

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA**  
**INSTITUTO DE CIÊNCIAS EXATAS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

**Izaque Esteves da Silva**

**DT-Create – Suíte de serviços para especificação de  
Digital Twins na Indústria 5.0**

Juiz de Fora  
2024

Izaque Esteves da Silva

**DT-Create – Suíte de serviços para especificação de  
Digital Twins na Indústria 5.0**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Ciência da Computação.

Orientadora: Prof. Dr. Regina Maria Maciel Braga

Coorientador: Prof. Dr. José Maria Nazar David

Juiz de Fora

2024

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF  
com os dados fornecidos pelo(a) autor(a)

Esteves da Silva, Izaque.

DT-Create – Suíte de serviços para especificação de  
Digital Twins na Indústria 5.0 / Izaque Esteves da Silva. – 2024.  
100 f. : il.

Orientadora: Regina Maria Maciel Braga

Coorientador: José Maria Nazar David

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Instituto  
de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação,  
2024.

1. Digital Twins. 2. Manutenção Preditiva. 3. Indústria 5.0. I. Braga,  
Regina. II. David, José Maria Nazar.

**Izaque Esteves da Silva**

**DT-Create – Suíte de serviços para especificação de Digital Twins na Indústria 5.0**

Dissertação apresentada ao Programa de Pós-graduação em Ciência da Computação da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Ciência da Computação. Área de concentração: Ciência da Computação.

Aprovada em 22 de março de 2024.

**BANCA EXAMINADORA**

**Prof<sup>ª</sup>. Dra. Regina Maria Maciel Braga Villela** - Orientadora

Universidade Federal de Juiz de Fora

**Prof. Dr. José Maria Nazar David** - Coorientador

Universidade Federal de Juiz de Fora

**Prof. Dr. Victor Ströele de Andrade Menezes**

Universidade Federal de Juiz de Fora

**Prof. Dr. Tadeu Moreira Classe**

Universidade Federal do Estado do Rio de Janeiro

Juiz de Fora, 14/03/2024.

---



Documento assinado eletronicamente por **Regina Maria Maciel Braga Villela, Professor(a)**, em 22/03/2024, às 10:29, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Victor Stroele de Andrade Menezes, Professor(a)**, em 22/03/2024, às 11:12, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Jose Maria Nazar David, Professor(a)**, em 22/03/2024, às 12:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Tadeu Moreira de Classe, Usuário Externo**, em 26/04/2024, às 12:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



A autenticidade deste documento pode ser conferida no Portal do SEI-Uffj ([www2.uffj.br/SEI](http://www2.uffj.br/SEI)) através do ícone Conferência de Documentos, informando o código verificador **1750336** e o código CRC **5B704F9D**.

---

Dedico este trabalho a meus pais, que sempre deram o seu melhor, muitas vezes passando por privações, desejando e se esforçando para me oferecer perspectivas de futuro. Pelo amor, pela fé, pelo cuidado e carinho.

## **AGRADECIMENTOS**

Agradeço a meus orientadores pelo apoio, confiança e profissionalismo na condução de todo o trabalho e, principalmente pela paciência. Sem vocês não seria possível a conclusão dessa jornada. Também à minha esposa Fernanda, pelo apoio incondicional em todos os momentos e por acreditar que era possível.

*"Porque o SENHOR dá a sabedoria, e da sua boca vem a inteligência e o entendimento.". Provérbios 2:6*



## RESUMO

A manutenção preditiva nos cenários da Indústria 4.0 e da Indústria 5.0 demanda soluções inteligentes que garantam a sustentabilidade da indústria no que diz respeito a sua eficiência, minimizando custos associados à manutenções corretivas. Um dos desafios da manutenção preditiva é tomar decisões baseadas em dados de forma ágil e assertiva. Sensores e equipamentos conectados produzem dados operacionais que favorecem o uso de técnicas de processamento inteligente para enriquecer informações e permitir a tomada de decisão. Digital Twins (DTs) são uma representação em tempo real de máquinas físicas e geram dados que a manutenção preditiva pode utilizar a fim de tornar decisões assertivas e rápidas. A principal contribuição deste trabalho é a especificação de uma suíte de serviços para especificação de DTs, a DT-Create, focada no suporte à decisão na manutenção preditiva. A DT-Create é baseada no uso de técnicas inteligentes, processamento semântico de dados e auto adaptação. A DT-Create foi desenvolvida a partir da abordagem epistemológica Design Science Research (DSR), através de dois ciclos de desenvolvimento e avaliada em dois estudos de caso, sendo o primeiro no domínio da indústria têxtil e o segundo no domínio da indústria metalúrgica. Os resultados demonstraram a viabilidade de uso da DT-Create na especificação de DTs considerando os seguintes aspectos: (i) coleta, armazenamento e processamento inteligente de dados gerados por sensores, (ii) enriquecimento de informações por meio de aprendizagem de máquina e ontologias, (iii) uso de técnicas inteligentes para seleção de modelos preditivos aderentes ao conjunto de dados disponível, e (iv) suporte à decisão e auto adaptação a partir de processamentos ontológico.

Palavras-chave: indústria 4.0, indústria 5.0, técnicas inteligentes, *Internet* das coisas, sustentabilidade, manutenção preditiva, suporte à decisão, digital twins, sistemas autoadaptativos

## ABSTRACT

Predictive maintenance in Industry 4.0 and Industry 5.0 scenarios demands intelligent solutions that guarantee the sustainability of the industry in terms of its efficiency, minimizing costs associated with corrective maintenance. One of the challenges of predictive maintenance is making decisions based on data in an agile and assertive way. Connected sensors and equipment produce operational data that favors the use of intelligent processing techniques to enrich information and enable decision-making. Digital Twins (DTs) are a real-time representation of physical machines and generate data that predictive maintenance can use to make assertive and quick decisions. The main contribution of this work is the specification of a suite of services for specifying DTs, DT-Create, focused on decision support in predictive maintenance. DT-Create is based on the use of intelligent techniques, semantic data processing, and self-adaptation. DT-Create was developed using the Design Science Research (DSR) methodology, through two development cycles and evaluated in two case studies, the first in the textile industry and the second in the metallurgical industry. The results demonstrated the feasibility of using DT-Create in specifying DTs considering the following aspects: (i) collection, storage and intelligent processing of data generated by sensors, (ii) enrichment of information through machine learning and ontologies, (iii) use of intelligent techniques to select predictive models that adhere to the available data set, and (iv) decision support and self-adaptation based on ontological processing.

Keywords: *industry 4.0, industry 5.0, intelligent techniques, internet of things, sustainability, predictive maintenance, decision support, digital twins, self-adaptive systems*

## LISTA DE ILUSTRAÇÕES

Figura 1 - Fluxo de trabalho da manutenção preditiva . . . . .	14
Figura 2 - Fluxo DSR . . . . .	17
Figura 3 - Modelo conceito de um sistema auto adaptativo . . . . .	25
Figura 4 - Modelo conceitual de um Digital Twin . . . . .	27
Figura 5 - Processo de filtragem de artigos após consulta à base <i>Scopus</i> . . . . .	32
Figura 6 - Aplicação dos critérios de exclusão. . . . .	33
Figura 7 - Produção científica anual entre 2018 e 2023. . . . .	36
Figura 8 - Redes temáticas. . . . .	37
Figura 9 - Metodologias de avaliação das soluções. . . . .	40
Figura 10 - Elementos DSR. . . . .	47
Figura 11 - DT-Create no primeiro ciclo DSR. . . . .	50
Figura 12 - DT-Create no segundo ciclo DSR. . . . .	52
Figura 13 - Ciclo de AutoML. . . . .	56
Figura 14 - Elementos DSR . . . . .	61
Figura 15 - <i>Dataset</i> com dados de falhas. . . . .	64
Figura 16 - Preparação dos dados antes da criação do modelo preditivo. . . . .	64
Figura 17 - Validação cruzada do modelo preditivo preditivo. . . . .	65
Figura 18 - Matriz de confusão para análise do modelo preditivo. . . . .	66
Figura 19 - Principais classes e associações da ontologia <i>SmartMaintenance</i> . . . . .	67
Figura 20 - Classes e propriedades da ontologia <i>SmartMaintenance</i> . . . . .	68
Figura 21 - Propriedades inferidas pela ontologia ( <i>software</i> Protegé). . . . .	69
Figura 22 - Regra SWRL para geração de alerta representada no <i>software</i> Protegé. . . . .	69
Figura 23 - Adição de coluna ao <i>dataset</i> após processamento estatístico. . . . .	76
Figura 24 - Amostra de dados de um sensor. . . . .	77
Figura 25 - Tabela de comparação entre modelos de aprendizagem de máquina. . . . .	78
Figura 26 - Modelo de aprendizagem de máquina ajustado. . . . .	79
Figura 27 - Ontologia para representação da condição operacional de máquinas. . . . .	80
Figura 28 - Processamento do modelo preditivo com um novo conjunto de dados. . . . .	83
Figura 29 - Dashboard da Aplicação Mobile. . . . .	84

## LISTA DE TABELAS

Tabela 1 – Questões do mapeamento sistemático . . . . .	30
Tabela 2 – Questões do mapeamento sistemático . . . . .	31
Tabela 3 – Trabalhos selecionados . . . . .	33
Tabela 4 – Algoritmos de inteligência artificial utilizados . . . . .	39
Tabela 5 – Arquiteturas para análise de dados . . . . .	41
Tabela 6 – Avaliando temperatura e umidade por regras SWRL . . . . .	68
Tabela 7 – Faixa considerada normal para dados capturados por sensores. . . . .	72
Tabela 8 – Rótulos definidos após a verificação de intervalos de confiança e parâmetros	74
Tabela 9 – Relevância de cada sensor na determinação da condição de uma máquina.	74
Tabela 10 – Combinação entre peso rótulo para estimar a relevância. . . . .	75
Tabela 11 – Cenários de ajuste de valores esperados. . . . .	75
Tabela 12 – Estimativa de condição de máquina a partir de dados de sensores. . . .	76
Tabela 13 – Avaliando a condição operacional de máquinas por regras SWRL. . . .	82

## LISTA DE ABREVIATURAS E SIGLAS

IoT	Internet of Things
DT	Digital Twin
CPS	Sistemas ciber-físicos
DSR	Design Science Research
PLC	Controlador Lógico Programável
IoS	Internet dos Serviços
MQTT	Message Queuing Telemetry Transport
OBDA	Acesso a dados baseado em ontologia
OWL	Web Ontology Language
SI	Sistema Inteligente
IA	Inteligência Artificial
ES	Engenharia de Software
ML	Machine Learning
OI	Ontological Inference
AS	Actual System
SB	Snowballing backward
SF	snowballing forward
LSTM	Rede de memória de longo e curto prazo
RUL	Vida útil restante do sistema
RFID	Identificação por radiofrequência
SWRL	Semantic Web Rule Language
DBMS	MySQL Database Management System
GQM	Goal, Questions, Metrics
CBM	Manutenção Baseada em Condição
MCC	Coefficiente de Correlação de Matthews
MSL	Mapeamento Sistemático da Literatura

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>13</b>
1.1	CONTEXTUALIZAÇÃO . . . . .	13
1.2	MOTIVAÇÃO . . . . .	15
1.3	OBJETIVOS . . . . .	17
1.4	ORGANIZAÇÃO . . . . .	18
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>19</b>
2.1	MANUTENÇÃO PREDITIVA . . . . .	19
2.2	SENSORES E COMUNICAÇÃO DE DADOS . . . . .	19
2.3	ONTOLOGIAS . . . . .	21
2.4	SISTEMAS INTELIGENTES . . . . .	23
2.5	ARQUITETURAS AUTO ADAPTATIVAS . . . . .	24
2.6	DIGITAL TWINS . . . . .	26
2.7	CONSIDERAÇÕES FINAIS DO CAPÍTULO . . . . .	28
<b>3</b>	<b>MAPEAMENTO SISTEMÁTICO DA LITERATURA . . . . .</b>	<b>29</b>
3.1	PLANEJAMENTO . . . . .	29
3.2	CONDUÇÃO . . . . .	32
3.3	ANÁLISES . . . . .	38
3.4	RESULTADOS . . . . .	41
3.5	DISCUSSÃO . . . . .	45
3.6	CONSIDERAÇÕES FINAIS DO CAPÍTULO . . . . .	45
<b>4</b>	<b>DT-CREATE – SUÍTE DE SERVIÇOS PARA ESPECIFICAÇÃO DE DIGITAL TWINS NA INDÚSTRIA 5.0 . . . . .</b>	<b>46</b>
4.1	DSR . . . . .	46
4.2	REQUISITOS FUNCIONAIS E NÃO FUNCIONAIS . . . . .	48
4.3	PRIMEIRO CICLO DA DSR . . . . .	49
4.4	SEGUNDO CICLO DA DSR . . . . .	51
4.5	DT-CREATE: PRINCIPAIS COMPONENTES . . . . .	52
4.6	USO DE ALGORITMOS DE APRENDIZAGEM DE MÁQUINA . . . . .	54
4.7	USO DE ONTOLOGIAS . . . . .	56
4.8	ONTOLOGIA E APRENDIZAGEM DE MÁQUINA . . . . .	57
4.9	AUTO ADAPTAÇÃO . . . . .	57
4.10	ESTRATÉGIAS DE IMPLEMENTAÇÃO . . . . .	57
4.11	CONSIDERAÇÕES FINAIS DO CAPÍTULO . . . . .	58
<b>5</b>	<b>AVALIAÇÃO DA DT-CREATE . . . . .</b>	<b>59</b>
5.1	DEFINIÇÃO DE ESCOPO . . . . .	59
5.2	ESPECIFICAÇÃO DAS ONTOLOGIAS . . . . .	60
5.3	CENÁRIOS DE AVALIAÇÃO . . . . .	62

5.4	ESTUDO DE CASO 1 (PRIMEIRO CICLO DSR) . . . . .	63
5.4.1	CONDUÇÃO . . . . .	63
5.4.2	ONTOLOGIA SMARTMAINTENANCE . . . . .	66
5.4.3	ANÁLISE . . . . .	70
5.4.4	RESULTADOS . . . . .	70
5.5	ESTUDO DE CASO 2 (SEGUNDO CICLO DSR) . . . . .	71
5.5.1	CONDUÇÃO . . . . .	71
5.5.2	ONTOLOGIA SENSOREQUIPAMENT . . . . .	80
5.5.3	ANÁLISE . . . . .	83
5.6	RESULTADOS . . . . .	85
5.7	TRIANGULAÇÃO DOS RESULTADOS . . . . .	86
5.8	AMEAÇAS À VALIDADE . . . . .	87
5.9	CONSIDERAÇÕES FINAIS DO CAPÍTULO . . . . .	88
6	<b>CONCLUSÃO . . . . .</b>	<b>90</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>93</b>

# 1 INTRODUÇÃO

## 1.1 CONTEXTUALIZAÇÃO

A percepção atual das características da indústria resulta de uma série de transformações motivadas pela necessidade de produção com alta qualidade e capacidade de entregas cada vez mais rápidas. Ao longo dos anos as necessidades dos consumidores mudaram consideravelmente, o que demandou alta capacidade de flexibilização das fábricas para um atendimento cada vez mais direcionado ao cliente (Cheng et al., 2015). Esse processo constante de transformação marcou a história do desenvolvimento industrial em quatro estágios de mudança significativa nos métodos de produção e tecnologias aplicadas.

