos são as análises executadas via Aprendizado de Máquina Computacional a fim de identificar padrões ocultos nos dados. Alguns exemplos são a classificação, regressão, agrupamento e regras de associação. Os processos implícitos de análise de dados são especialmente aplicáveis aos atuais cenários de *Big Data*, que exigem recursos mais robustos para extrair conhecimento de grandes volumes de dados.

#### 2.1.5 Descarte

A última fase do ciclo de vida dos dados é o descarte ou eliminação. No caso do Brasil, a LGPD (Lei Geral de Proteção de Dados) regulamenta a eliminação dos dados como fase essencial. De acordo com BRASIL (2020), a eliminação é a “exclusão de dado ou de conjunto de dados armazenados em banco de dados, independentemente do procedimento empregado”. Os dados devem ser descartados ao término do ciclo de tratamento, salvo condições previstas em lei. Os dados devem ser conservados, dentre outros motivos, para cumprimento de obrigação ou regulatória e estudo por órgão de pesquisa, desde que anonimizados.

Em suma, a Tabela 4 apresenta a relação, elaborada pela autora, entre as fases do ciclo de vida dos dados genérico e IoT.

Tabela 4 – Ciclo de vida de dados em IoT

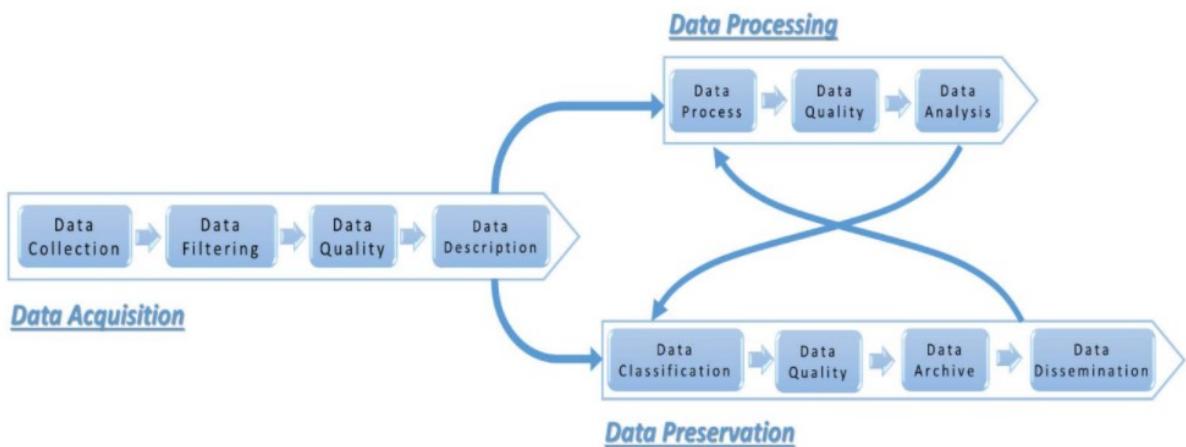
Fase	Aplicação
Produção	Coleta de dados de sensores, realidade virtual, manuais e registros de operação.
Armazenamento	Guarda dos dados em bancos de dados, geralmente <i>NoSQL</i> .
Transformação	Limpeza e correção de inconsistências; uniformização.
Análise	Geração de conhecimento através de diversas formas de estudo dos dados.
Descarte	Eliminação dos dados após o término do ciclo.

Fonte: Própria autora.

## 2.2 Ciclo de vida dos dados aplicado a *Big Data*

Além do modelo elaborado por Amaral (2016), para fins de aplicação em ambientes de *Big Data*, Sinaeepourfard *et al.* (2016) apresenta a proposição de um ciclo de dados específico. Aplicações de Manutenção 4.0 envolvem, muitas vezes, a utilização de *Big Data* devido ao grande volume, velocidade e variedade de dados gerados e processados. O modelo é composto por três grandes fases, que se dividem em elementos. São elas: aquisição de dados, composta por coleta, filtragem, qualidade e descrição dos dados; processamento de dados, composto por processo, qualidade e análise de dados e preservação de dados, composta por classificação, qualidade, armazenamento e disseminação de dados. A Figura 3 apresenta, de forma gráfica, os componentes do modelo proposto:

Figura 3 – Modelo proposto de ciclo de vida de dados para *Big Data*



Fonte: Sinaeepourfard *et al.* (2016)

## 2.3 Formas de processamento de dados

Podem-se destacar, dentre as formas comumente utilizadas para processamento de dados, as formas de processamento *batch* (em lote) e *streaming* (em fluxo). Nas próximas seções são abordados os conceitos e alguns exemplos de aplicação.

### 2.3.1 Processamento *batch* (em lote)

Os motores de processamento em lote são projetados para processar, de forma eficiente, grandes conjuntos de dados. Executam operações complexas, tais como a distribuição de recursos computacionais por diversos nós de processamento. Uma vez que o processamento é finalizado, os resultados são coletados e disponibilizados para consumo. Esta abordagem pode ser útil se um conjunto de dados precisa ser processado em conjunto e é requerido que todos os dados estejam disponíveis antes de o processamento iniciar, de acordo com Pfandzelter e Bermbach (2019). Um exemplo de aplicação é a geração de relatório de gastos, geralmente disponibilizados ao cliente mensalmente para pagamento.

### 2.3.2 *Streaming*

O processamento em *streaming*, de acordo com AWS (2021d), possui como principal característica serem gerados por diversas fontes de dados que enviam registros de tamanho pequeno em fluxo contínuo. Estes dados podem pertencer a diversas variedades, tais como registros de atividades de clientes em aplicativos, compras em *e-commerce*, redes sociais, serviços financeiros e geoespaciais e telemetria de dispositivos conectados. Devem ser processados por registro ou durante períodos móveis a fim de alimentar análises diversas, que permitem ao negócio acompanhar em tempo real suas operações e agir em tempo hábil quando um comportamento anormal é detectado.

A Tabela 5 apresenta uma comparação entre características dos processamentos em lote e *streaming*, de acordo com AWS (2021d):

Tabela 5 – Comparação entre processamentos em lote e *streaming*

	Processamento em lotes	Processamento de <i>streams</i>
<b>Escopo de dados</b>	Consultas ou processamento de todos ou da maioria dos dados no conjunto de dados.	Consultas ou processamento de dados dentro de um período sem interrupções, ou apenas no registro de dados mais recente.
<b>Tamanho dos dados</b>	Grandes lotes de dados.	Registros individuais ou micro lotes compostos de alguns registros.
<b>Desempenho</b>	Latências em minutos a horas.	Exige latência na ordem dos segundos ou milissegundos.
<b>Análise</b>	Análise complexa.	Métricas simples de funções, agregação e rotação de respostas.

Fonte: AWS (2021d).

De acordo com Bahri *et al.* (2021), “*data stream*” pode ser definido como sequências esporádicas e transitórias de dados obtidos através do tempo. Uma infinidade de instâncias de dados pode ser gerada através deste processo; assim, valiosos *insights* de dados podem ser obtidos através de processos de análise em tempo real. Para tal, é necessário empregar recursos computacionais eficientes e paralelização de processamento.

## 2.4 Considerações sobre o capítulo

O Capítulo 2 apresenta o ciclo de vida de dados genérico e aplicado ao contexto de IoT. Este é composto por cinco fases: 1. Produção (coleta de dados originados de sensores inteligentes, Realidade Aumentada, manuais de operação e registros de operação), 2. Armazenamento (bancos de dados relacionais, não relacionais e, no contexto de IoT, *NoSQL* e *NewSQL*), 3. Transformação (processo de uniformização e tratamento dos dados de forma a prepará-los para a fase de análise), 4. Análise (aplicação de técnicas para extrair conhecimento dos dados) e 5. Descarte (remoção dos dados quando tiverem cumprido sua função ou segundo leis e normas). Além disso, é apresentado um modelo de ciclo de vida dos dados adaptado às características de *Big Data*, composto pelas fases Aquisição, Processamento e Preservação de dados.