A Indústria 1.0 foi representada pela primeira revolução industrial onde, no final do século XVIII máquinas a vapor substituíram a força do trabalho manual. A Indústria 2.0 ou segunda revolução industrial, foi marcada pela introdução da energia elétrica, ampliando o trabalho e sua divisão. No terceiro estágio da revolução industrial (ou Indústria 3.0) temos a inserção de novas formas de energia, como a energia nuclear, e muitos avanços na eletrônica. Passa-se então a utilizar sistemas computacionais para projeto de fabricação de produtos, como sistemas CAD e CAM. Também há a utilização da robótica no ambiente industrial, aumentando a precisão e performance das plantas industriais (Dorst et al., 2016). Como consequência, com o aumento da capacidade produtiva passou-se a pensar em como a manutenção poderia ser realizada de forma menos impactante em tempo e recursos, minimizando perdas. Assim, surgem estudos sobre a manutenção preventiva e preditiva de equipamentos (Womack et al., 2007).

O estágio atual consolidado de desenvolvimento da indústria é denominado Indústria 4.0, caracterizado basicamente pelo uso de tecnologias avançadas e pela *Internet*, novamente transformando o panorama industrial (Lasi et al., 2014). Nesse estágio temos "coisas"(equipamentos) que são capazes de realizar processamento de dados e de estabelecer conexões com outros dispositivos por meio de redes de comunicação, o que é chamado de *Internet* das Coisas (do inglês *Internet of Things* – IoT) (Ashton, 2009), importante componente na Indústria 4.0. Tais dispositivos são capazes de enviar e receber informações por meio da *Internet* (Santos et al., 2016).

Entretanto, estamos em um momento de transição para uma nova evolução: a Indústria 5.0. De acordo com a Comissão Europeia para Pesquisa e Inovação (Breque et al., 2021) a Indústria 5.0 é uma extensão da Indústria 4.0, que tem o ser humano como ator central do processo produtivo, além de focar na sustentabilidade da indústria em relação aos recursos cada vez mais escassos e na resiliência. O objetivo é criar uma cadeia de valor que vai além da relação custo-eficiência ou maximização de lucros, mas favorecer todas as partes interessadas, ou seja, investidores, trabalhadores, consumidores, sociedade e ambiente (Breque et al., 2021).



Figura 1 - Fluxo de trabalho da manutenção preditiva



Fonte: Achouch et al. (2022).

Observando as tecnologias habilitadoras da Indústria 4.0 englobando modelos de processos cada vez mais informacionais, e com vistas aos objetivos da Indústria 5.0, um dos desafios relacionados aos processos de manutenção é a viabilização da transição dos modelos de manutenção corretiva e preventiva para o modelo de manutenção preditiva. A manutenção preditiva antecipa os problemas, a partir da coleta e preparação dos dados para posterior análise, definindo um modelo para análise de dados, que permite a criação de mecanismos para auxiliar na tomada de decisão (Achouch et al., 2022).

A Figura 1 mostra um ciclo de vida da manutenção preditiva. O ciclo se inicia com a compreensão dos problemas e restrições de um projeto, estabelecendo as grandezas a serem medidas, quais serão os sensores e onde serão instalados e o levantamento dos possíveis tipos de falhas. Com a infraestrutura estabelecida, a coleta de dados é realizada por sensores nos equipamentos. Na etapa de compreensão e preparação, busca-se identificar quais dados serão analisados e construir significado para os dados. Assim, os dados passam por processos de limpeza, gerenciamento do volume de dados, correções de dados ausentes ou anômalos e seleção de tipos de dados de maior interesse, entre outras etapas. Na etapa de modelagem de dados é feita a escolha do procedimento mais adequado para criar um modelo capaz de realizar previsões com a maior precisão possível. Na etapa de implantação o modelo é posto à prova quanto à sua precisão e relevância. O ciclo termina com a tomada de decisão, que inclui resolver problemas e tomar decisões, selecionar os procedimentos mais adequados para minimizar custos e atrasos e avaliar decisões a fim de melhorar intervenções futuras.

Analisando este ciclo de vida da manutenção preditiva, podemos verificar que os processos estão cada vez mais voltados para antecipar decisões, para apoiar de forma mais automática pessoas (humanos) envolvidas nos processos industriais. Além disso,

pode-se evidenciar uma preocupação com a sustentabilidade, uma vez que os processos são voltados para a longevidade de equipamentos. Neste sentido, a manutenção preditiva pode ser considerada um processo importante com vistas à indústria 5.0.

Um dos principais desafios para a criação de sistemas de manutenção preditiva é a falta de dados de eventos de interesse, uma vez que a máquina é frequentemente reparada antes de ocorrer qualquer evento característico para manutenção. Neste sentido, para um ciclo virtuoso com vistas a manutenção preditiva, é necessário dispositivos que capturem, sem intervenções no sistema produtivo, dados de eventos de interesse. Os *Digital Twins - DTs*, ou Gêmeos Digitais, fornecem uma representação em tempo real da máquina física e geram dados que podem ser utilizados na manutenção preditiva (van Dinter et al., 2022). Assim, é possível operacionalizar a manutenção preditiva sem afetar diretamente os equipamentos em produção.

O uso de DTs ajuda a lidar de forma sistematizada com os desafios de controle em tempo real de sistemas de produção, uma vez que o DT representa de forma simulada um sistema de produção, onde análises podem ser antecipadas, permitindo a sincronização entre sistemas reais e virtuais (Qi and Tao, 2018).

A literatura atual e Indústria 4.0 discutem o uso de DTs em processos produtivos (van Dinter et al., 2022). O rápido aumento de dados de sensores digitais de máquinas e a disponibilidade para usar esses dados por meio de IoT permitem que as empresas possam tomar decisões baseadas em dados. Assim, nesse contexto, DTs são réplicas digitais de um sistema ou componente físico. Além disso, eles utilizam dados do sistema físico como entradas para simular como esses dados afetam o sistema produtivo. Os DTs têm muitas variantes, como DTs que simulam componentes, ativos, sistemas e processos.

No entanto, mesmo com o uso de DTs, o cenário da manutenção no contexto da Indústria 4.0 é cada vez mais desafiador no que diz respeito à velocidade com que a tomada de decisão deve ocorrer. Adicionalmente, a colaboração entre os atores, se mostra necessária nos processos decisórios, pois se aproveita das *expertises* de diferentes áreas e posições. Também a agilidade na produção de informações confere às decisões maior grau de antecipação em relação às falhas e assertividade. É necessário, portanto, criar um cenário de tomada de decisão que possua riqueza de informações e utilize tecnologias que permitam a descoberta antecipada de informações, de maneira ágil.

## 1.2 MOTIVAÇÃO

IoT e o uso de DTs envolvem um novo paradigma tecnológico que reúne o mundo físico e virtual para fornecer sistemas de *software* em todos os lugares através de objetos da vida cotidiana, sendo um dos principais componentes da Indústria 4.0. Entretanto, ainda há uma falta de consenso e compreensão de conceitos e recursos que melhor se adaptam ao desenvolvimento de sistemas de *software* complexos considerando IoT e DTs (Weyns,

2020).

Nesse sentido, com o uso de IoT e DTs, as tecnologias relacionadas ao desenvolvimento de *software* enfrentam desafios para desenvolver serviços inteligentes conectados a produtos inteligentes e sistemas ciber-físicos (CPS). Assim, cresce o interesse por novas estratégias de desenvolvimento de *software*, derivando novas abordagens para o desenvolvimento de sistemas complexos (Weyns, 2020). A engenharia de *software* auto adaptativa (Eramo et al., 2021), abrangendo consciência do contexto e estratégias inteligentes, é uma abordagem que visa resolver as complexidades sob uma nova perspectiva.

Por outro lado, na Indústria 5.0, a manutenção precisa lidar com a rapidez com que a tomada de decisões precisa ocorrer, de forma a considerar consistentemente os fatores humanos e a sustentabilidade. Uma das dificuldades está relacionada justamente ao monitoramento em tempo real e tratamento de problemas operacionais nos processos produtivos. Além disso, tem-se desafios que envolvem gerenciamento de recursos, produção cada vez mais customizada e processos que necessitam de gestão com respostas cada vez mais rápidas. Nesse contexto, é necessário utilizar tecnologias que favoreçam o desenvolvimento de soluções inteligentes para essas demandas.

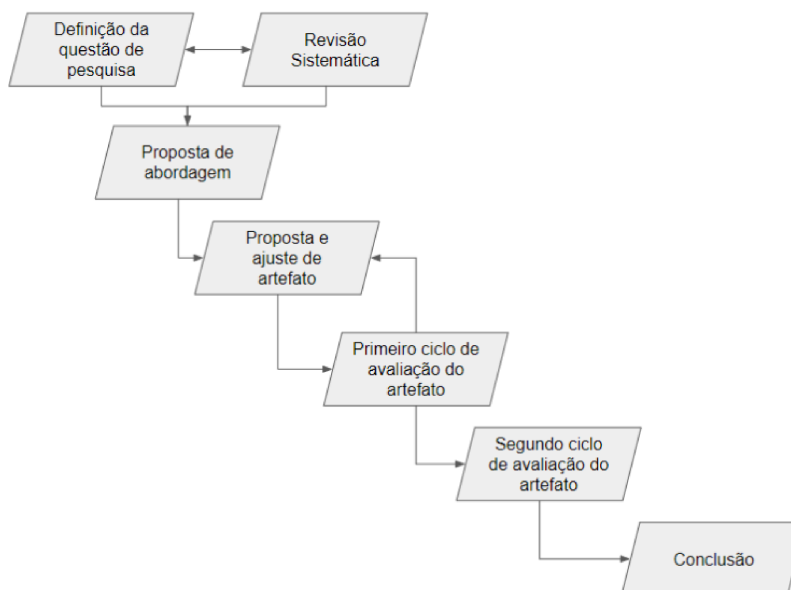
Portanto, este trabalho é orientado pela seguinte questão de Pesquisa (QP): “Como apoiar a tomada de decisão relacionada à manutenção preditiva na Indústria 5.0 com o uso de DTs?”. Para isso, utilizamos a abordagem *Design Science Research - DSR* (Hevner et al., 2008) para apoiar o objetivo da pesquisa por meio do artefato produzido.

A DSR é uma abordagem caracterizada pela melhoria contínua uma solução por meio da introdução de novos artefatos e de processos de construção dos mesmos (Simon, 1996). Tem como objetivos (i) desenvolver artefatos, que podem ser considerados produtos resultantes de um projeto, criados para resolver problemas reais em contextos específicos; e (ii) gerar novo conhecimento técnico e científico (Pimentel et al., 2020). A DSR, portanto, envolve a criação de modelos, protótipos ou sistemas e sua avaliação em termos de utilidade, eficácia e impacto.

Além disso, a DSR envolve ciclos de avaliação. Nesta dissertação, foram realizados 2 ciclos de avaliação, conforme apresentado na Figura 2, sendo o primeiro no contexto de processos produtivos relacionados à indústria têxtil e o segundo ciclo com dados de sensores instalados em fornos de tratamento térmico.

Considerando os problemas relacionados à manutenção preventiva, sistemas auto adaptativos podem fornecer estratégias para melhorar as decisões relacionadas aos sistemas complexos demandados pela Indústria 5.0. Conjuntamente com a IoT e o processamento de grandes volumes de dados, esses sistemas permitem a captura e processamento inteligente de dados que podem enriquecer decisões. Desta forma, este trabalho apresenta, a partir da especificação de DTs, uso de arquiteturas auto adaptativas, técnicas inteligentes e análises de dados, um conjunto de serviços para o apoio à manutenção preditiva na Indústria 5.0.

Figura 2 - Fluxo DSR



Fonte: Elaborada pelo autor (2023).

### 1.3 OBJETIVOS

A utilização de DTs no gerenciamento da manutenção preditiva representa ganho tanto no planejamento de atividades de manutenção quanto na otimização de tarefas, o que em linhas gerais representa ganhos financeiros para o sistema de produção (Negri and Fumagalli, 2017).

No contexto desta dissertação, a utilização de técnicas de aprendizagem de máquina (do inglês *Machine Learning* - *ML*) está inserida em um ambiente onde modelos preditivos são selecionados de acordo com os conjunto de dados, permitindo a descoberta de conhecimento sobre dados com maior assertividade. De forma complementar, modelos semânticos, como ontologias (Gruber, 1995), são utilizados para enriquecimento das informações obtidas a partir da avaliação dos modelos preditivos, além de informações contextuais importantes para a tomada de decisão. Desta forma, aliamos análises sintáticas e semânticas para o enriquecimento de dados (Amershi et al., 2019).

Este trabalho tem como objetivo apoiar processos de tomada de decisão relacionados à manutenção preditiva, através de uma abordagem que utiliza técnicas inteligentes para enriquecer informações sobre processos produtivos. Uma suíte de serviços é proposta com o objetivo de fornecer serviços específicos para auxiliar na especificação de DTs, com vistas a apoiar o suporte à manutenção na Indústria 5.0. Para isso, a suíte de serviços utiliza arquiteturas auto adaptativas, ontologias e aprendizagem de máquina como tecnologias habilitadoras dos serviços.

Portanto, desenvolvemos a DT-Create, uma suíte de serviços para apoio a espe-

cificação de DTs que tem como objetivo apoiar a análise de dados de sensores e outros dispositivos utilizados nos processos de gerenciamento de manutenção industrial. A DT-Create é utilizada no contexto da Indústria 5.0 para apoiar o gerenciamento de máquinas e equipamentos. Utilizamos a abordagem DRS-Model (Pimentel et al., 2020) para criar duas versões do artefato, que foram avaliadas em dois estudos de caso. O primeiro cenário de avaliação utilizamos dados de uma indústria têxtil e o segundo cenário dados de uma indústria metalúrgica.

#### 1.4 ORGANIZAÇÃO

Este trabalho está dividido em seis capítulos. O Capítulo 2 apresenta a fundamentação teórica. O Capítulo 3 discute os trabalhos relacionados, a partir de um Mapeamento Sistemático da Literatura - MSL. O Capítulo 4 apresenta a metodologia utilizada baseada na abordagem Design Science Research, a suíte de serviços proposta para apoiar a tomada de decisão na manutenção preditiva, detalhando os aspectos conceituais e a implementação da suíte de serviços. O Capítulo 5 detalha dois ciclos de avaliação da suíte de serviços no apoio ao desenvolvimento de DTs, destacando o planejamento, execução e resultados obtidos nos dois ciclos de avaliação. O Capítulo 6 apresenta as considerações finais, discute as contribuições do trabalho, suas limitações e trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, discutimos os principais conceitos relacionados à proposta desta dissertação, com o objetivo de facilitar o entendimento de tecnologias específicas utilizadas.

### 2.1 MANUTENÇÃO PREDITIVA

Podemos conceituar Manutenção Preditiva como o monitoramento regular da condição mecânica real, da eficiência operacional e de outros indicadores da condição operacional de máquinas e sistemas de produção, que visa minimizar o número e o custo de interrupções não programadas ocasionadas por falhas de máquinas (Mobley, 2002). A manutenção preditiva é realizada considerando a condição dos equipamentos, definida a partir da coleta dados obtidos de sensores e outros equipamentos de instrumentação. A partir da aplicação de recursos computacionais, como aprendizagem de máquina e inteligência artificial, é possível detectar padrões que podem sinalizar desgaste e permitir ações de manutenção em estágios ainda iniciais, onde falhas ainda são consideradas potenciais (Hansen, 2001).

Como estratégia de sustentabilidade industrial, a manutenção preditiva contribui para o aumento da vida útil de equipamentos, redução de desperdícios decorrentes de falhas de máquinas ou substituições frequentes de peças. Também contribui para a eficiência operacional das empresas, aumentando a eficiência dos processos a partir da detecção de necessidades de manutenção com maior grau de antecipação e assertividade (Jardine e Tsang, 2005).

A manutenção preditiva utiliza tecnologias habilitadoras da Indústria 4.0, como IoT e ML, que permitem maior integração entre dados e sistemas, favorecendo a tomada de decisão em tempo real e contribuindo para a sustentabilidade industrial. Além disso, a capacidade de monitoramento constante do estado e do desempenho dos equipamentos em uma fábrica, conferem agilidade na resposta às demandas operacionais e às necessidades de mercado (Wang, 2016).

### 2.2 SENSORES E COMUNICAÇÃO DE DADOS

Sensores são componentes críticos em sistemas industriais automatizados, pois fornecem dados a sistemas de controle, favorecendo sua operação precisa, eficiente e segura (Soloman, 2009). São projetados para detectar, medir e responder a diferentes estímulos físicos, químicos ou biológicos em um ambiente industrial. Também são capazes de medir grandezas como temperatura, umidade, pressão, níveis de sinais elétricos, entre outros e o resultado das medições pode ser processado por meio de Controladores Lógicos Programáveis (do inglês *Programmable Logical Controls - PLCs*) ou computadores industriais

(Bolton, 2021).

Para criar um ambiente informacional ativo na Indústria, os sensores são componentes essenciais no contexto da Indústria 4.0, contribuindo para a produção de conhecimento a partir de dados de equipamentos e processos (Santos et al., 2018). De acordo com (Cornago et al., 2022), as tecnologias habilitadoras da Indústria 4.0 reforçaram a necessidade do uso de sensores para coleta de dados em tempo real. Permitem o monitoramento de processos de produção, a garantia da eficiência e a qualidade dos produtos. Os sensores são utilizados tanto para a coleta de dados operacionais quanto para as informações em tempo real sobre equipamentos e ambiente, como, por exemplo, temperatura, pressão, umidade e vibração. As informações podem ser, posteriormente, analisadas para o planejamento das atividades de manutenção e tomada de decisão sobre os processos.

Assim, a adoção crescente de sensores na indústria está moldando a maneira como as empresas operam, tornando os processos mais eficientes, reduzindo custos e melhorando a qualidade. De acordo com (Zhong et al., 2017), essa revolução sensorial possibilita um ambiente de produção mais conectado e inteligente, onde dados precisos e em tempo real são centrais para a tomada de decisões estratégicas das organizações. A junção da utilização em larga escala de sensores e a presença cada vez maior de dispositivos conectados, permitem a rápida aquisição e comunicação das informações através da IoT, criando uma rede de comunicação rápida e que permite a análise de dados para obtenção de *insights* para apoio na tomada de decisão.

No entanto, o panorama atual da indústria contempla várias outras tecnologias digitais associadas à aquisição e análise de dados em tempo real, como, por exemplo, computação em nuvem e análise de grandes volumes de dados (Frank et al., 2019). Da mesma forma, a *Internet* dos Serviços - IoS contempla sensores, atuadores e dispositivos IoT integrados por uma rede de comunicação comum industrial e sistemas de produção conectados à *Internet*, além de técnicas de análise de dados, figurando assim aplicações inteligentes (Drath and Horch, 2014).

Para permitir a comunicação entre os vários dispositivos que são a base das aplicações inteligentes no ambiente industrial, é necessário um protocolo adequado para trafegar os dados tanto em quantidade quanto em velocidade. O protocolo *Message Queuing Telemetry Transport - MQTT* (Hillar, 2017) é amplamente utilizado em situações nas quais conjugar pouca capacidade de processamento computacional e pouca largura de banda são componentes críticos, como por exemplo equipamentos instalados em locais remotos com pouca infraestrutura de energia elétrica e comunicação de dados.

O protocolo MQTT funciona como um modelo de publicação-assinatura para troca de mensagens entre dispositivos. Foi projetado para ser mais leve que o protocolo HTTP 1.1<sup>1</sup> e ser mais confiável em cenários de redes com conectividade intermitente. Dessa forma, o

---

<sup>1</sup> Primeira versão do Protocolo de Transferência de Hipertexto, usado principalmente para a

MQTT é mais adequado para cenários onde vários dispositivos precisam trocar dados quase em tempo real usando a menor largura de banda possível (Hillar, 2017). Funciona a partir de duas entidades principais: *client* e *broker*. Os clientes são os próprios dispositivos que produzem e comunicam dados e informações, já o *broker* é um intermediário responsável por gerenciar e distribuir as mensagens através da rede. A troca de mensagens é feita a partir da conexão de um cliente com o *broker* e essa conexão pode ter duas intenções: a de assinatura de tópicos, quando o cliente espera receber informações do *broker*, ou de publicação de mensagens (Soni and Makwana, 2017). Portanto, o protocolo MQTT é simples no que diz respeito ao consumo de recursos, uma vez que a comunicação é feita sob demanda, ou seja, só há transmissão de informações quando realmente é necessário, o que contribui para economia de processamento de informações e uso do meio de transmissão.

### 2.3 ONTOLOGIAS

Ontologias são representações formais de conhecimento que definem conceitos, relações e regras que regem essas relações dentro de um determinado domínio. Gruber (Gruber, 1995) define ontologias no contexto computacional como uma especificação formal e explícita de uma conceituação compartilhada. Tal conceituação é uma visão do que se deseja representar em um nível simplificado e abstrato, de acordo com o propósito.

No contexto da Indústria 4,0 e conseqüente na 5.0, grande quantidade de dispositivos e sistemas interconectados produzem um vasto volume de dados. O uso de ontologias apoia o compartilhamento de conhecimento entre estes sistemas e dispositivos, através de um vocabulário comum. Além disso, permitem a descoberta de novos relacionamentos entre os dados, permitindo a descoberta de novas informações. Essas características estão alinhadas diretamente com as vantagens no uso de ontologias, como: processamento inteligente, complementação de dados faltantes e compreensão de domínio, permitindo um melhor suporte a descoberta de práticas que levem à sustentabilidade (Ushold and Gruninger, 1996).

Para a especificação de ontologias a OWL (*Web Ontology Language*) é geralmente utilizada. A OWL é uma linguagem projetada para representar o conhecimento complexo sobre termos, grupos de termos e relações entre termos de uma forma legível por máquinas, permitindo o compartilhamento e reuso de dados entre aplicações diferentes. De acordo com (McGuinness, 2008) a OWL supera outras linguagens como XML, RDF e RDF Schema (RDF-S) pois combina um vocabulário adicional com semântica. Sua principal vantagem está relacionada ao processamento de informações contidas em documentos por meio de aplicativos, ao contrário de outras soluções que focam na representação da informação a humanos.

---

comunicação de dados na World Wide Web



Existem vários *frameworks* e ferramentas que permitem a manipulação de ontologias, facilitando a criação, manipulação, e aplicação de ontologias em sistemas de informação. O *Protege* (Musen, 2015) é uma ferramenta gratuita e de código aberto para modelagem de ontologias e construção de sistemas inteligentes. Possui interface gráfica que facilita a criação e edição de classes, propriedades e instâncias. Suporta *reasoners* padrão OWL como *Pellet*, *MermiT* e *FacT++*. Também suporta diferentes formatos de serialização de ontologias como *Turtle*, OWL, XML, RDF/XML, entre outros (Musen, 2015). Semelhante ao *Protege*, temos a OWL API escrita em Java e capaz de manipular e serializar ontologias OWL e temos também o *Apache Jena* que além de manipulação de ontologias oferece funcionalidades de persistência e consulta a dados semânticos.

O framework OWL2ready (Jean-Baptiste, 2021) é uma ferramenta de modelagem ontológica que permite a integração e interoperabilidade de dados semânticos em sistemas de informação. Proporciona um ambiente robusto para o desenvolvimento de ontologias complexas com suporte a raciocínio lógico e alinhamento de vocabulários diversos. Este *framework* é destacado por sua eficiência na facilitação da criação e manutenção de ontologias, tornando-se uma escolha popular para projetos de dados relacionados e semânticos, especialmente para aplicações escritas na linguagem de programação *Python*.

No contexto processos produtivos, uma ontologia pode ser utilizada para representar os diferentes componentes Indústria 5.0, como máquinas, produtos, processos, recursos e operadores, bem como suas interações e propriedades. Também podem ser utilizadas para padronizar a representação dos dados sobre processos, facilitando a troca e o compartilhamento de informações entre sistemas heterogêneos, como é apresentado por (Noy and McGuinness, 2001). Dessa forma, ontologias são aplicadas no contexto da Indústria 5.0 para organizar, analisar e permitir a integração semântica de dados de fontes distintas.

Na manutenção preditiva e preventiva as ontologias podem ser utilizadas na modelagem do domínio da manutenção, capturando informações sobre tipos de equipamentos, estados funcionais, tipos de falhas, informações de contexto entre outros. Podem atuar também na integração de dados de diferentes fontes, como dados históricos de manutenção e dados de monitoramento, oferecendo como contribuição a possibilidade de realizar inferências lógicas com base nos dados obtidos de sensores, fazendo verificações e atuando na correção de inconsistências que podem ocorrer na aquisição de dados (Steinmetz, 2018).

Além de representar e integrar informações, as ontologias podem produzir conhecimento novo, a partir do uso de algoritmos que atuam sobre os dados, auxiliando na detecção de padrões e falhas ou ainda na geração de alertas relacionados à manutenção, oferecendo suporte à decisão e permitindo a especificação de sistemas inteligentes, considerando o processamento semântico dos dados.

## 2.4 SISTEMAS INTELIGENTES

De acordo com (Russell and Norvig, 2010) Sistemas Inteligentes (SIs) são estruturas computacionais capazes de realizar tarefas que normalmente exigiriam inteligência humana. SIs simulam processos de cognição humana, como aprendizado, análise de dados, tomada de decisões entre outros e associam tecnologias como Aprendizado de Máquina, Lógica *Fuzzy*, Redes Neurais, ontologias e outras técnicas inteligentes para criar sistemas capazes de realizar tarefas complexas de forma autônoma ou com mínima supervisão humana (Negnevitsky, 2005). Estes sistemas podem ser aplicados em diversos setores, como automotivo, manufatura, saúde, finanças, entretenimento, entre outros, oferecendo soluções que vão desde robôs autônomos e assistentes virtuais até sistemas de análise de dados e monitoramento inteligente de processos.