As particularidades do ciclo de vida dos dados em IoT estão concentradas nas etapas Produção (dados gerados por dispositivos inteligentes em tempo real) e Análise (produção de conhecimento baseado em dados de tempo real). Os processos de manutenção podem ser beneficiados por este recurso através da medição e coleta de dados relacionados a variáveis críticas. Estes dados podem gerar alertas preditivos ou disparar ações sem a necessidade de intervenção humana. Assim, são evitadas paradas de equipamentos e prejuízos à produção devido a falhas.

### 3 MANUTENÇÃO: DO TRATAMENTO DE FALHAS À PREDIÇÃO

Qualquer instalação ou sistema, que possui como objetivo a produção de um bem, utiliza dispositivos simples ou complexos. Independente de qual seja o dispositivo, podem ocorrer diversas formas de incidentes durante o processo de produção. Assim, toda atividade produtiva requer manutenção, sob o risco de a produção entrar em colapso sob a ocorrência de falhas.

Segundo apresentado por Nepomuceno (2014), em instalações industriais clássicas, as atividades de manutenção recebem pouca importância e são consideradas apenas quando ocorrem falhas. Porém, quando há um planejamento adequado, os custos de manutenção tornam-se baixos em comparação aos retornos de longo prazo. Boa conservação dos equipamentos aumenta sua vida útil e proporciona retorno satisfatório, muitas vezes superior ao investido na manutenção praticada. À medida que a complexidade dos equipamentos aumenta, os problemas relacionados à manutenção também aumentam. Isto torna necessário o investimento em mudanças organizacionais de forma a dar à manutenção um papel de maior destaque.

Este capítulo mostra um breve histórico das práticas de manutenção e sua evolução, desde a simples correção de falhas até a manutenção digital e baseada em dados. Identifica diferenças entre a manutenção tradicional e a Manutenção 4.0. Apresenta recursos inerentes à Manutenção 4.0, tais como o conceito de Internet Industrial das Coisas e sensores inteligentes.

#### 3.1 Manutenção tradicional: evolução histórica

De acordo com Poor *et al.* (2019) são encontradas, na literatura, diversas definições para “manutenção”. Em suma, referem-se a ações aplicadas com o objetivo de manter a qualidade esperada de um equipamento, reduzir os efeitos de falha e melhorar o desempenho.

Conforme descrevem Slack *et al.* (2018) e Jacobs e Chase (2012), apesar dos processos de produção serem planejados de forma a evitar a ocorrência de falhas, a manutenção continua a ser uma necessidade. Isto ocorre através de procedimentos que indicam os cuidados requeridos pelas instalações e equipamentos de uma organização. Alguns dos benefícios derivados de um processo eficaz de manutenção são maior segurança, maior confiabilidade, melhor qualidade, custos operacionais menores, maior vida útil da tecnologia de processo e maior valor final.

Stevan Jr *et al.* (2018), Nepomuceno (2014) e Slack *et al.* (2018) definem a execução de processos de manutenção sob três formas: corretiva, preventiva e preditiva. Segundo estas referências, as definições de cada um dos termos são:

- Corretiva: a manutenção corretiva é executada em caso de falha ou baixo desempenho do equipamento. Apresenta alta probabilidade de ocorrência de erros humanos durante a execução e há custos referentes ao estoque de componentes que suprem o processo. A parada inesperada da operação em caso de falhas traz prejuízos ao processo produtivo e, consequentemente, à organização.
- Preventiva: parada programada do equipamento para substituição de componentes e execução de procedimentos planejados de manutenção. O equipamento deve ser parado para que a manutenção seja efetuada, o que pode trazer prejuízos econômicos devido à interrupção ou estoque e substituição desnecessária de componentes. Por conta dos custos envolvidos nas manutenções corretivas e preventivas, faz-se necessário um levantamento comparativo entre custo de manutenção e custo de falha.
- Preditiva: dados sobre variáveis críticas são coletados pela própria máquina em tempo real, a fim de evitar a parada do equipamento devido a falhas. Através desta é possível programar, com a devida antecedência, operações de manutenção de acordo com o comportamento das variáveis mensuradas. Viabiliza a alteração dos parâmetros de condição ou desempenho através de

acompanhamento sistemático. Assim, o estudo contínuo dos dados obtidos pode determinar com antecedência a necessidade de manutenção.

O conceito de manutenção preditiva não é recente; porém, uma limitação apresentada por Nepomuceno (2014) ao final da década de 80 é a não existência de métodos de manutenção preditiva aptos a detectar defeitos em tempo hábil. A solução vigente na época era a instalação de sensores de medição nos equipamentos; porém, tais sensores funcionavam com a energização do equipamento e, em caso de interrupção do fornecimento de energia, a medição era interrompida. Conforme será tratado adiante, a evolução da manutenção nas próximas décadas supre esta limitação.

Os conceitos referentes à gestão de manutenção começaram a receber maior foco no século XX, com o advento do Sistema Toyota de Produção e popularização da manufatura enxuta. A Figura 4, elaborada por Stevan Jr *et al.* (2018), apresenta a evolução histórica dos conceitos de manutenção.

Figura 4 – Evolução da manutenção

<b>Primeira geração: corretiva</b> Conserto após falha; Gestão reativa.	<b>Segunda geração: preventiva</b> Revisões programadas; Sistemas de PPCM; Técnicas preditivas; Gestão proativa.	<b>Terceira geração: produtiva</b> Produção monitorada; Projeto, Confiabilidade e Manutenibilidade. Estudos de <i>retrofit</i> ; Microcomputadores; Sistemas especialistas; Versatilidade; Trabalho em equipe; FMEA manutenção; Gestão TPM Lean.	<b>Quarta geração: digital</b> Monitoramento total; Projeto, confiabilidade e manutenibilidade; Conservação de energia; Segurança total; Fábrica digital e PLM; Sistemas MES e ERP; Simulação e modelagem; Estudos de riscos; <i>Netbooks</i> e <i>tablets</i> ; Sistemas especialistas; Funções múltiplas; FMEA ciclo produtivo; Gestão digital integrada.
1940	1950	1960	1970
1980	1990	2000	2010

Fonte: Stevan Jr *et al.* (2018).

Conforme apresentado na Figura 4, a primeira geração de manutenção, entre as décadas de 1940 e 1950, costumava ser puramente reativa. A segunda geração, entre as décadas de 1960 e 1970, possuía enfoque mais proativo. Foram incluídas revisões programadas às rotinas de manutenção dos equipamentos e deu-se início às primeiras formas de predição. A terceira geração (décadas de 1980 e 1990) passa a adotar atividades de monitoramento do processo produtivo apoiadas por recursos computacionais. Finalmente, a quarta geração (2000-presente) incorpora recursos tecnológicos mais avançados, tais como sistemas especialistas, dispositivos para acompanhamento visual e integração da gestão digital.

### 3.2 Manutenção 4.0

Os processos de manutenção descritos anteriormente evoluíram e deram origem à Manutenção 4.0 (também conhecida por “manutenção digital”). A Manutenção 4.0, de acordo com Slack *et al.* (2018), permite monitorar totalmente o sistema através de mecanismos de confiabilidade, otimização de recursos e segurança, integração dos sistemas *Manufacturing Execution System* (MES) (do inglês, sistema de execução de manufatura) e *Enterprise Resource Planning* (ERP) (do inglês, planejamento de recursos de empreendimento) com mecanismos de modelagem e simulação, além de um completo sistema de gestão integrada. No ambiente da Indústria 4.0, os sistemas computacionais estão inseridos e integrados entre si. Com o advento da inteligência artificial, começou-se a desenvolver o conceito de manutenção inteligente: as máquinas e equipamentos identificam as condições anormais de operação e emitem notificações para que ações sejam tomadas. Além disso, o sistema permite que o manutentor não se preocupe com o fluxo de manutenção, pois o sistema inteligente realiza todos os procedimentos para que as máquinas possam ser corrigidas, incluindo previsão de recursos (equipe, peças, tempo).

Assim, processos de manutenção corretiva são minimizados no ambiente da Indústria 4.0 por meio da contínua coleta e análise de dados, o que fornece embasamento para a atualização de informações. Em sistemas de Manutenção 4.0 existe um fluxo contínuo de dados trafegando pela arquitetura. Assim, toda a informação está disponível para qualquer parte da estrutura industrial.

Um recurso utilizado neste processo é a virtualização de um elemento industrial e acesso a todos os seus dados via equipamentos móveis. Isto inclui manuais, modelo, parâmetros, histórico de manutenção e outras informações relevantes. Assim, o processo de acompanhamento de um equipamento é composto por sua virtualização e criação de perfis de manutenção preventiva e programada, o que inclui dados históricos. Estes dados ficam armazenados em nuvem, o que permite acompanhar individualmente os elementos a fim de apresentar planos de manutenção para otimizar custos e reduzir o número de paradas. Além disso, os dados são facilmente acessíveis a diferentes níveis da organização e são mantidos durante todo o tempo de vida útil do equipamento, o que permite melhor gestão de suprimentos, interações humanas e redistribuição de ativos.

### 3.2.1 A função de *Internet of Things* na Manutenção 4.0

A fim de viabilizar a Manutenção 4.0, descrita o subcapítulo anterior, algumas tecnologias habilitadoras são necessárias. A *Industrial Internet of Things* (IIoT) (do inglês, Internet Industrial das Coisas) é uma delas.

De acordo com Sisinni *et al.* (2018), IoT é, em suma, um conceito que descreve a conexão ubíqua à Internet e a conversão de objetos comuns em dispositivos conectados. A IIoT é um subconjunto da IoT que associa tecnologias de comunicação a aplicações de automação industrial. Viabiliza a melhor compreensão do processo produtivo a fim de melhorar sua eficiência e sustentabilidade, além de enfatizar a integração entre diferentes plantas produtivas e locais de trabalho. Conecta todos os ativos da indústria a sistemas de informação e aplicações de negócios. O grande volume de dados produzidos e trocados pela IIoT possui como objetivo final análises preditivas, tais como manutenção e logística aprimorada. Assim, a IIoT é um subconjunto da IoT, específico para aplicações industriais. A intersecção entre a Indústria 4.0 e a IoT ocorre na IIoT.

Segundo apresentado por O'Donovan *et al.* (2015, a. 25), a fim de tratar de alguns problemas relacionados ao ambiente industrial tradicional, emergiu a manufatura

inteligente, um campo interdisciplinar. A manufatura inteligente pode ser descrita como orientada a dados, onde dados em tempo real coletados via sensores diretamente na fábrica podem ser analisados em processos de decisão. É uma especialização de *Big Data* que atende especificamente as necessidades do processo de produção e utiliza como tecnologias de apoio o Aprendizado de Máquina e a IIoT. É importante o foco em manutenção e diagnóstico por conta do impacto nos custos operacionais totais, que podem exceder 30%. O tempo de vida dos equipamentos pode ser melhorado através da redução das retiradas de operação; técnicas preventivas e preditivas podem ser empregadas na programação das paradas para manutenção. Porém, a manufatura inteligente possui extremo foco em tratamento de dados e compartilhamento de informação em tempo real entre processos físicos e computacionais a fim de alimentar aplicações analíticas.

De acordo com Gilchrist (2016), em conjunto com *Big Data*, computação em nuvem, sensores, tecnologias *machine-to-machine* (M2M) (do inglês, máquina a máquina) e outras habilitadoras, a IIoT provê recursos para se obter melhor visibilidade dos processos e recursos de uma indústria. Dados são coletados através deste sistema e transformados em conhecimentos úteis para o negócio através de *analytics* avançado. Benefícios para o negócio são obtidos pois tais tecnologias fornecem recursos para a elaboração de análises que trazem ganho operacional, aceleração da produtividade e redução do tempo em que a operação precisa ficar parada.

Em suma, pode-se afirmar que a IIoT é requisito primordial para a implantação da manufatura inteligente, da qual a manutenção preditiva faz parte. Foco é dado à coleta e análise de dados em tempo real a fim de permitir que decisões sejam tomadas o quanto antes e, quando possível, sem necessidade de intervenção humana.

### 3.3 Considerações sobre o capítulo

Este capítulo aborda o histórico de evolução da gestão de manutenção e a crescente atenção que tal conceito recebe com o passar do tempo. Os processos de manutenção tiveram seu princípio em atividades corretivas; isto é, ações eram

tomadas apenas na presença de falhas e não se tomavam ações para evitar novas ocorrências. Posteriormente foram desenvolvidos procedimentos preventivos, nos quais rotinas de manutenção programada retiravam os equipamentos de operação. Finalmente, adotou-se a manutenção preditiva, na qual variáveis críticas do processo são monitoradas a fim de prever a necessidade de manutenção antes que a falha ocorra.

A manufatura inteligente, da qual a manutenção 4.0 é integrante, caracteriza-se pela comunicação ubíqua entre dispositivos. Para tal, é essencial a adoção da IIoT a fim de garantir o tráfego de dados em tempo real entre os dispositivos e entre as instalações físicas, que podem estar distantes entre si. O capítulo apresentou algumas diferenças entre IoT e IIoT e o papel da IIoT como geradora de dados em tempo real para a Manutenção 4.0.

## 4 DATA STREAMING E A MANUTENÇÃO 4.0

Neste capítulo é apresentada a utilização de *data streaming* (fluxo de dados gerados em tempo real) como requisito de evolução da manutenção tradicional. A fim de se obter o melhor aproveitamento de tais dados, são necessários recursos para análise. Alguns conceitos, algoritmos e recursos que apoiam estas análises, apresentados nas referências citadas neste trabalho, são descritos. A Manutenção 4.0 é a evolução da manutenção segundo a Quarta Revolução Industrial; assim, é parte integrante da Indústria 4.0. Sob esta ótica, são investigados alguns fatores viabilizadores que devem ser considerados para a implantação bem-sucedida de processos de manutenção inteligente. Por fim são apresentados alguns cases de aplicação em áreas diferentes, extraídos da literatura citada.

### 4.1 A importância do *data streaming* na Manutenção 4.0

A respeito das possíveis fontes de dados utilizadas na Manutenção 4.0, Sahal *et al.* (2020) afirmam que aplicações de manutenção preditiva devem ser capazes de identificar falhas com antecedência, de modo que os responsáveis tomem ações a tempo; sejam elas reposição de componentes ou parada de máquina planejada. Proporcionam economia na manutenção de equipamentos e melhoram a produtividade através do aumento de tempo em que permanecem em operação. Muitas vezes a ocorrência de falha em um equipamento causa impactos negativos em todo o processo produtivo. Porém, para que as soluções de manutenção preditiva sejam exatas, é essencial coletar e analisar grandes volumes de dados relevantes em espaços de tempo razoáveis. Desta forma, recursos de análise avançada e processamento em tempo real são requisitos-chave para a manutenção preditiva.

A visão apresentada por Munirathinam (2020) declara que a IoT deu origem a um subsegmento voltado à indústria: a Internet Industrial das Coisas. A Internet Industrial das Coisas baseia-se fortemente em interconectividade, automação, autonomia, aprendizado de máquina e dados em tempo real. A captura e

transmissão de dados entre equipamentos proporciona oportunidades de crescimento às empresas. A IoT e seus derivados baseiam-se em conectividade, o que significa maior quantidade de dados obtidos em diversos lugares e em tempo real. Seus objetivos são o aumento de produtividade, lucro e eficiência, além da viabilização de decisões tomadas em tempo real.