A construção de sistemas inteligentes difere das abordagens tradicionais de desenvolvimento de *software*, tornando as técnicas mais desafiadoras. A Engenharia de *Software* (ES) tradicional envolve a coleta de requisitos, análise e projeto detalhado para implementar um programa executável. O desenvolvimento de *software* com componentes de IA envolve recursos adicionais, como coleta de dados, seleção de um algoritmo apropriado (por exemplo, algoritmos de ML ou *Ontological Inference* (OI)) e treinamento de modelos com base na entrada/saída desejada (Ahmad et al., 2023).

Na indústria, os SIs contribuem para aumento de eficiência de processos em cadeias de suprimentos, que inclui previsão da demanda, gestão de estoque, a otimização de rotas de entrega e conseqüente redução de custos de transporte (Min, 2010). Também podem contribuir em processos de inspeção em linhas de produção, para identificação de defeitos ou desvios de padrões de qualidade, por meio do aprendizado a partir de dados históricos e evolução de modelos preditivos (Chouchene et al., 2020). No contexto da Indústria 4.0, SIs geralmente são especificados considerando tecnologias como IoT, grandes volumes de dados e outras tecnologias habilitadoras para coletar, analisar e ser proativo, considerando informações em tempo real. Na Indústria 5.0, SIs podem envolver novas funcionalidades, envolvendo o suporte a sustentabilidade de equipamentos e processos e auxiliando atores humanos no desempenho otimizado de suas tarefas.

Com SIs, a previsibilidade no gerenciamento de fábricas se torna possível a partir de soluções de *software* integradas ao ambiente de produção. A partir do uso de sensores IoT, os equipamentos são capazes de produzir dados durante suas operações e com o uso de algoritmos inteligentes, esses dados podem ser analisados para detecção de padrões de interesse e anomalias (Zhong et al., 2017). No contexto da manutenção preditiva, é possível prever falhas potenciais ou identificar a necessidade de manutenção, permitindo que as indústrias atuem de forma proativa e preventiva, em direção a sustentabilidade, economia de recursos e livrando operadores de tarefas consideradas extenuantes e tediosas. No entanto, para que SIs possam ser especificados, técnicas específicas de desenvolvimento de

*software* devem ser utilizadas. Neste sentido, uma das possibilidades é o uso de arquiteturas auto adaptativas, discutida na próxima seção.

## 2.5 ARQUITETURAS AUTO ADAPTATIVAS

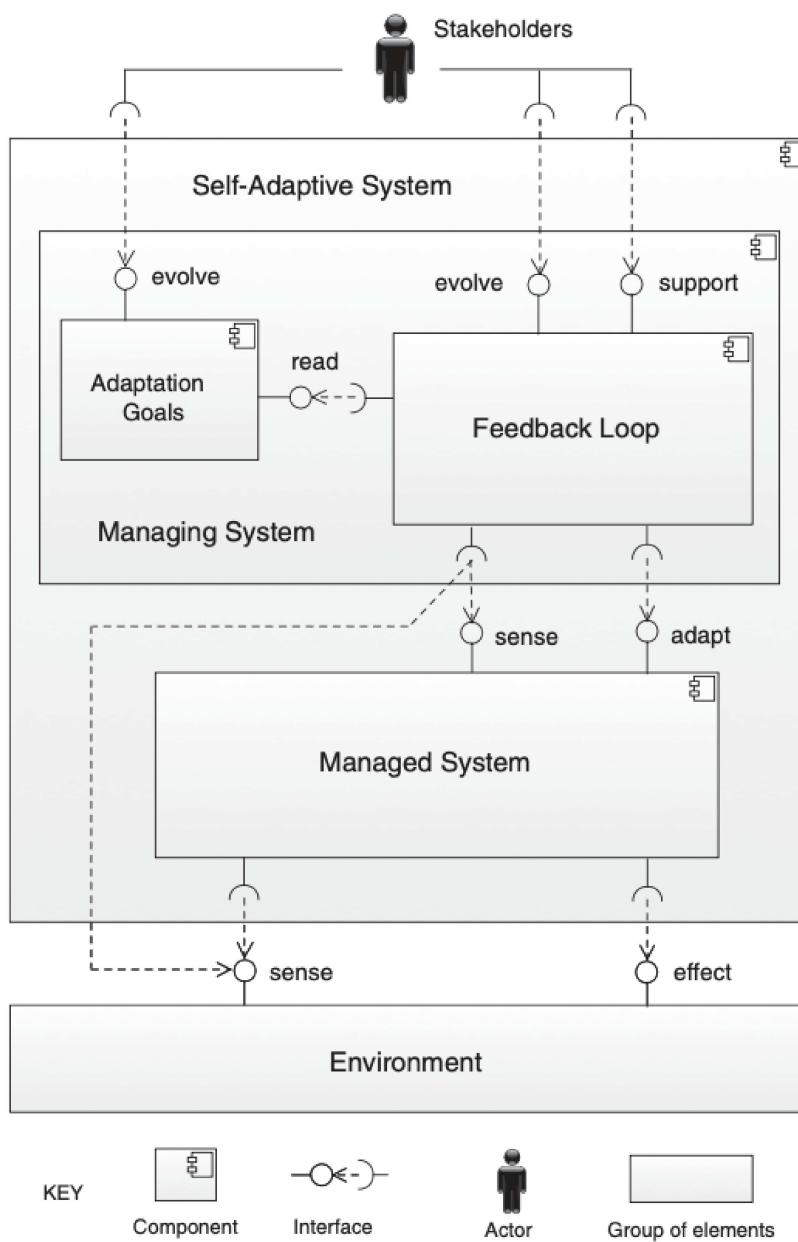
Arquiteturas de *software* auto adaptativas (ou sistemas auto adaptativos) são projetadas para ajustar automaticamente seu comportamento em tempo de execução, em resposta a mudanças no ambiente ou no próprio sistema (Weyns, 2020). Estas mudanças podem ser causadas por diversos fatores, incluindo falhas de *hardware*, variações na demanda do sistema, incidentes de segurança ou mudanças nas condições operacionais. O objetivo principal é manter ou melhorar determinadas características de qualidade, como desempenho, segurança e confiabilidade, sem intervenção humana (De Lemos et al., 2013).

Para isso, as características dos sistemas auto adaptáveis envolvem a utilização de sensores ou agentes no sistema e no entorno, detectando as condições de adaptação necessárias. Em seguida, um mecanismo de tomada de decisão determina se uma adaptação é necessária e que tipo de adaptação deve ser feita. Consequentemente, as modificações são então realizadas em tempo de execução pelo sistema auto adaptativo. A Figura 3 apresenta um modelo conceitual de um sistema auto adaptativo especificado em (Weyns, 2020), que é composto de duas partes distintas: a primeira parte interage com o ambiente e é responsável pelos interesses do domínio, ou seja, pelos interesses dos usuários para os quais o sistema é construído; a segunda parte consiste em um ciclo de *feedback* que interage com a primeira parte (e monitora seu ambiente), e é responsável pela adaptação (Weyns, 2020).

Os sistemas de manutenção preventiva e preditiva podem ser otimizados a partir de sistemas auto adaptativos, pois pode fornecer estratégias para melhoria das decisões relacionadas a esses tipos de sistemas complexos. A IoT associada a máquinas e equipamentos industriais oferece um cenário favorável à captura e processamento de dados que podem enriquecer decisões e realizar auto-adaptações nos sistemas produtivos. Assim, é possível descobrir informações estratégicas por meio de técnicas inteligentes que podem facilitar análises e tomadas de decisões relacionadas à manutenção. O conhecimento adquirido através da análise de falhas, por exemplo, permite a criação de soluções de manutenção capazes de mitigar falhas de baixa criticidade que podem ter relação com a ocorrência de falhas de maior impacto ou criticidade, podendo, por exemplo, permitir o desligamento automático de equipamentos ou realizar adaptações na velocidade do sistema produtivo para evitar o desgaste de certos equipamentos.

Assim, a manutenção preditiva utilizando técnicas de auto-adaptação fornece um método para tornar os processos de manutenção mais eficientes, pois as máquinas equipadas com sensores e unidades de processamento de dados podem enviar sinais de status durante o uso (Aivaliotis et al., 2019) e a partir do processamento destes dados, a auto-adaptação

Figura 3 - Modelo conceito de um sistema auto adaptativo



Fonte: (Weyns, 2020).

pode ocorrer. No entanto, é importante que os dados sejam capturados e processados sem a necessidade de interrupções e/ou atrasos na operação dos equipamentos. Neste sentido, o uso de DTs se mostra promissor.

Neste trabalho abordamos o uso de sensores e análise de dados para manutenção de máquinas aliado ao uso de sistemas auto adaptativos para compor uma suíte de serviços de apoio à especificação de DTs capazes de auxiliar nos processos de manutenção preditiva. O componente de auto adaptação pode atuar após a análise de dados e predição de falhas para melhorar o funcionamento de um equipamento ou até mesmo exercer uma intervenção em seu funcionamento. Desta forma, o planejamento das intervenções de manutenção

pode ser mais assertivo e rápido, resultando em menos equipamentos fora de uso por falhas detectadas de forma preditiva.

## 2.6 DIGITAL TWINS

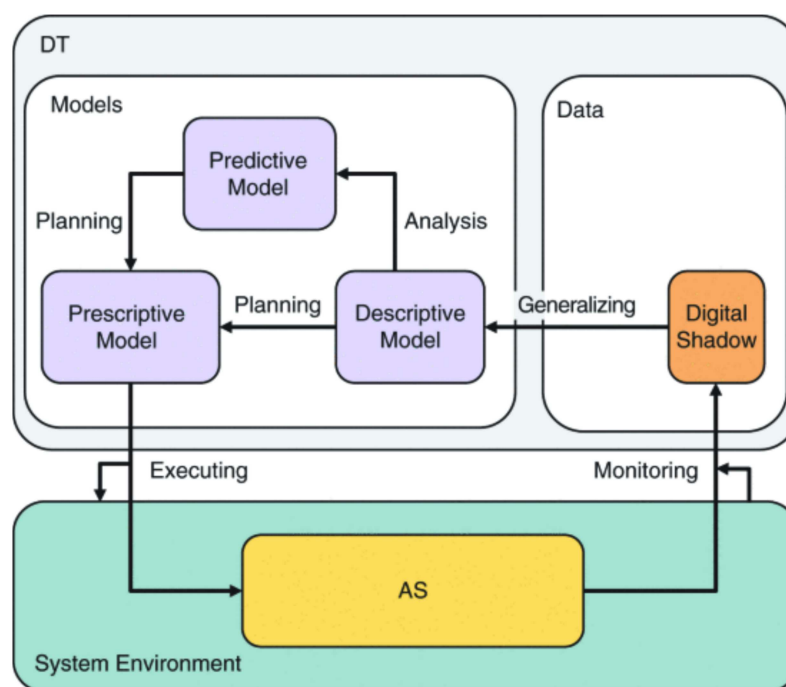
*Digital Twins* ou Gêmeos Digitais são aplicações de *software* usadas para monitorar, compreender e otimizar o comportamento de outro sistema (Eramo et al., 2021) durante seu tempo de execução. Eles surgiram como uma nova abordagem para auxiliar no enfrentamento de alguns dos problemas da Indústria 4.0 (Qi and Tao, 2018) e podem ser usados nas diversas áreas que requerem sistemas computacionais complexos, como a indústria siderúrgica com processos de fabricação de aços, a indústria aeroespacial e os processos de manutenção de aeronaves que são críticos e de alto custo, no comércio eletrônico com sistemas de recomendação cada vez mais próximos dos gostos e interesses dos usuários, entre outras aplicações potenciais como cuidados médicos nas atividades de ensaios clínicos prevendo o comportamento de um paciente ao utilizar determinado medicamento (Hinduja et al., 2020). O grande diferencial do uso de um DT é que o sistema complexo em execução não é afetado. Uma cópia do mesmo é feita pelo DT, que analisa os dados e propõem intervenções no sistema em operação quando necessário.

Especificamente, o DT é uma representação virtual (ou réplica) de um sistema real ou sistema atual (do inglês *Actual System - AS* - Figura 4) que é continuamente atualizado com dados em tempo real ao longo de seu ciclo de vida e, ao mesmo tempo, pode interagir e influenciar o AS. Uma característica proeminente de um DT é que ele deve estar conectado aos dados obtidos do AS e de seu ambiente. Estes dados são processados através de modelos e estes modelos e dados permitem a implementação de serviços relacionados ao AS, entre estes, serviços relacionados a análise inteligente dos dados. Esta análise inteligente desses dados permite que o DT preveja eventos. Para fazer estas análises, a utilização de algoritmos de IA é uma das abordagens utilizadas. A IA oferece tecnologias para desenvolver DTs pois apresenta soluções que envolvem a predição de eventos, dando suporte à tomada de decisão.

A Figura 4 apresenta um modelo conceitual de DTs. O AS refere-se ao sistema real conectado ao DT para fornecer dados e o DT por meio de modelos realiza ações/operações no AS. O componente *Data* está relacionado à representação e armazenamento de dados, sejam eles históricos ou atuais para utilização pelo DT. A partir dos dados são criados modelos digitais específicos para o funcionamento do DT. Também o modelo apresenta o elemento *Digital Shadow* que representa um conjunto variado de dados que são estruturados e organizados, que apresentam o estado do AS e compõem sua representação digital (Eramo et al., 2021). Já o componente *Models* possui três papéis que podem ser representados por modelos em um DT: descrição, predição e prescrição.