Com base nestas referências, pode-se concluir que a coleta e a análise de *data streams* são essenciais para o sucesso de um processo de manutenção preditiva nos moldes atuais. Dados em tempo real possibilitam identificar possíveis futuras falhas e permitem a adoção de planos de ação em tempo hábil, de forma a evitar e tratar tais falhas com antecedência.

Conforme destacado no Capítulo 3, a engenharia de manutenção pode ser beneficiada com a adoção de tecnologias inovadoras relacionadas à Indústria 4.0, a fim de prover mitigação da ocorrência de falhas, aumento da eficiência na utilização de recursos e maior precisão no fornecimento de valor ao cliente. A utilização de IIoT neste processo é a maior viabilizadora do sucesso de um projeto de Manutenção 4.0 devido à possibilidade da coleta e intercâmbio de dados em tempo real entre dispositivos e instalações físicas.

De acordo com Cohen *et al.* (2017), a Indústria 4.0 é guiada por quatro princípios: conectividade, informação, conhecimento e inteligência. Consequentemente, a Manutenção 4.0 (também denominada “manutenção inteligente”) deve ser caracterizada pelos mesmos fatores. Bokrantz *et al.* (2020) definem a manutenção inteligente, em linhas simples, como um projeto organizacional de gerenciamento de manutenção em ambientes com tecnologias digitais pervasivas. Trata-se de um conceito multidimensional baseado em decisões baseadas em dados, recursos de capital humano, integração interna e integração externa.

Desta forma, a Manutenção 4.0 é caracterizada pelo constante fluxo de dados coletados nos dispositivos, compartilhamento da informação por toda a organização e utilização destes dados em análises de manutenção preditiva e processos de tomada de decisão. Além disso, os dados devem estar facilmente acessíveis via Internet, de qualquer lugar do mundo. Desta forma é possível atuar em possíveis

incidentes dentro do tempo apropriado, evitando-se prejuízos como paradas na operação e perda de componentes.

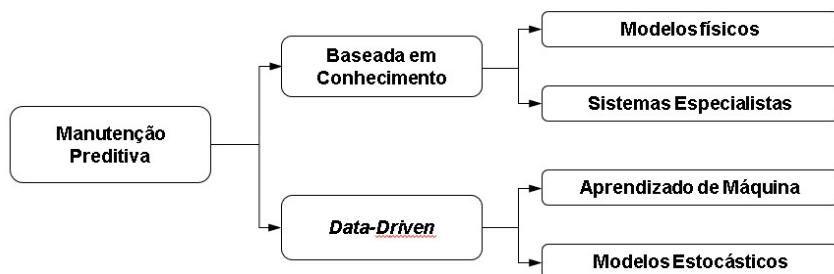
A fim de viabilizar a implantação e uso eficiente da Manutenção 4.0 é necessário identificar alguns fatores essenciais. Navas *et al.* (2020) apresentam alguns dos requisitos técnicos básicos para a construção de um ambiente de Manutenção 4.0 de qualidade:

- *Design* ergonômico de sistemas para reduzir o tempo de parada para manutenção, montagem e desmontagem de componentes e ajuste de elementos. Os sistemas ergonômicos são, de acordo com Ciccone *et al.* (2021), desenvolvidos para melhorar a acessibilidade de uma maior variedade de perfis de usuários;
- Sensores que viabilizam o monitoramento de variáveis físicas e operacionais para a implantação de técnicas de manutenção preditiva em tempo real. O uso de sensores na manutenção industrial não é recente; porém, com o advento da Indústria 4.0, foram adotados os sensores inteligentes. Os sensores inteligentes possuem pequena memória e trocam dados com a rede através de comunicação padronizada, segundo Aksa *et al.* (2021). Estes são elementos básicos da Manutenção 4.0 e podem-se adotar diversas variedades de sensores em conjunto;
- Portas exclusivas de comunicação para a gestão do processo de manutenção, estados, alarmes e variáveis preditivas. A integração do equipamento ao centro de manutenção deve ser feita por meio da implantação de protocolos padronizados de comunicação a uma porta dedicada, a fim de simplificar e reduzir os custos de integração;
- Autoteste e testes funcionais de sistemas projetados para tarefas de manutenção, a fim de garantir que cumpra os requisitos particulares do projeto.

De acordo com Navas *et al.* (2017), a manutenção tradicional opera de forma pouco eficiente ao utilizar uma grande quantidade de recursos humanos e técnicos. Complexos modelos foram desenvolvidos nas últimas décadas para esta função; porém, nem sempre há comprovação real de benefícios. A fim de viabilizar ações de manutenção precisas e no tempo exato, dados devem ser coletados em tempo real. Atuar em incidentes em um intervalo de tempo inapropriado pode causar prejuízos à operação e a outros departamentos da empresa, além dos prejuízos financeiros.

Le Nguyen *et al.* (2020) apresentam duas abordagens de dados que podem ser utilizadas na Manutenção 4.0: *knowledge-based* (baseada em conhecimento) e *data-driven* (orientada a dados). A primeira pode ser desdobrada em modelos físicos e sistemas especialistas, enquanto a segunda é composta por Aprendizado de Máquina e modelos estocásticos. Tal desdobramento é apresentado na Figura 5:

Figura 5 – Taxonomia das abordagens de manutenção preditiva



Fonte: Le Nguyen *et al.* (2020).

A abordagem *knowledge-based* conta com especialistas de domínio para a construção de modelos de manutenção preditiva. Pode ser dividida em duas classes: modelos físicos e sistemas especialistas.

Os modelos físicos são o conjunto de equações matemáticas que descrevem o mecanismo de deterioração de um equipamento. Estes são formados por vasto conhecimento de mecânica e especialização de domínio. Desta forma, os modelos adequados podem refletir com acurácia o comportamento físico de deterioração e fornecer *insights* confiáveis para a manutenção da saúde do equipamento. Porém,

este modelo pode não ser prático pois a complexidade dos sistemas do mundo real exige uma modelagem precisa e não é possível adaptar o mesmo modelo para sistemas diferentes. A fim de validar os parâmetros da equação, resultados de testes físicos são utilizados ao invés de dados. A operação do equipamento precisa ser interrompida para que tais testes sejam feitos.

Os sistemas especialistas são compostos por uma base sólida de conhecimento gerada por pessoas que detêm extenso conhecimento sobre o domínio, composta por fatos formalizados, regras e resultados de consultas. Esta solução é útil para diagnóstico de falhas através de alertas e agendamento de manutenções quando os indicadores excedem determinado limite ou quando um componente está próximo de sofrer uma falha. Contam com equipamento computacional poderoso e algoritmos que permitem a geração de soluções mais rapidamente que especialistas humanos. Porém, a modelagem necessária é complexa e não consegue lidar com situações não cobertas pelas regras. Desta forma, ainda de acordo com Le Nguyen *et al.* (2020), a abordagem *data-driven* é a mais apropriada para a Manutenção 4.0.

A abordagem *data-driven*, diferente da abordagem *knowledge-based*, estrutura-se em dados coletados por dispositivos conectados, tais como *logs* de sistemas, históricos de manutenção e leituras de sensores ou fluxos de dados de qualquer tipo ou fonte. Com base nestes dados é possível reconhecer padrões e predizer futuras falhas.

De acordo com Bahri *et al.* (2021), *data streams* podem ser definidos como sequências esporádicas e transitórias de dados obtidos através do tempo. A fim de extrair o maior valor destes dados, é necessário explorar formas de análise em tempo real. Um *streamer* gera altos volumes de dados em tempo real e *insights* ricos podem ser obtidos através de sua análise. Para tal, é necessário empregar recursos computacionais eficientes e paralelização de processamento. Porém, é necessário levar em consideração alguns desafios para a adoção de sistemas baseados em *data streaming*. São eles: flexibilidade de evolução, curto tempo de processamento, uso de memória, alta dimensionalidade, *concept drifts*, rotulagem atrasada e rotulagem desbalanceada.

O *data streaming* viabiliza a geração de alto volume de dados em tempo real. A fim de utilizar ao máximo a resultante capacidade de obtenção de conhecimento, faz-se necessário o uso de recursos de análises de dados, tais como o *stream learning* (do inglês, aprendizado de fluxo), em aplicações de Manutenção 4.0.

#### 4.2 Técnicas e ferramentas para extração de conhecimento do *data streaming*

Algumas das técnicas empregadas na análise de *data streams* para a Manutenção 4.0 são a aplicação de algoritmos de aprendizado supervisionado e não supervisionado, *concept drifts* (desvios de conceito) e *learning and forgetting* (aprendizado e esquecimento). Le Nguyen *et al.* (2020) apresentam os principais algoritmos utilizados na Manutenção 4.0 e as definições dos algoritmos são apresentados de acordo com as referências citadas nas Tabelas 6 e 7.

Segundo definição apresentada por Le Nguyen *et al.* (2020), o aprendizado supervisionado consiste em regressão ou classificação de *streams*. Dois exemplos de algoritmos de classificação utilizados são o *Hoeffding Tree* (HT) e o *Naïve Bayes*. De acordo com Antony *et al.* (2021), HT é um modelo baseado em árvore, apropriado para lidar com grandes volumes de dados com aprendizado contínuo por instância. Caso sejam fornecidos suficientes exemplos para treinamento, o resultado será bastante semelhante ao que outros algoritmos baseados em árvore podem fornecer. Por sua vez, o clássico *Naïve Bayes*, segundo Youssef, A. e Pourghasemi, H. (2021), é facilmente adaptável ao modelo de dados em *streaming* ao atualizar as ocorrências dos valores de atributos de forma incremental.

As regressões de *streams* podem ser baseadas em árvore ou regras, apresentadas a seguir.

- Regressão baseada em árvore: *Fast Incremental Model Trees with Drift Detection*; de acordo com Wibisono *et al.* (2016), funciona através da segmentação por seleção de atributos. Cada folha possui um modelo linear atualizado a cada vez que recebe uma nova instância. O modelo então executa regressões para instâncias não rotuladas na folha.

- Regressão baseada em regras: *Adaptive Model Rules from High Speed Data Streams*. Segundo a definição apresentada por Duarte *et al.* (2016), inicia-se com um conjunto vazio de regras, que são incrementalmente redimensionadas conforme novos dados chegam. Cada regra contém um modelo linear que é treinado incrementalmente nos dados cobertos por tal regra. O valor predito para uma instância não vista é a média das regressões individuais dadas pelas regras que cobrem tal instância.

A Tabela 6 apresenta um resumo dos algoritmos de aprendizado supervisionado utilizado na análise de *data streams* em Manutenção 4.0.

De acordo com Le Nguyen *et al.* (2020), por meio de clusterização ou detecção de anomalias (aprendizado não supervisionado) são feitas análises para descobrir padrões e rotular os dados originários de *data stream*. A seguir são apresentados alguns algoritmos de aprendizado não supervisionado utilizados em *data streaming*.

- Clusterização em duas fases. *CluStream*, segundo Hu *et al.* (2021), é um exemplo representativo de algoritmo de clusterização em duas fases. A primeira fase (*online*) coleta e atualiza de forma eficiente a síntese estatística de dados no formato de *micro-clusters*. A fase *offline* trata estes *micro-clusters* como pontos e executa o algoritmo *K-Means* para apresentar o resultado da clusterização em determinado horizonte de tempo. Outros exemplos de algoritmos que utilizam a abordagem em duas fases são *DenStream* e *D-Stream*, segundo Putina *et al.* (2018).
- Algoritmos de detecção de anomalias. A detecção de anomalias procura pontos anormais comparados à distribuição dos dados dominantes nos *data streams*. Exemplos: algoritmos *Hierarchical Temporal Memory* que, de acordo com Ahmad *et al.* (2017), detecta *outliers* em *streams* que apresentam ruídos e alterações dinâmicas. O *xStream*, de acordo com Akyildiz *et al.* (2018) executa detecção de anomalias baseada em densidade, na qual uma instância não chega com todos os atributos esperados, mas estes chegam parte a parte através do tempo.

Tabela 6 – Algoritmos de aprendizado supervisionado

Algoritmo	Classificação ou regressão	Definição	Fonte
<i>Hoeffding Tree</i>	Classificação	Estrutura de árvore de decisão incrementalmente induzida. Os nós internos contêm testes baseados em um atributo e direcionam a instância de dados que chega a uma folha, baseado em seus atributos. As folhas contêm preditores de classificação e classificam cada instância. A diferença entre a indução da <i>Hoeffding Tree</i> e as árvores de decisão em lote é processar a instância apenas uma vez, em vez de iterar por todos os dados.	Antony et al. (2021)
<i>Naïve Bayes</i>	Classificação	Método de classificação baseado no teorema de Bayes. Assume que todos os fatores são independentes, dada a classe de saída. Vantagens: robusto contra ruídos e variáveis irrelevantes, facilidade na aplicação e esquemas iterativos complexos são desnecessários.	Youssef, A.; Pourghasemi, H (2021)
<i>Fast Incremental Model Trees with Drift Detection</i>	Regressão	Algoritmo de aprendizado incremental de regressão e árvores de modelo a partir de <i>streams</i> que mudam através do tempo. Ao invés de processar todos os dados em memória, o algoritmo permite que, a cada chegada de um <i>stream</i> , dados sejam processados e aprendidos.	Wibisono et al. (2016)
<i>Adaptive Model Rules from High Speed Data Streams</i>	Regressão	Primeiro algoritmo de aprendizado baseado em regras para problemas de regressão. O antecedente de uma regra é um conjunto de condições nos valores dos atributos e a saída é uma combinação linear dos atributos.	Duarte et al. (2016)

Fonte: Própria autora.

A Tabela 7 resume os algoritmos de aprendizado não supervisionado apresentados neste trabalho.

Tabela 7 – Algoritmos de aprendizado não supervisionado

Algoritmo	Definição	Fonte
<i>CluStream</i>	Possui como objetivo executar clusterizações de forma mais rápida enquanto garante sua acurácia e é utilizado na correlação de alarmes. O algoritmo é formado por clusterização <i>online</i> e <i>offline</i> . Na fase online, a meta é analisar rapidamente os <i>data streams</i> e gerar um micro registro que se refere ao conjunto de condições anormais. Na fase offline, a ideia principal é determinar se o alarme está relacionado às características das condições anormais.	Hu et al. (2021)
<i>DenStream</i>	Permite descobrir clusters de formato arbitrário ao manter um conjunto de micro clusters normais e outliers. O <i>DenStream</i> pode unir novos pontos de dados a um micro cluster já existente (normal/outlier) ou gerar um novo micro cluster de outliers, o que dispara um alarme.	Putina et al. (2018)
<i>Hierarchical Temporal Memory</i>	Suas implementações operam em tempo real e trabalham bem com tarefas de predição, além de aprender continuamente e modelar as características espaço-temporais de suas entradas. Porém, não aplicam diretamente modelagem em anomalias, o que não permite obter uma classificação útil de anomalias	Ahmad et al. (2017)
<i>xStream</i>	Plataforma flexível que provê dupla comunicação entre aplicações que executam em dispositivos conectados à Internet e na rede. Pode ser usada para melhorar o desempenho de vários tipos de tráfego.	Akyildiz et al. (2018)

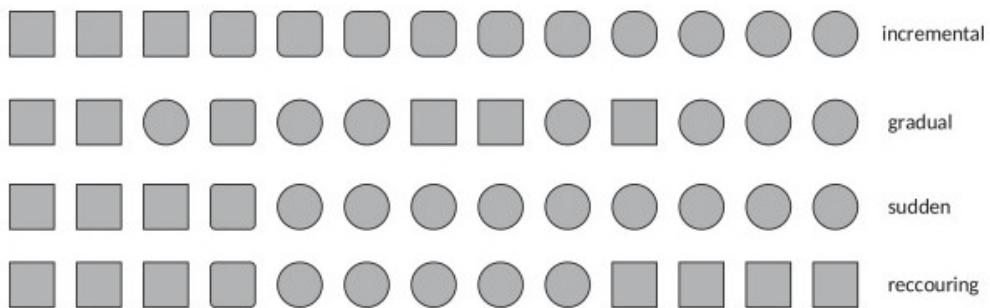
Fonte: Própria autora.