O modelo descritivo apresenta o AS, apenas refletindo-o. O modelo preditivo é

Figura 4 - Modelo conceitual de um Digital Twin



Fonte: (Eramo et al., 2021).

utilizado para prever dados que não foram obtidos do AS que serão usados para processos de tomadas de decisão. Por exemplo, a partir de dados do DT já processados é possível, a partir de modelos inteligentes, prever a ocorrência de uma falha considerando dados históricos do modelo descrito e previsões do modelo. Já o modelo prescritivo orienta o processo evolutivo do sistema em tempo de execução em sistemas auto adaptativos (Eramo et al., 2021).

Os DTs são um componente essencial da Indústria 4.0 e também o serão na Indústria 5.0, evoluindo a maneira como as indústrias projetam, produzem e mantêm seus produtos e processos. Em relação aos processos de manutenção preditiva, DTs desempenham um papel importante na detecção antecipada de falhas, na programação eficiente da manutenção e na redução de paradas não planejadas. Na manutenção preditiva de equipamentos, um dos principais benefícios da utilização de DTs é a capacidade de apoiar o planejamento da manutenção preditiva pois, ao coletar dados em tempo real dos equipamentos e compará-los com os modelos virtuais correspondentes, é possível identificar padrões que indicam o desgaste ou o mau funcionamento eminente. Isso permite que as equipes de manutenção intervenham antes que ocorra uma falha, reduzindo assim o tempo de inatividade e os custos associados.

No Capítulo 4 detalhamos uma suíte de serviços proposta nesta dissertação para apoiar a manutenção preditiva de equipamentos, com vistas à Indústria 5.0. A suíte de serviços é baseada na especificação de DTs, a partir de arquiteturas auto adaptativas, IA e ontologias.

## 2.7 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Apresentamos neste capítulo uma revisão das principais tecnologias relacionadas à proposta desta dissertação. Foi realizada uma revisão bibliográfica sobre sensores, sistemas inteligentes, arquiteturas auto-adaptativas, ontologias e *Digital Twins*, com o objetivo de auxiliar no desenvolvimento da abordagem proposta. No capítulo seguinte, um MSL é apresentado, com o objetivo de analisar as principais soluções propostas na literatura relacionada a abordagem apresentada nesta dissertação.

### 3 MAPEAMENTO SISTEMÁTICO DA LITERATURA

A análise de dados oriundos de dispositivos IoT e sensores apresenta alguns desafios, considerando o gerenciamento e processamento do grande volume de dados gerados, de forma a possibilitar análises que apoiem a tomada de decisão (Neves et al., 2016) e a complexidade e variedade dos dados na Indústria 4.0 (Hashem et al., 2015).

Considerando especificamente a manutenção preditiva na indústria 4.0 ou 5.0, a análise de dados em tempo real desempenha um papel importante, produzindo informações relevantes para a tomada de decisão relacionada à manutenção de equipamentos. A partir da utilização de técnicas como aprendizado de máquina, mineração de dados e processamento semântico, é possível detectar padrões, tendências e anomalias em dados coletados de equipamentos e sistemas de produção, produzindo *insights* e conhecimento para apoiar processos de decisão (Neves et al., 2016).

Para verificar como a literatura discute o assunto, conduzimos uma revisão sistemática considerando artigos científicos que discutem análise de dados de dispositivos IoT, processados por técnicas inteligentes, no contexto da manutenção preditiva. Uma revisão sistemática baseada em evidências é um tipo de revisão que utiliza uma abordagem sistemática e transparente para identificar, avaliar e sintetizar as melhores evidências sobre um tópico específico. É um estudo abrangente e rigoroso de todas as evidências disponíveis sobre um determinado tópico, que é conduzido de acordo com critérios pré-especificados (Kitchenham and Brereton, 2013). Nesta dissertação, a abordagem híbrida proposta por (Mourão et al., 2020) será utilizada, que abrange as seguintes etapas: planejamento, condução e resultados.

#### 3.1 PLANEJAMENTO

Para este trabalho adotamos a estratégia híbrida que combina buscas em bibliotecas digitais através *snowballing backward* (SB) e *snowballing forward* (SF). A utilização combinada das técnicas SB e SF é justificada pela capacidade de proporcionar uma revisão sistemática da literatura mais abrangente e robusta. Enquanto o SB busca referências citadas no trabalho inicial, o SF procura por trabalhos que citaram a referência inicial, permitindo assim identificar estudos relevantes que emergiram após a publicação do artigo original. Essa abordagem combinada pode revelar uma evolução teórica e metodológica no campo de estudo (Mourão et al., 2020).

Para isso, as buscas foram realizadas na biblioteca digital *Scopus* que, de acordo com a análise comparativa com outras bibliotecas digitais feita por (Mourão et al., 2020), apresentou resultados mais adequados quando o protocolo utilizado envolve a combinação da busca com SB e SF, em comparação com outras bibliotecas digitais. Dessa forma, utilizamos uma estratégia híbrida que combina a busca na base de dados e o processo de



*snowballing* (Scopus + BS||FS).

Escolhida a base a ser utilizada, as etapas seguintes do planejamento foram executadas. Assim, foram definidos os objetivos e a seleção do protocolo a ser utilizado durante o mapeamento, que tem como objetivo responder às questões apresentadas na Tabela 1.

Tabela 1 – Questões do mapeamento sistemático

Questões do Mapeamento	Objetivos
MQ1: Quais soluções de <i>software</i> propõem suporte à decisão na manutenção na indústria 4.0 ou 5.0?	Esta questão tem como objetivo identificar as arquiteturas utilizadas nas soluções envolvendo manutenção preditiva na indústria 4.0 ou 5.0 que promovem resultados no contexto de manutenção, por exemplo, arquiteturas auto adaptativas.
MQ2: Quais foram as técnicas inteligentes utilizadas?	Esta questão tem como objetivo identificar quais foram as abordagens técnicas nas soluções, quais recursos tecnológicos foram aplicados na elaboração das soluções.
MQ3: Quais foram as técnicas utilizadas para avaliação das soluções?	Esta questão tem como objetivo identificar como as soluções propostas foram avaliadas, quais critérios foram considerados pelos autores na avaliação.
MQ4: Como as soluções realizam análise de dados ?	Esta questão tem como objetivo identificar como as soluções fazem a análise de dados e para suporte à decisão.

Fonte: Elaborada pelo autor (2023).

Com base nas questões de pesquisa, utilizamos o método PICOC proposto por (Petticrew and Roberts, 2008) para definir o escopo da busca. Os seguintes termos relacionados foram utilizados:

- População (P): soluções que envolvam análise de dados para apoio à tomada de decisão;
- Intervenção (I): soluções que utilizam técnicas inteligentes sobre dados de dispositivos IoT e a auto adaptação;
- Comparação (C): comparação entre as técnicas as tecnologias inteligentes selecionadas para apoio nos processos de manutenção preditiva e preventiva;
- Resultados (R): maior agilidade e assertividade nas tomadas de decisão e intervenções em processos de manutenção preditiva e preventiva;

- Contexto (C): dispositivos IoT na indústria 4.0 ou 5.0;

Para criar a *string* de busca foram selecionados os principais termos e seus sinônimos relacionados às questões de pesquisa definidas. Com base nos termos, a *string* de busca foi formada aplicando operadores booleanos E/OU. Termos semelhantes foram associados através do operador OR e em seguida os subconjuntos foram integrados pelos operadores booleanos AND. Ao final, a *string* de busca ficou assim especificada:

String de busca
"iot data"AND ( "industry"OR "industry 4.0"OR "industry 5.0") AND ( "machine learning"OR "deep learning"OR "artificial intelligence"OR “intelligent techniques” ) AND "predictive maintenance"AND ( "decision support"OR "self adaptive systems")

Para inclusão e exclusão de trabalhos foram adotados 06 critérios de inclusão e 13 critérios de exclusão, apresentados na Tabela 2:

Tabela 2 – Questões do mapeamento sistemático

Critérios de inclusão	Critérios de exclusão
A solução tecnológica está dentro do contexto da indústria 4.0 ou 5.0 OU envolve dispositivos IoT (E)	Corresponde à palavra-chave na string de pesquisa, mas o contexto é diferente dos objetivos da pesquisa.
A solução tecnológica faz uso de técnicas inteligentes (E)	O resumo não abordou nenhum aspecto das questões de pesquisa.
A solução tecnológica trata de análise de dados para tomada de decisão (E)	O artigo não contém resumo.
O trabalho está escrito em inglês (E)	Não é um estudo primário
O trabalho é um artigo completo (E)	Não disponível para as credenciais universitárias (UFJF)
O trabalho foi publicado entre 2018 a 2023	O estudo foi publicado como um artigo curto
O trabalho está disponível como artigo completo em bibliotecas digitais	O estudo não foi publicado em congresso ou periódico relacionado à Ciência da Computação
	O estudo não foi publicado em veículo de revisão por pares
	O estudo é uma revisão bibliográfica

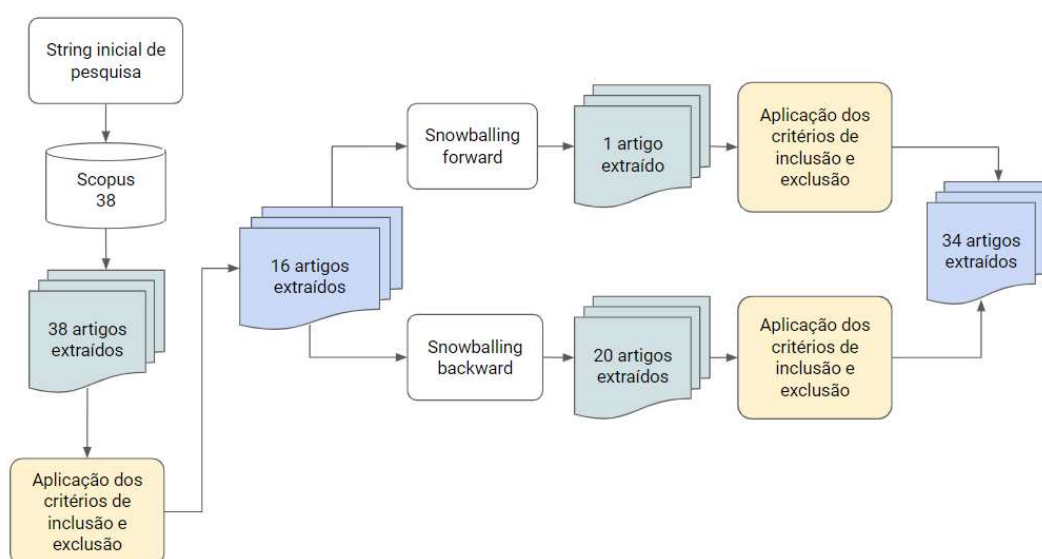
Fonte: Elaborada pelo autor (2023).

### 3.2 CONDUÇÃO

Definidas as questões de pesquisa e os critérios de inclusão e exclusão, os passos seguintes foram: (i) definir as fontes para executar a pesquisa e (ii) definir a *string* de busca que será executada nas fontes de pesquisa selecionadas. Nesta dissertação, como ressaltado anteriormente, a fonte de pesquisa escolhida para a busca foi a *Scopus* ([www.scopus.com](http://www.scopus.com)).

Após a execução da *string* de busca foram retornados 38 artigos. A Figura 5 detalha como o processo de filtragem foi executado.

Figura 5 - Processo de filtragem de artigos após consulta à base *Scopus*

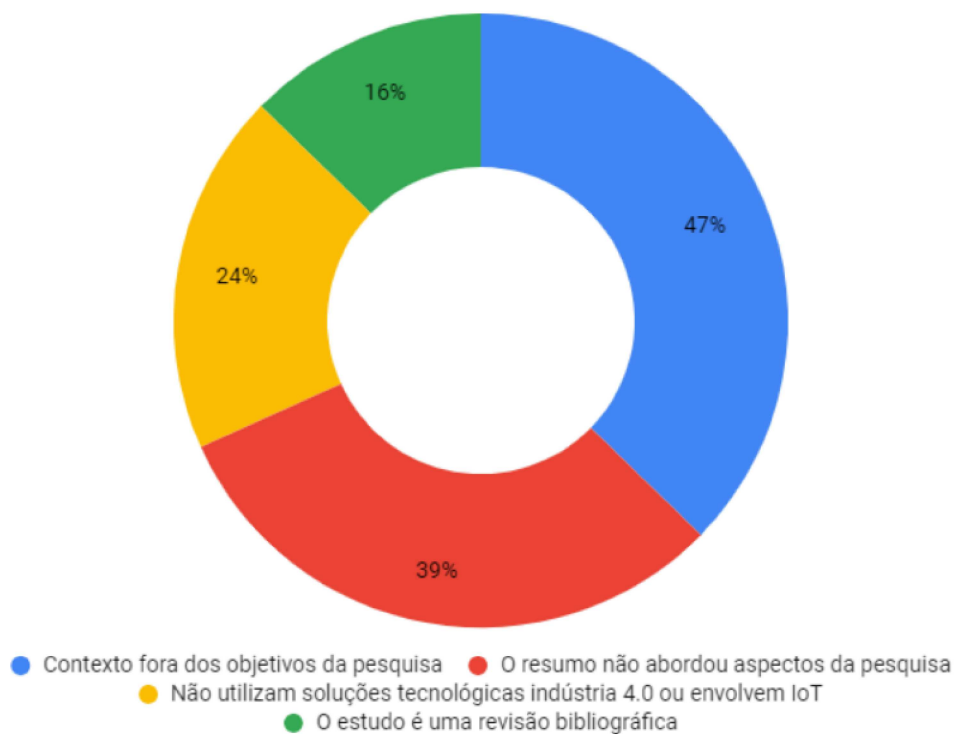


Fonte: Elaborada pelo autor (2023).