Além dos algoritmos de aprendizado supervisionado ou não supervisionado, podem ser abordados outros conceitos e práticas próprias da análise e mineração de *stream setting*, tais como *concept drifts* (desvios de conceito) e *learning and forgetting* (aprendizagem e esquecimento). A seguir são apresentadas as considerações de Le Nguyen et al. (2020) para estes itens.

*Concept drifts* (desvios de conceito): Krawczyk et al. (2017) definem os *concept drifts* como problemas resultantes de alterações na distribuição do *data stream*. Refletem-se nas instâncias de entrada e deterioram a acurácia de modelos de classificação ou regressão aprendidos nas instâncias anteriores de treinamento. A Figura 6 ilustra algumas padronizações de ocorrências de desvios de conceito, tais como

*incremental* (incremental), *gradual* (gradual), *sudden* (repentina) e *reccouring* (recorrente).

Figura 6 – Tipos de desvio



Fonte: Krawczyk *et al.* (2017).

Um algoritmo incremental, por natureza, pode se adaptar a desvios incrementais através da incorporação contínua de novos dados. Porém, desvios abruptos podem prejudicar a acurácia do modelo até que receba uma quantidade suficiente de dados para adaptar-se ao novo cenário. E, caso ocorram desvios graduais, existe a possibilidade de o modelo não se adaptar às mudanças.

*Learning and forgetting* (aprendizagem e esquecimento): os algoritmos que trabalham de forma *online* possuem a característica de aprender conforme novas instâncias chegam, o que pode ocorrer através de um registro individual ou um conjunto de registros por vez. A segunda forma depende da técnica conhecida por *windowing* (janelas de aprendizagem): um número de registros é acumulado em uma janela de tamanho fixo ou variável e o modelo é atualizado conforme os dados recebidos na última janela. Então os registros antigos são excluídos da janela a fim de receber os novos. O tamanho da janela é um parâmetro essencial. Janelas de tamanho pequeno tornam o modelo sensível a mudanças; porém, são submetidos várias vezes a mudança. Enquanto isso, uma janela grande implica em adaptação mais lenta a mudanças, mas o modelo é mais estável em fases estacionárias dos *streams*.

Por conta da limitação no espaço de memória, a absorção de novos registros implica na exclusão e “esquecimento” dos antigos. Porém, esta propriedade não é adequada para casos em que desvios graduais ou recorrentes ocorrem. A fim de manter registros mais consistentes de janelas anteriores sem recorrer ao armazenamento, a estes dados são atribuídos pesos de acordo com sua idade. Então tais dados são sumarizados e atualizados de acordo com funções estatísticas que determinam a relevância de dados passados.

#### 4.3 Utilização de *data streaming* em aplicações de Manutenção 4.0

É possível aplicar os princípios da Manutenção 4.0 (conectividade entre dispositivos e análise de dados em tempo real) em diversas áreas, além da manufatura. As principais investigações, independente do campo de aplicação, são relacionadas ao monitoramento de variáveis e investigação de causa raiz. A seguir são apresentados três casos em diferentes áreas de aplicação: saúde conforme apresentado por Maktoubian e Ansari (2019), energia a carvão conforme se verifica em Wang *et al.* (2018) e energia eólica conforme trabalho de pesquisa apresentado por Wu *et al.* (2021).

Maktoubian e Ansari (2019) apresentam um caso de aplicação na área da saúde. Trata-se do monitoramento do estado de operação de diversas máquinas e equipamentos hospitalares, baseado em *Internet of Things*. Sensores são instalados em tais máquinas para coletar dados, em tempo real, de variáveis consideradas críticas (vibração, temperatura, umidade e outras), além de alimentar análises preventivas e preditivas. Após a coleta, tais *streams* são enviados para uma camada de análise genérica, na qual poderão ser adotados diversos algoritmos. O resultado da pesquisa foi a proposição de um *framework* de monitoramento de integridade de equipamentos com base em ferramentas de *Big Data* e dados coletados em tempo real.

Wang *et al.* (2018) desenvolveram melhorias na operação de uma central de energia a carvão. Existia um sistema de sensores; porém, eram obtidas apenas leituras individuais e os equipamentos não eram conectados entre si. Além disso, não eram feitas análises preditivas. Entre 2015 e 2016 foram instalados 629 sensores

inteligentes em 21 equipamentos, tais como ventiladores, esteiras e alimentadores, a fim de coletar dados referentes a vibração, eletricidade e alimentação de carvão. É proposto um *framework* de predição de anomalias baseadas em correlação temporal e construído a partir de um conjunto de dados reais e requisitos de manufatura. Minera os grafos de dependência a partir dos dados coletados pelos sensores.

Wu *et al.* (2021) descrevem que, em turbinas eólicas, medições via sensores inteligentes apresentam valores esparsos. Algumas das variáveis medidas são velocidade do vento, velocidade da turbina, temperatura e energia. O modelo sugerido busca a identificação das possíveis causas da grande variação nas medições através de regressão e correlação entre variáveis. O artigo propõe um *framework* de pesquisa para monitoramento em tempo real e diagnóstico de falha na presença dos quatro desafios de dispositivos de IoT para monitoramento: distribuição de dados complexos, estrutura de correlação, esparsidade e eficiência de computação.

A Tabela 8 apresenta os casos destacados acima, de forma resumida, referentes aos exemplos de aplicação de *data streaming* na Manutenção 4.0.

Tabela 8 – Exemplos de aplicação de *data streaming* na Manutenção 4.0

Área	Descrição do caso	Forma de análise	Fonte
Saúde	<i>Data streaming</i> e sensores para coletar leituras de variáveis críticas (vibração, temperatura, umidade e outras) e alimentar análises de falha.	Camada genérica de análise que pode utilizar múltiplos algoritmos.	Maktoubian e Ansari (2019)
Energia a carvão	Habilitar análises preditivas através da instalação de sensores inteligentes que coletam dados referentes a vibração, eletricidade e alimentação de carvão.	Correlação temporal	Wang <i>et al.</i> (2018)
Energia eólica	Identificar possíveis causas pelas quais as medições tomadas pelos sensores inteligentes apresentavam valores esparsos de velocidade do vento, velocidade da turbina, temperatura e energia.	Regressão e correlação entre variáveis	Wu <i>et al.</i> (2021)

Fonte: Própria autora.

Assim, é possível afirmar que os princípios relacionados à Manutenção 4.0 podem ser utilizados para a solução de problemas e implantação de melhorias em infinitas áreas de aplicação, além da indústria.

#### 4.4 Considerações sobre o capítulo

Este capítulo apresenta as relações entre a evolução da manutenção tradicional, conhecida como Manutenção 4.0 e o *data streaming*. Demonstra também que o uso de dados, especialmente em tempo real, é essencial no apoio a iniciativas de manutenção inteligente para garantir a eficiência do processo. A fim de extrair todo o conhecimento que esta grande massa de dados pode fornecer, técnicas de análise de dados são úteis. Técnicas de Manutenção Preditiva não são utilizadas exclusivamente pela manufatura: é possível a aplicação nas mais diversas áreas que podem ser beneficiadas por estudos de causa raiz de problemas e detecção antecipada de possíveis falhas.

## 5 CONSIDERAÇÕES FINAIS

A Manutenção 4.0 é uma realidade tecnológica em plena expansão nos próximos anos. Para que sejam atingidos os objetivos de evitar a ocorrência de falhas, aprimorar a qualidade do produto final, otimizar a utilização de recursos e viabilizar a coleta de dados em tempo real, a adoção de sistemas de manutenção inteligente baseadas em *data streaming* é essencial. Sem estes recursos, as soluções poderão impossibilitar as tomadas de decisão em tempo real caso seja detectado um comportamento anormal e predecessor de falhas. Isto poderá interferir negativamente na atuação no tempo requerido pelo contexto da produção.

### 5.1 Conclusões

A previsão de aumento no investimento na Manutenção 4.0 até 2026 e o crescente interesse de pesquisa neste campo ilustram a relevância do tema na atualidade. Este trabalho está alinhado ao contexto pois a utilização de *data streaming* na manutenção é incipiente e apresenta a necessidade de maior aprofundamento. Estudos de casos em empresas de diversas áreas de atuação demonstram que a adoção de sistemas de manutenção inteligente podem trazer grandes benefícios em pouco tempo, o que possibilita que os recursos outrora empregados na manutenção e recuperação de prejuízos derivados da ocorrência de falhas podem ser aplicados no aprimoramento da cadeia de valor da linha de produção.

Dos conteúdos apresentados neste trabalho, podem-se destacar a apresentação, no Capítulo 2, do ciclo de vida dos dados nos contextos tradicionais, *Internet of Things* e *Big Data*. No Capítulo 3 foram exploradas as abordagens de manutenção, desde as ações corretivas, passando-se pelas preventivas, até chegar à manutenção preditiva. O Capítulo 4 apresentou o uso de dados no contexto da manutenção com ênfase na utilização do *data streaming* como parte essencial em sistemas de Manutenção 4.0.

O método utilizado foi a pesquisa nas bases de artigos científicos *Web of Science* e *Scopus*. Foram selecionados artigos publicados a partir de 2016 a fim de identificar

estudos mais recentes na área de Manutenção 4.0. Além dos artigos científicos, foram utilizados livros especializados e publicações de organizações fornecedoras de tecnologia e relatórios de publicações corporativas, tais como artigos de sites e estatísticas apresentadas por um instituto de pesquisa do posicionamento de *Internet of Things* no contexto organizacional.

Em suma, foi possível identificar que o *data streaming* originado em dados coletados por dispositivos de *Internet of Things* é componente essencial na Manutenção 4.0. Tal aplicação não é exclusivamente utilizada na manufatura e pode ser adotada por diversas áreas que requeiram análises em tempo real de possíveis falhas.

## 5.2 Contribuições do Trabalho

Este trabalho procurou apresentar como o *data streaming*, com origem em dispositivos de *Internet of Things*, apoiam a Manutenção 4.0. Trata-se de sistemas de manutenção preditiva baseados em *data streaming* que apresentam a maior independência possível de intervenção humana e que fornecem informações que possibilitem tomadas de decisão no tempo apropriado ao contexto da organização.

É apresentada a engenharia de dados necessária a fim de viabilizar a implantação de um sistema de Manutenção 4.0. Esta estrutura inicia-se pela coleta de dados através de sensores e dispositivos baseados em *Internet of Things* e termina na fase de descarte, seguindo o ciclo de vida de dados exigido pelas leis gerais de proteção a dados.

Além de apresentar os conceitos e estudos mais recentes no que se refere à aplicação de *data streaming* a sistemas de manutenção, identifica-se a oportunidade de aprofundamento das pesquisas referentes a *stream learning* no contexto da Manutenção 4.0 para melhor aproveitamento dos possíveis *insights* derivados dos dados obtidos.

## 5.3 Trabalhos Futuros

Sugere-se, como trabalhos futuros, maior aprofundamento nos algoritmos e técnicas de *stream learning* na Manutenção 4.0. Além disso, a fim de evidenciar a eficiência dos sistemas de manutenção baseados em dados, podem-se utilizar os conceitos apresentados neste trabalho no projeto e implantação de um sistema de manutenção preditiva baseada em *data streaming* em ambiente real.

## REFERÊNCIAS

- AHMAD, S. *et al.* Unsupervised real-time anomaly detection for streaming data. **Neurocomputing**, v. 262, p. 134-147, 2017.
- AKSA, K. *et al.* Developing a Web Platform for the Management of the Predictive Maintenance in Smart Factories. **Wireless Personal Communications**, p. 1-29, 2021.
- AKYILDIZ, I. *et al.* ARBAT: A flexible network architecture for QoE-aware communications in 5G systems. **Computer Networks**, v. 147, p. 262-279, 2018.
- AMARAL, F. **Introdução à Ciência de Dados**. 1. ed. Rio de Janeiro: Alta Books, 2016.
- AMAZON. **Definição de banco de dados de chave-valor.** Disponível em: <https://aws.amazon.com/pt/nosql/key-value>. Acesso em: 31 ago. 2021.
- AMAZON. **Definição do banco de dados grafo.** Disponível em: <https://aws.amazon.com/pt/nosql/graph>. Acesso em: 31 ago. 2021.
- AMAZON. **O que é um banco de dados em colunas?.** Disponível em: <https://aws.amazon.com/pt/nosql/columnar>. Acesso em: 31 ago. 2021.
- AMAZON. **O que são dados em streaming?.** Disponível em: <https://aws.amazon.com/pt/streaming-data>. Acesso em: 23 jul. 2021.
- ANGELOV, P.; Kordon, A. **Adaptive Inferential Sensors Based on Evolving Fuzzy Models.** In: IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), v. 40, no. 2, p. 529-539. IEEE, 2010.

ANTONY, A.; DEVI, A.; VARGHESE, K. **High Throughput Hardware for Hoeffding Tree Algorithm with Adaptive Naive Bayes Predictor.** In: 2021 6th International Conference for Convergence in Technology (I2CT), p 1-6. IEEE, 2021.

BAHRI, M. *et al.* **Data stream analysis: Foundations, major tasks and tools.** Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, v. 11, n. 3, p. 1-17, 2021.

BOKRANTZ, J. *et al.* Smart Maintenance: an empirically grounded conceptualization. **International Journal of Production Economics**, v. 223, p. 1-17, 2020.

BRASIL. **LEI No 13.709, DE 14 DE AGOSTO DE 2018 Lei Geral de Proteção de Dados Pessoais (LGPD).** Organizado por TEMER, M. *et al.*, 2020. Brasília, DF, 14 ago 2018.

BRUEGGE, F. **Predictive Maintenance Market: The Evolution from Niche Topic to High ROI Application,** 2021. Disponível em: <https://iot-analytics.com/predictive-maintenance-market-evolution-from-niche-topic-to-high-roi-application>. Acesso em: 16 jun. 2021.

CICCONE, B.; BAILEY, S.; LEWIS, J. The Next Generation of Virtual Reality: Recommendations for Accessible and Ergonomic Design. **Ergonomics in Design**, p. 1, 2021.

COHEN, Y. *et al.* Assembly system configuration through Industry 4.0 principles: the expected change in the actual paradigms. **IFAC-PapersOnLine**, v. 50, n. 1, p. 14958-14963, 2017.

CORRAL-PLAZA, D. *et al.* A stream processing architecture for heterogeneous data sources in the Internet of Things. **Computer Standards & Interfaces**, 70, p. 1-13, 2020.