A Figura 6 apresenta o resultado da aplicação dos critérios de exclusão considerando a análise do título e resumo de cada trabalho resultante da busca (EC1, EC2, EC12 e EC13). Os demais critérios de exclusão são objetivos e foram aplicados na própria busca da *Scopus*, através de campos de filtragem. Após a aplicação dos critérios de exclusão o conjunto de trabalhos foi reduzido de 38 para 16 artigos.

De acordo com o fluxo de execução do protocolo de pesquisa apresentado na Figura 6 e após a aplicação dos critérios de inclusão e exclusão, foi realizado o SB e SF para o conjunto resultante de artigos (16 artigos). Após o *snowballing* 21 novos artigos foram encontrados e analisados a partir de seus títulos e resumos. Durante a análise destes artigos, 3 deles estavam inacessíveis a partir da base *Scopus*, o que reduziu o conjunto final para 34 artigos. A Tabela 3 mostra a lista com o título de cada um dos trabalhos selecionados.

Figura 6 - Aplicação dos critérios de exclusão.



Fonte: Elaborada pelo autor (2023).

Tabela 3 – Trabalhos selecionados

#	Título
S1	A machine learning modeling framework for predictive maintenance based on equipment load cycle: an application in a real world case (Vallim Filho et al., 2022)
S2	Proposed industry 4.0 maintenance framework for critical and demanding infrastructures and processes. (Karaiskos et al., 2022)
S3	A blockchain-enabled deep residual architecture for accountable, in-situ quality control in industry 4.0 with minimal latency. (Leontaris et al., 2023)
S4	Human knowledge centered maintenance decision support in digital twin environment. (Naqvi et al., 2022)
S5	Enterprise integration and interoperability for big data-driven processes in the frame of industry 4.0. (Bousdekis et al., 2021)

Continua na próxima página

**Tabela 3 – continuação**

#	Título
S6	Integrating preventive and predictive maintenance policies with system dynamics: a decision table approach. (Yıldız and Soylu, 2023)
S7	An artificial intelligence approach for improving maintenance to supervise machine failures and support their repair. (Rojek et al., 2023)
S8	Real-time context-aware microservice architecture for predictive analytics and smart decision-making. (Ortiz et al., 2019)
S9	Deep learning stranded neural network model for the detection of sensory triggered events. (Kontogiannis et al., 2023)
S10	Comparison of different features and neural networks for predicting industrial paper press condition. (Rodrigues et al., 2022)
S11	From knowledge-based to big data analytic model: a novel iot and machine learning based decision support system for predictive maintenance in industry 4.0. (Rosati et al., 2023)
S12	Using supervised and one-class automated machine learning for predictive maintenance. (Ferreira et al., 2022)
S13	A data-driven predictive maintenance strategy based on accurate failure prognostics. (Chen et al., 2021)
S14	A framework for big data analytical process and mapping—baprom: Description of an application in an industrial environment. (de Carvalho Chrysostomo et al., 2020)
S15	A model for predictive maintenance based on asset administration shell. (Cavalieri and Salafia, 2020)
S16	A novel decision support system for managing predictive maintenance strategies based on machine learning approaches. (Arena et al., 2022)
S17	Building predictive maintenance framework for smart environment application systems. (Katona and Panfilov, 2018)
S18	CBR-Based Decision Support System for Maintenance Text Using NLP for an Aviation Case Study. (Naqvi et al., 2022)

Continua na próxima página

**Tabela 3 – continuação**

#	Título
S19	Cloud-based analytics module for predictive maintenance of the textile manufacturing process. (Chang et al., 2021)
S20	Enabling of Predictive Maintenance in the Brownfield through Low-Cost Sensors, an IIoT-Architecture and Machine Learning. (Strauss et al., 2018)
S21	IDARTS – Towards intelligent data analysis and real-time supervision for industry 4.0. (Peres et al., 2018)
S22	Machine Learning approach for Predictive Maintenance in Industry 4.0. (Paolanti et al., 2018)
S23	Machine learning for predictive maintenance of industrial machines using IoT sensor data (Kanawaday and Sane, 2017)
S24	Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture (Cachada et al., 2018)
S25	Predictive Maintenance and Fault Monitoring Enabled by Machine Learning: Experimental Analysis of a TA-48 Multistage Centrifugal Plant Compressor (Achouch et al., 2023)
S26	Predictive maintenance architecture development for nuclear infrastructure using machine learning. (Gohel et al., 2020)
S27	Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time (Ayvaz and Alpay, 2021)
S28	SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4.0 (Calabrese et al., 2020)
S29	An Artificial Intelligence Approach for Improving Maintenance to Supervise Machine Failures and Support Their Repair (Rojek et al., 2023)
S30	IoT-based data-driven predictive maintenance relying on fuzzy system and artificial neural networks (Aboshosha et al., 2023)
S31	Predictive Maintenance in Industry 4.0 for the SMEs: A Decision Support System Case Study Using Open-Source Software (Pejić Bach et al., 2023)

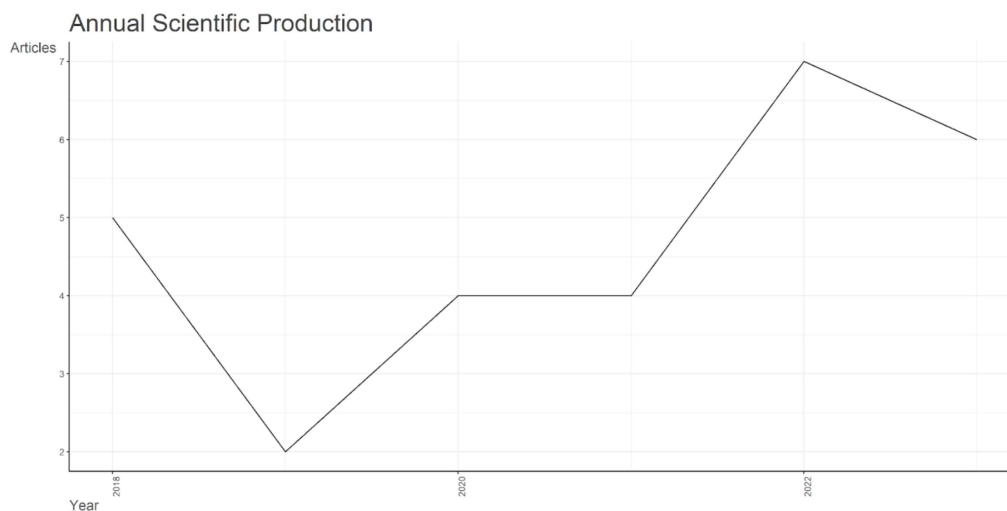
Continua na próxima página

**Tabela 3 – continuação**

#	Título
S32	Digital twin-driven decision support system for opportunistic preventive maintenance scheduling in manufacturing (Neto et al., 2021)
S33	Predictive maintenance of baggage handling conveyors using IoT (Gupta et al., 2023)
S34	Predictive Maintenance with Sensor Data Analytics on a Raspberry Pi-Based Experimental Platform (Chuang et al., 2019)

Embora a busca não tenha sido realizada em ordem cronológica, os resultados foram obtidos considerando o intervalo de tempo entre os anos de 2018 e 2023. Analisando este período, foi possível observar uma queda na produção científica no ano de 2019 e estabilização na curva de crescimento nos anos 2020 e 2021, como mostra a Figura 7. A pandemia iniciada no final de 2020 pode ser um fator importante para esse decréscimo. De forma geral, a produção científica nessa área tem crescido a uma taxa superior a 3% por ano, a partir da amostra analisada.

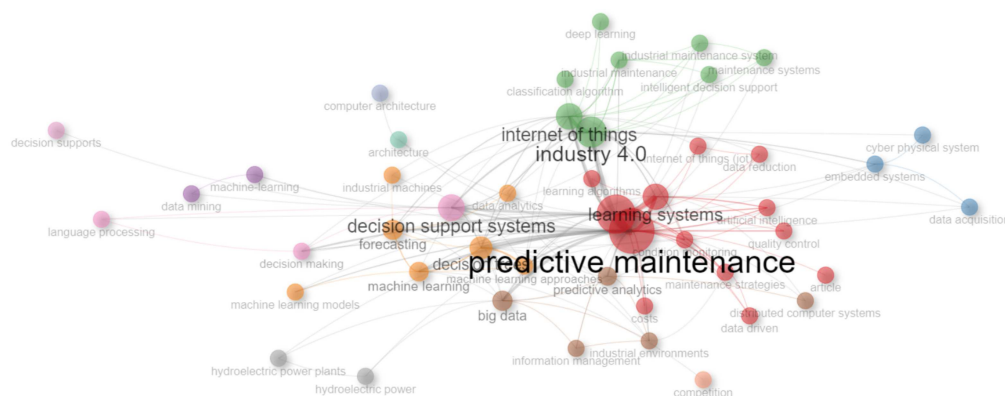
Figura 7 - Produção científica anual entre 2018 e 2023.



Fonte: R-Studio (Bibliometrix).

Considerando os autores ativos na área, a pesquisa retornou um quantitativo de 130 autores nas 34 publicações selecionadas. Cerca de 10% das publicações foram produzidas a partir de coautoria envolvendo autores de diferentes nacionalidades e a média de autores por publicação foi de 5 autores por publicação.

Figura 8 - Redes temáticas.



Fonte: R-Studio (Bibliometrix).

Considerando as áreas relacionadas, foi possível perceber que existem grupos de termos com densidades distintas, como mostra a Figura 8. O termo “Manutenção Preditiva” está relacionado fortemente com termos como “Sistemas de Aprendizagem”. Da mesma forma, o termo “Suporte a Decisão” está fortemente relacionado aos termos “Previsão”, “Análise de Dados” e “Modelos de Aprendizagem de Máquina”. Por fim, o termo “Internet das Coisas” também apresenta conexões com densidade considerável com os termos “Algoritmos de Classificação”, “Aprendizagem Profunda”, “Sistemas de Manutenção Industrial” e “Suporte de Decisão Inteligente”.

A Figura 8 destaca que as soluções para manutenção preditiva no contexto da Indústria 4.0 ou 5.0 são baseadas em aplicação de sistemas de aprendizagem que fazem uso de dados de dispositivos da IoT para produção de conhecimento necessário aos sistemas de suporte à decisão. De fato, durante a análise dos trabalhos selecionados fica clara a ocorrência dessa tríade.

A partir do conjunto de trabalhos selecionados também foi possível perceber diferentes abordagens para produção de dados de apoio à decisão. O trabalho S1 aborda a seleção de variáveis de interesse como sendo um processo crítico para a coleta de dados relevantes e úteis. Uma vez identificadas as variáveis críticas, como eficiência de produção, tempo de inatividade de equipamento, ou custo operacional, podem-se obter *insights* para a tomada de decisão. Um passo à diante está o trabalho S31 que oferece um sistema de apoio à decisão que analisa dados de sensores em tempo real informando desvios e fazendo correções de forma automática.

O uso de algoritmos avançados e técnicas de aprendizagem de máquina podem ajudar a identificar quais dessas variáveis são de fato relevantes, maximizando o potencial de análise. Da mesma forma, os trabalhos S10, S22, S23, S24, S25 e S26 exploram os recursos de aprendizagem de máquina para predição de comportamentos de máquinas e informações de sensores.



A manutenção inteligente (S4, S18, S33 e S34) e o planejamento inteligente da manutenção baseado em conhecimento (S6, S14 e S21) foram explorados a partir da análise de dados de desempenho de máquinas e equipamentos. Neste sentido, as organizações podem prever e prevenir falhas antes que elas ocorram, otimizando a eficiência do processo e reduzindo os custos de reparo e tempo de inatividade. Nesse sentido, o trabalho (S32) se destaca por propor um DT para planejamento de intervenções de manutenção.

A extração de recursos (S10, S13 e S14) para modelos de previsão e a análise de dados contextuais (S16, S21 e S22) para apoio na predição foram utilizados nos trabalhos considerando informações que não estão diretamente relacionadas ao processo controlado, mas exercem influência. O trabalho (S18) faz uso de técnicas de raciocínio baseado em casos para analisar situações passadas semelhantes para orientar as decisões presentes e futuras, o que também é feito pelo trabalho (S30) por meio de busca de padrões em dados de falhas para minimizar erros humanos na classificação de eventos de falhas operacionais.

Sobre os algoritmos de inteligência artificial utilizados nas soluções selecionadas na pesquisa, organizamos em categorias e identificamos quais trabalhos utilizam qual técnica, conforme a Tabela 4. Grande parte dos trabalhos utiliza redes neurais artificiais como base para construção das soluções inteligentes. Árvores de Decisão e Regressão Logística também aparecem de forma recorrente nos trabalhos, mas em menor escala. Considerando as categorias nas quais os algoritmos foram organizados, destacamos os modelos de aprendizado supervisionado como método adotado por grande parte dos trabalhos.

### 3.3 ANÁLISES

A busca por trabalhos relacionados à manutenção no contexto da indústria 4.0 teve por objetivo encontrar soluções que contribuam para a manutenção preditiva ou que auxiliem no planejamento da manutenção. Dessa forma, os trabalhos selecionados trazem propostas de soluções e/ou arquiteturas para desenvolvimento de soluções e a partir dessa referência buscamos destacar a forma como as soluções foram avaliadas. Identificamos que a maioria das soluções utiliza modelos preditivos baseados em aprendizagem de máquina e, por isso, a metodologia utilizada para avaliar tais modelos é direcionada para a medição de acurácia dos modelos, que é uma métrica utilizada para avaliar o desempenho de um modelo em sua capacidade de classificação correta de amostras. Além da acurácia de modelos, alguns trabalhos utilizaram análise estatística, estudos de caso com dados reais, simulação e comparação entre modelos. A Figura 9 mostra a distribuição das metodologias de avaliação.

Tabela 4 – Algoritmos de inteligência artificial utilizados

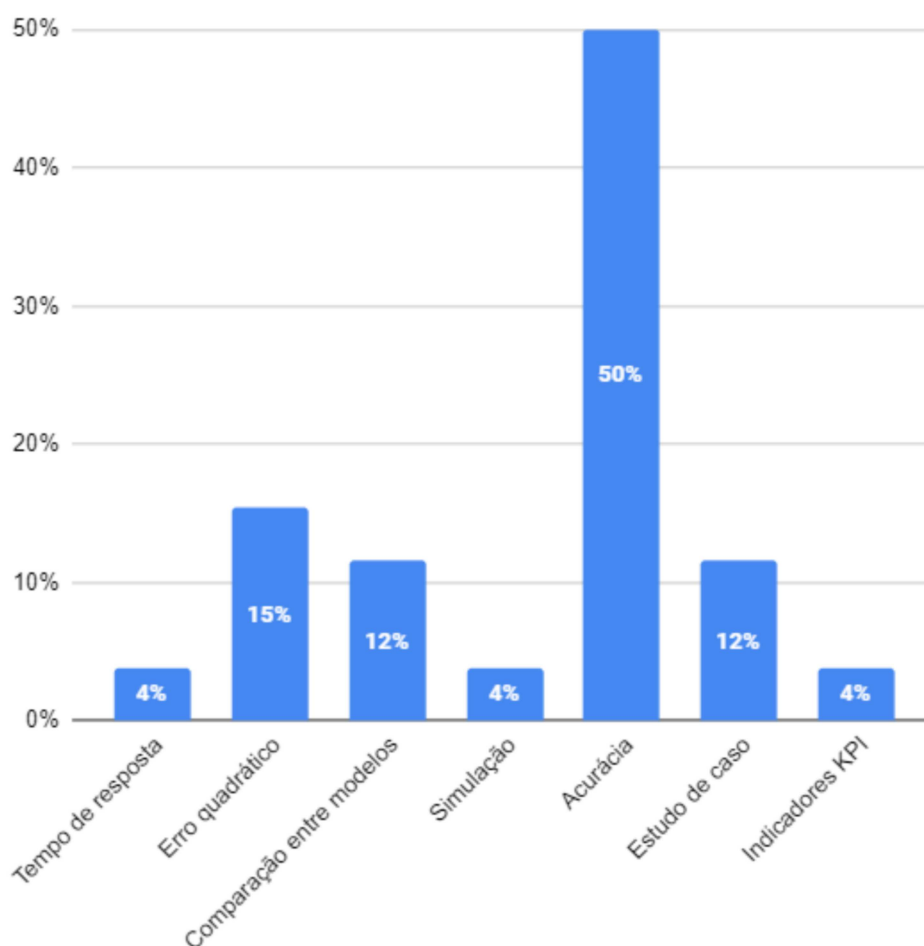
<b>Categorias</b>	<b>Algoritmos</b>	<b>Trabalhos</b>
Modelos de Aprendizado Supervisionado (75%)	Árvore de decisão(6)	S1, S6, S14, S15, S16, S34
	Rede Neural Artificial(15)	S1, S6, S7, S8, S9, S10, S13, S14, S15, S23, S25, S27, S29, S33, S34
	Regressão Logística(8)	S1, S19, S21, S25, S26, S30, S33, S34
	Floresta Aleatória(5)	S1, S11, S27, S28, S33
	Máquina de Vetor de Suporte(6)	S6, S20, S23, S26, S27, S33
	XG Boost(2)	S10, S27
	Naive Bayes(1)	S25
	Gradient Boosting(1)	S28
Modelos de Aprendizado Não Supervisionado (5%)	Classificadores K-Means(1)	S21
	Floresta de isolamento(2)	S17, S20
Modelos de Aprendizado Profundo (8%)	Redes Residuais Profundas(1)	S3
	Aprendizagem Profunda(3)	S5, S9, S30
	Redes Neurais Recorrentes(1)	S4
Modelos de Séries Temporais (2%)	Modelo ARIMA(2)	S8, S31
Ferramentas e Técnicas de Aprendizado de Máquina (5%)	Auto Machine Learning(1)	S12
	Floresta de Decisão(1)	S22
	Árvores de Classificação(1)	S23
Modelagem e Simulação (5%)	Digital Twins(3)	S2, S4, S32

Fonte: Elaborada pelo autor (2023).

A análise de dados é uma funcionalidade importante no processo de tomada de decisão e perpassa pela utilização de algoritmos. Assim, considerando o conjunto de trabalhos selecionados, identificamos técnicas utilizadas para análise de dados. Neste sentido, identificamos a utilização de algoritmos de inteligência artificial para classificar dados ou executar algum tipo de predição sobre dados como a principal técnica utilizada. De acordo com o conjunto de trabalhos obtidos, foi possível observar que as publicações

relacionadas à análise de dados estão organizadas em quatro categorias, como mostra a Tabela 5.

Figura 9 - Metodologias de avaliação das soluções.



Fonte: Elaborada pelo autor (2023).

Foi possível identificar quatro categorias distintas de abordagens, conforme a Tabela 5. A primeira categoria engloba trabalhos que se dedicam à proposição de técnicas eficientes para a aprendizagem de dados. Essas abordagens são fundamentadas na exploração de dados históricos como base para o treinamento de modelos de predição, proporcionando uma análise de padrões e tendências. A segunda categoria engloba pesquisas que se concentram em técnicas específicas de análise de dados sensíveis ao contexto. Estas propostas têm como objetivo principal realizar ajustes em tempo real em sistemas de manufatura, com base nas informações obtidas a partir dos dados contextualizados. Essa abordagem representa um avanço na otimização e eficiência dos processos industriais. A terceira categoria compreende trabalhos que se aplicam técnicas para a classificação de dados em processos de controle da qualidade. Essa abordagem é relevante para garantir a precisão e a confiabilidade dos produtos e serviços, tornando-se um elemento-chave na excelência operacional. Por fim, a quarta categoria explora a simulação através de DT

Tabela 5 – Arquiteturas para análise de dados

<b>Categorias de técnicas</b>	<b>Trabalhos relacionados</b>
Aprendizado a partir de dados históricos	S4, S6, S7, S13, S18, S24, S27 e S28, S33, S34
Ajustes em sistemas de produção a partir de dados sensíveis a contexto / processamento dados em tempo real	S8 e S21
Classificação de dados em processos de controle da qualidade	SS3 e S23
Planejamento de manutenção através de simulação	S32

Fonte: Elaborada pelo autor (2023).

para otimizar o planejamento de intervenções de manutenção a partir de dados históricos e em tempo real usando simulações. Cada uma dessas categorias de técnicas apresenta contribuições considerando o uso de aprendizado de máquina e análise de dados, ainda que com sua diversidade e foco distintos.

### 3.4 RESULTADOS

Com base nos dados extraídos e na análise detalhada dos artigos resultantes, pudemos responder as questões de pesquisa.

Respondendo à Questão 01 do mapeamento: “Quais arquiteturas propõem suporte à decisão na manutenção na indústria 4.0?” percebemos que o conjunto de trabalhos analisados apresenta diversas estratégias para a produção de dados de apoio à decisão. No entanto, a base técnica na qual soluções foram construídas é semelhante, ou seja, o aprendizado de máquina foi aplicado na identificação de variáveis mais importantes em conjuntos de dados e na previsão de comportamentos de máquinas por meio de análise de dados de falha ou em tempo real. De forma semelhante, os trabalhos concentrados no planejamento da manutenção utilizaram aprendizagem de máquina e simulação por DT para otimizar o planejamento.

Sobre os algoritmos de inteligência artificial utilizados (“Quais foram as técnicas inteligentes utilizadas?”, Questão 02 do mapeamento), a maioria dos trabalhos usa redes neurais artificiais para desenvolver soluções inteligentes. Entre as categorias de algoritmos de aprendizagem de máquina, os modelos de aprendizado supervisionado são comumente adotados. Os trabalhos relacionados apresentam um ponto fraco do ponto de vista do suporte à decisão e da auto adaptação, uma vez que apenas um modelo é considerado ou então é feita uma comparação entre vários modelos de forma desconectada da proposta. A

proposta desta dissertação também utiliza aprendizado supervisionado, mas o aprendizado de máquina não tem enfoque em apenas um modelo. Entendemos que é interessante avaliar dentro de um conjunto de modelos qual deles tem maior aderência com o conjunto de dados. Sobre esse aspecto a proposta desta dissertação apresenta um ponto forte, pois permite a seleção automática do melhor modelo preditivo, diferentemente dos trabalhos relacionados que até chegam a fazer a comparação entre modelos destacando dentre eles o melhor, mas essa informação não é gerada de forma automática e o resultado dessa análise não é aproveitado pela solução proposta.

Respondendo à Questão 03 do mapeamento (“Quais foram as técnicas utilizadas para avaliação das soluções?”), os trabalhos selecionados propõem soluções e arquiteturas para o desenvolvimento dessas soluções. A maioria emprega modelos preditivos baseados em inteligência artificial, cuja eficácia é medida pela acurácia, uma métrica que avalia a capacidade de classificação correta do modelo, evidenciando a aplicação de métodos quantitativos para avaliar as soluções. Além disso, a análise estatística, estudos de caso com dados reais, simulações e comparações entre modelos são também propostas. Aqui surge uma oportunidade de avanço na pesquisa, pois percebemos que as soluções avaliadas utilizam diferentes modelos de predição e os resultados da comparação entre estes modelos não são aproveitados na própria solução. Selecionar o modelo preditivo com melhor desempenho em um conjunto de dados pode tornar a solução técnica mais adaptável, ou seja, a solução responderá com a maior precisão possível independentemente do conjunto de dados de entrada. Por exemplo, em um cenário onde temos conjuntos de dados diferentes, sobre processos diferentes, a tarefa de escolha do modelo preditivo dependerá sempre da avaliação de um especialista e essa avaliação só será possível após a análise de desempenho entre modelos, consumindo tempo e não necessariamente garantindo uma abordagem técnica adequada para comparar modelos preditivos. A proposta desta dissertação traz contribuições nessa etapa, pois propomos uma suíte de serviços que permite a seleção automática de um modelo específico que melhor se adapta a um dado conjunto de dados.

Por fim, respondendo à Questão 04 do mapeamento (“Como as soluções realizam análise de dados?”), as soluções se dividem em três categorias distintas: arquiteturas eficientes para a aprendizagem de máquina a partir de dados históricos para treinar modelos preditivos, arquiteturas para análise de dados contextuais com objetivo de realizar ajustes em tempo real em sistemas de manufatura e arquiteturas focadas na classificação de dados em controle de qualidade. Também aqui a proposta desta dissertação apresenta contribuições, pois propomos o enriquecimento de informações através de ontologias e processamento semântico, utilizando dados de contexto, sensores e dispositivos IoT em um DT para apoiar a manutenção preditiva na Indústria 5.0.

Como resultado da análise do conjunto de trabalhos retornados, podemos destacar que muitos trabalhos consideram a manutenção preventiva como uma questão-chave nos

sistemas da Indústria 4.0. No entanto, não consideram a complexidade dos sistemas de *software* que a Indústria 4.0 exige. O trabalho S7 apresenta uma solução para implementar uma estratégia de manutenção baseada na análise histórica das falhas, com o objetivo de classificá-las quanto aos seus tipos e suas causas, como também as atividades de manutenção. A solução proposta utiliza IA para detectar as causas das falhas, apoiando ações de manutenção necessárias. A solução é apresentada como um sistema de monitoramento combinado com diagnósticos através de IA. A solução apresentada oferece ainda um modelo para predição de falhas, mas não avança na questão do apoio às decisões da manutenção preventiva, uma vez que analisa dados históricos e não oferece recursos de auto adaptação e gerenciamento, nem aponta necessidades de intervenções.

O trabalho S27 também apresenta uma solução para predição de falhas tal como o trabalho S7, mas aprofunda a seleção de um modelo de aprendizagem de máquina com maior capacidade preditiva. Dados de dispositivos IoT são registrados em bancos de dados e preparados para processamento através de modelos de aprendizagem de máquina. Saídas dos modelos de predição e alertas são mostrados aos usuários. No entanto, as informações não são enriquecidas a partir de dados de contexto ou de outra ordem, que considere informações além do que é recebido a partir dos dispositivos IoT. A solução aborda a seleção de recursos mais importantes no conjunto de dados, mas não faz a seleção de um modelo preditivo baseado nas características dos conjuntos de dados, e também essa seleção não é feita de forma automática, ou seja, a cada mudança de conjunto de dados, um novo modelo predito deve ser escolhido de acordo com a sua precisão.

O trabalho S8 traz uma solução baseada em microsserviços, favorecendo o baixo acoplamento entre as partes da solução e permitindo a escalabilidade. Além disso, chama a atenção a abordagem contextual que a solução propõe, considerando que, de forma geral, sistemas IoT são sensíveis ao contexto, oferecendo um sistema de análise preditiva sensível ao contexto e adaptável e cenários diferentes de aplicação. Como características funcionais destacam-se: a capacidade de processamento de grandes quantidades de dados heterogêneos, detecção de eventos com riscos potenciais em um determinado domínio, decisões inteligentes automáticas e comunicação com usuários através de alertas. Tais alertas são gerados considerando os usuários conectados ao sistema, de forma que as informações de alertas são geradas de acordo com o contexto de interesse onde o usuário está inserido. No entanto, a solução proposta não permite a seleção do modelo preditivo mais adequado e os dados produzidos pelo modelo preditivo são aplicados apenas para alertas aos usuários, mas não são utilizados para auto ajuste do sistema avaliado.