- DUARTE, J.; GAMA, J.; BIFET, A. Adaptive model rules from high-speed data streams. **ACM Transactions on Knowledge Discovery from Data (TKDD)**, v. 10, n. 3, p. 1-22, 2016.
- EBERENDU, C. Unstructured Data: an overview of the data of Big Data. **International Journal of Computer Trends and Technology**, v. 38, n. 1, p. 46-50, 2016.
- FATIMA, H.; WASNIK, K. **Comparison of SQL, NoSQL and NewSQL databases for internet of things**. In: 2016 IEEE Bombay Section Symposium (IBSS). IEEE, 2016. p. 1-6, 2016.
- GAMA, J. *et al.* (Ed.). **IoT Streams for Data-Driven Predictive Maintenance and IoT, Edge, and Mobile for Embedded Machine Learning**: Second International Workshop, IoT Streams 2020, and First International Workshop, ITEM 2020, Co-located with ECML/PKDD 2020, Ghent, Belgium, September 14-18, 2020.
- GILCHRIST, A. **Introducing Industry 4.0**. In: Industry 4.0. Apress, Berkeley, CA, p. 195-215, 2016.
- HASHEM, H.; RANC, D. **Evaluating NoSQL document-oriented data model**. In: 2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), p. 51-56. IEEE, 2016.
- HELU, M. *et al.* Scalable data pipeline architecture to support the industrial internet of things. **CIRP Ann.**, vol. 69, no. 1, p. 385–388, 2020.
- HU, Z.; YU, X., SHI, J., YE, L. Abnormal Event Correlation and Detection Based on Network Big Data Analysis. **Computers, Materials & Continua**, vol. 69, no. 1, p. 695-711, 2021.

IOT-ANALYTICS. **Global IoT Market Forecast** Disponível em: <https://iot-analytics.com/wp/wp-content/uploads/2021/09/Global-IoT-market-forecast-in-billion-connected-iot-devices-min.png>. Acesso em: 24 out. 2021.

IOT-ANALYTICS. **Predictive Maintenance: Niche Topic to Killer Application in 10 Years.** Disponível em: <https://iot-analytics.com/wp/wp-content/uploads/2021/05/Predictive-maintenance-market-niche-topic-to-killer-application-in-10-years-min.png>. Acesso em 24 out 2021.

JACOBS, F.; CHASE, R. **Administração de Operações e da Cadeia de Suprimentos**, 13. ed. Porto Alegre: Mc Graw Hill e Bookman, 2012.

KHARE, S.; TOTARO, M. **Big data in IoT**. In: 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), p. 1-7. IEEE, 2019.

KRAWCZYK, B. *et al.* Ensemble learning for data stream analysis: A survey. **Information Fusion**, v. 37, p. 132-156, 2017.

LE NGUYEN, M.; TURGIS, P.; FAYEMI, E.; BIFET, A. Challenges of Stream Learning for Predictive Maintenance in the Railway Sector. **Cham: Springer International Publishing**, vol. 1325, p. 14-29, 2020.

LUETH, K. **Top 10 IoT applications in 2020**, 2020. Disponível em: <https://iot-analytics.com/top-10-iot-applications-in-2020>. Acesso em: 16 jun. 2021.

MAKTOUBIAN, J.; Ansari, K. An IoT architecture for preventive maintenance of medical devices in healthcare organizations. **Health and Technology**, v. 9, p. 233-243, 2019.

MEHDIPOUR, F.; NOORI, H.; JAVADI, B. **Energy-efficient big data analytics in datacenters**. In: Advances in Computers. Elsevier, p. 59-101, 2016.

MOHAMMADI, M. et al. Deep learning for IoT big data and streaming analytics: A survey. **IEEE Communications Surveys & Tutorials**, v. 20, n. 4, p. 2923-2960, 2018.

MUNIRATHINAM, S. **Chapter Six - Industry 4.0: Industrial Internet of Things (IIOT)**. In: Advances in Computers. Elsevier, v. 117, l. 1, p. 129-164, 2020.

NAVAS, M.; SANCHO, C.; CARPIO, J. Disruptive maintenance engineering 4.0. **International Journal of Quality & Reliability Management**, 2020.

NAVAS, A.; SANCHO, C.; CARPIO, J. Reliability analysis in railway repairable systems. **International Journal of Quality & Reliability Management**, 2017.

NEPOMUCENO, L. **Técnicas de Manutenção Preditiva**. Editora Blucher, 2014.

O'DONOVAN, P. et al. An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. **Journal of Big Data**, a. 15, 2015.

OZTEMEL, E.; GURSEV, S. Literature review of Industry 4.0 and related technologies. **Journal of Intelligent Manufacturing**, 2020.

PFANDZELTER, T.; BERMBACH, D. **IoT data processing in the fog: Functions, streams, or batch processing?**. In: 2019 IEEE International conference on fog computing (ICFC), p. 201-206. IEEE, 2019.

POOR, P.; BASL, J.; ZENISEK, D. **Predictive Maintenance 4.0 as next evolution step in industrial maintenance development**. In 2019 International Research Conference on Smart Computing and Systems Engineering (SCSE), p. 245–253. IEEE, 2019.

PUTINA, A. et al. **Unsupervised real-time detection of BGP anomalies leveraging high-rate and fine-grained telemetry data**. In: IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS),

p. 1-2. IEEE, 2018.

RIZKALLAH, J. **The Big (Unstructured) Data Problem**, 2017. Disponível em: <https://iot-analytics.com/top-10-iot-applications-in-2020>, 05 jun. 2017. Acesso em: 31 ago. 2021.

SAHAL, R.; Breslin, J.; Ali, M. Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. **Journal of Manufacturing Systems**, v. 54, p. 138-151, 2020.

SHIMEI, Lin *et al.* **Design of Industrial Internet of Things Gateway with Multi-source data Processing**. In: 2020 International Conference on Computer Engineering and Application (ICCEA). IEEE, p. 232-236, 2020.

SINAEEPOURFARD, A. *et al.* **Towards a comprehensive data lifecycle model for big data environments**. In: Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies. ACM, p. 100–106, 2016.

SINHA, S. **State of IoT 2021: Number of connected IoT devices growing 9% to 12.3 billion globally, cellular IoT now surpassing 2 billion**. Disponível em: <https://iot-analytics.com/wp/wp-content/uploads/2021/09/Global-IoT-market-forecast-in-billion-connected-iot-devices-min.png>. Acesso em: 24 out. 2021.

SISINNI, S. *et al.* Industrial Internet of Things: Challenges, Opportunities, and Directions. **IEEE Transactions on Industrial Informatics**, v. 14, i. 1, p. 4724 – 4734, 2018.

SLACK, N.; BRENDON-JONES, A.; JOHNSTON, R. **Administração da Produção**, 8. ed. São Paulo: Atlas, 2018.

SONY, S. *et al.* A systematic review of convolutional neural network-based structural condition assessment techniques. **Engineering Structures**, v. 226, p. 1-16, 2021.

STEVAN, S.; LEME, M.; SANTOS, M. **Indústria 4.0: Fundamentos, Perspectivas e Aplicações**. 1. ed. São Paulo: Ed. Érica, p. 140-150, 2018.

WANG, J.; Liu, C.; Zhu, M.; Guo, P.; Hu, Y. **Sensor Data Based System-Level Anomaly Prediction for Smart Manufacturing**. In: IEEE International Congress on Big Data (BigData Congress), p. 158-165. IEEE, 2018.

WIBISONO, A. *et al.* Traffic big data prediction and visualization using fast incremental model trees-drift detection (FIMT-DD). **Knowledge-Based Systems**, v. 93, p. 33-46, 2016.

WU, Z.; Li, Y.; Tsung, F.; Pan, E. Real-time monitoring and diagnosis scheme for IoT-enabled devices using multivariate SPC techniques. **IIE Transactions**, 2018.

YOUSSEF, A., POURGHASEMI, H. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia [J], 12(2): 639-655. **Geoscience Frontiers**, 2021.

ZHOU, X.; Hu, Y.; Liang, W. Ma, J.; Jin, O. Variational LSTM Enhanced Anomaly Detection for Industrial Big Data. **Transactions on Industrial Informatics**, v. 17, n. 5, p. 3469-3477, 2021.