O trabalho S13 propõe uma estratégia de manutenção preditiva baseada em dados para tomada de decisões em manutenção para sistemas de engenharia reparáveis. A proposta é baseada em dois pilares: (i) seleção de recursos de degradação e um modelo baseado em rede neural profunda chamada rede de memória de longo e curto prazo (LSTM) para predição de desgaste de componentes, a partir da análise de padrões em

séries temporais subjacentes para então prever a vida útil restante do sistema (RUL) e (ii) de tomada de decisão relacionada ao custo de manutenção, tomando como premissa que o melhor momento para manutenção é a combinação entre o menor custo de manutenção preventiva e as probabilidades de falhas previstas. Tal como ocorre em S8 a estratégia de modelagem preditiva considera apenas o algoritmo LSTM para analisar o conjunto de dados e fazer previsões. Também não há o enriquecimento de informações sobre o estado do sistema a partir da extração de conhecimento e não há uma interface do modelo preditivo com o sistema atual para além da previsão de manutenção, como por exemplo o disparo de alertas quer seja para manutenção próxima ou intervenção emergencial em algum equipamento.

O trabalho S21 apresenta um framework para análise inteligente de dados e supervisão em tempo real (IDARTS), combinando aquisição distribuída de dados, aprendizado de máquina e raciocínio em tempo de execução contribuindo para a manutenção preditiva e controle da qualidade, reduzindo impactos de eventos disruptivos na produção. A proposta se baseia na análise de dados históricos e de tempo real por meio de um modelo de aprendizagem de máquina, produzindo dados preditivos que podem ser utilizados para auto ajuste do sistema ou emissão de alertas a operadores e produção de conhecimento. A análise em tempo real está baseada em processamento de filas de mensagens, na detecção precoce de padrões ou tendências que possam representar uma avaria ou falha de processo em chão de fábrica. Os resultados do processamento podem ser comunicados a operadores para visualização ou ainda para um componente de tomada de decisão em tempo de execução, que pode sugerir ações para retomar as condições normais de operação. A proposta é muito alinhada com a proposta desta dissertação, mas não explora o potencial que o enriquecimento de informações e auto adaptação de sistemas a partir das saídas dos modelos preditivos proporciona.

Por fim, o trabalho S32 propõe um DT para programação da manutenção preditiva oportunista. O DT representa o estoque de produtos durante o processo de fabricação, com informações obtidas através de radiofrequência (RFID) e o estado operacional das máquinas, obtido por dados de sensores. O suporte à decisão é feito por meio de simulações periódicas a partir do modelo atualizado do chão de fábrica e eventos de oportunidade que surgem em tempo real. O objetivo das simulações é prever intervalos de tempo em que as máquinas estão fora de operação, favorecendo a execução de tarefas de manutenção, com a mínima interferência no rendimento da produção. Dessa forma, a proposta contribui para a manutenção apenas no seu planejamento, sem oferecer tecnologia para produção de conhecimento a partir das simulações. Também não são aplicados de aprendizagem de máquina em variedade que permita selecionar qual modelo é mais aderente ao modelo atual do sistema de produção, aumentando assim a precisão do modelo em encontrar oportunidades de manutenção.

### 3.5 DISCUSSÃO

Como já ressaltado, uma arquitetura auto adaptativa é um sistema ou *software* com a capacidade de monitorar o ambiente e o seu próprio desempenho, para ajuste dinâmico de seus componentes, parâmetros ou estratégias para otimizar seu funcionamento em resposta a mudanças ou demandas variáveis. Considerando este conceito, os trabalhos selecionados apresentam um ou mais elementos característicos de sistemas auto adaptativos, como nos trabalhos S2,S7, S8, S13, S21 e S27 mas não encontramos uma solução que apresente todas as características que a classifique como auto adaptativa, mais especificamente que apresentasse uma abordagem onde houvesse ajuste dinâmico de componentes, parâmetros ou estratégias.

A proposta desta dissertação utiliza técnicas de aprendizagem de máquina no contexto de grandes conjuntos de dados e apoio ao processo de manutenção preditiva por meio de um DT, permitindo ainda a possibilidade de seleção do melhor modelo preditivo para um determinado conjunto de dados. Diferentemente dos trabalhos retornados no MSL, nossa proposta permite enriquecimento de informações a partir de modelos de ontologia e inferências para auxiliar na tomada de decisão em relação aos processos industriais, proporcionando ao sistema a capacidade de ser auto adaptativo, uma vez que o próprio sistema se utiliza da seleção inteligente de modelos preditivos para tornar o sistema auto adaptável com o modelo mais assertivo para o conjunto de dados.

### 3.6 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Avaliando as contribuições existentes percebemos a necessidade de combinar dados de estado de equipamentos, com processamento inteligente para enriquecimento das informações favorecendo a tomada de decisões relacionada a manutenção. Também, de acordo com o MSL e análise dos trabalhos relacionados, percebemos lacunas nos trabalhos retornados no que diz respeito à capacidade de auto adaptação de sistemas a partir da análise de dados utilizando modelos de aprendizagem de máquina. Nenhum dos trabalhos selecionados utiliza modelos semânticos combinados com algoritmos de inteligência artificial para promover a auto adaptação. Neste sentido, no capítulo 4 propomos a especificação de uma suíte de serviços para apoio a especificação de DTs de suporte a manutenção preditiva.



## 4 DT-CREATE – SUÍTE DE SERVIÇOS PARA ESPECIFICAÇÃO DE DIGITAL TWINS NA INDÚSTRIA 5.0

No capítulo anterior, apresentamos os principais trabalhos relacionados através de um MSL. Com base nos resultados, identificamos lacunas relacionadas a auto adaptação relacionada à manutenção preditiva, e na análise de dados utilizando modelos de aprendizagem de máquina. Além disso, nenhum dos trabalhos selecionados utiliza modelos semânticos combinados com algoritmos de inteligência artificial para favorecer a auto adaptação. Considerando estas lacunas, desenvolvemos a DT-Create, uma suíte de serviços com o objetivo de apoiar a análise de dados na Indústria 5.0 para auxiliar a tomada de decisões na manutenção preditiva. A DT-Create foi desenvolvida considerando a abordagem DSR. Neste capítulo, detalhamos os passos da DSR conduzidos neste trabalho, como foco nos ciclos relacionados ao desenvolvimento da suíte de serviços.

A seção 4.1 detalha a abordagem epistemológica aplicada à pesquisa e os ciclos de desenvolvimento. A seção 4.2 detalha a suíte de serviços e o DT desenvolvido no primeiro ciclo DSR. A seção 4.3 detalha a evolução da suíte de serviços e o DT desenvolvido no segundo ciclo DSR. A seção 4.4 detalha os principais componentes da suíte de serviços DT-Create desenvolvida como suíte de serviços e a seção 4.5 apresenta as considerações finais do capítulo.

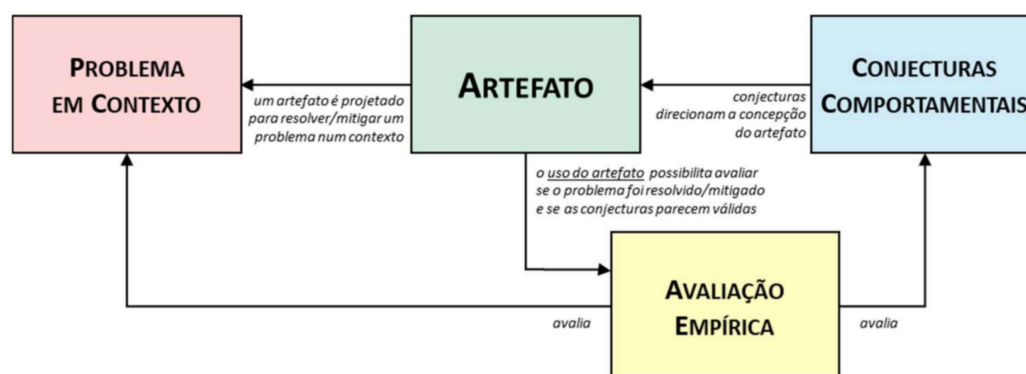
### 4.1 DSR

A abordagem DSR pode ser caracterizada como uma abordagem para construção de soluções por meio da elaboração, desenvolvimento e avaliação de artefatos. Em contraste com abordagens que se limitam à observação e análise de eventos, a DSR foca em projetar, implementar e verificar soluções que cumpram objetivos específicos (Hevner, 2007). De acordo com (Pimentel et al., 2020), existem diferentes propostas de condução da DSR, mas em geral todas têm dois objetivos claramente definidos: (i) produzir um artefato e (ii) produzir conhecimento técnico-científico.

A Figura 10 apresenta as etapas do ciclo iterativo DSR, que se inicia com a identificação e compreensão de um problema prático, que necessita de uma solução. A partir das suposições ou hipóteses, ou seja, conjecturas comportamentais, sobre como será a interação do artefato com o ambiente e como resolverá o problema, o artefato é criado como uma solução concreta para o problema. Fechando o ciclo de iteração, tem-se a avaliação empírica, que testa o artefato no ambiente para o qual foi projetado tendo como base as conjecturas comportamentais, através de estudos de caso, experimentos, ou simulações, entre outros.

Na DSR, existem ciclos bem definidos. O ciclo de relevância inicia a pesquisa com o levantamento dos requisitos para a pesquisa, ou seja, a definição dos critérios de aceitação

Figura 10 - Elementos DSR.



Fonte: (Pimentel et al., 2020).

para a avaliação final dos resultados da pesquisa. Este ciclo conecta a pesquisa à situação do mundo real, garantindo que os problemas enfrentados são relevantes e que a solução proposta é aplicável e útil (Hevner, 2007). Os resultados dessa etapa são as demandas mapeadas e os critérios de avaliação do artefato. O ciclo de rigor é aquele em que ocorre a fundamentação teórica, que norteará a construção do artefato, seja por métodos, teorias ou processos disponíveis na literatura. Neste ciclo, verifica-se se as conjecturas teóricas iniciais estão corretas com base na avaliação do artefato.

O ciclo de design (construção/avaliação) é o principal de todo projeto de pesquisa em DSR e onde o trabalho mais intensivo é realizado (Pimentel et al., 2020). Todo o processo de condução da pesquisa deve ser descrito e métodos rigorosos devem ser aplicados tanto na construção quanto na avaliação do artefato. O rigor da pesquisa decorre da utilização eficaz da base de conhecimento, que é o conjunto de fundamentos e metodologias necessárias à realização da pesquisa. A partir dos resultados obtidos nas avaliações é possível saber se novas iterações serão necessárias no projeto DSR. Novas interações podem ser necessárias porque o artefato criado pode ter necessidades de melhorias tanto nas funcionalidades quanto na qualidade, fatores que podem ser limitadores para utilização na prática. Dessa forma, em cada iteração é avaliado se o artefato atendeu aos requisitos e se resolveu o problema. Caso ele não atenda, o artefato é refinado para obter resultados mais precisos.

Assim, seguimos as seguintes etapas: definição do problema, revisão da literatura, discussão das soluções existentes, desenvolvimento do artefato, valoração e discussão dos resultados. Na primeira etapa, identificamos a Questão de Pesquisa: “Como apoiar a tomada de decisão relacionada à manutenção preditiva na Indústria 5.0, com uso de DTs?”. Como etapa seguinte investigamos trabalhos relacionados para encontrar propostas na literatura que tratam de apoio à decisão na Indústria 4.0 e 5.0. Para isso, conduzimos o MSL, já relatado no Capítulo 3.

Com base nos achados do MSL, observamos uma lacuna nos estudos que consideram

o apoio à manutenção preditiva no contexto da Indústria 5.0, o que revela que ainda é necessário explorar soluções capazes de lidar com a complexidade e diversidade de dados e possíveis soluções no contexto da Indústria 5.0. Acreditamos que ao agregar os resultados da aplicação de técnicas inteligentes com modelos semânticos focados na manutenção preditiva, processos de tomada de decisão em relação à manutenção podem ser melhorados, considerando tanto o apoio aos tomadores de decisão quanto questões relacionadas à sustentabilidade na Indústria 5.0. Desta forma, desenvolvemos a suíte de serviços em dois ciclos DSR. Em cada um dos ciclos, avaliações foram conduzidas através de estudos de caso, o que gerou conhecimento científico e teórico.

Portanto, utilizamos a DSR-Model (Pimentel et al., 2020) como base metodológica e desenvolvimento da suíte de serviços, DT-Create (artefato) para especificação de DTs, a partir dos requisitos funcionais e não funcionais a seguir.

## 4.2 REQUISITOS FUNCIONAIS E NÃO FUNCIONAIS

Como primeira etapa para a elaboração da DT-Create, identificamos quais requisitos funcionais e não funcionais eram prioritários para para o seu desenvolvimento.

**RF01:** A suíte de serviços deve ser capaz de processar diferentes conjuntos de dados industriais sobre manutenção ou mesmo operação.

**RF02:** A suíte de serviços deve ser capaz de integrar informações de contexto, para ajudar a enriquecer o conjunto de dados.

**RF03:** A suíte de serviços deve ser capaz de realizar análises semânticas e de aprendizagem de máquina.

**RF04:** A suíte de serviços deve permitir o armazenamento eficiente dos dados.

**RF05:** O DT (especificado a partir da suíte de serviços) deve fornecer mecanismos para visualização dos dados após seu processamento e, assim, apoiar os tomadores de decisão na interpretação das informações e na tomada de decisões.

**RF06:** A suíte de serviços deve apoiar os requisitos básicos de soluções relacionadas à Indústria 5.0, ou seja, foco no componente humano e na sustentabilidade.

**RNF01:** (Confiabilidade). A suíte de serviços deve permitir a comunicação com fontes de dados de maneira confiável.

**RNF02:** (Produto). A suíte de serviços deve respeitar os princípios da extensibilidade, para que englobe novas funcionalidades.

**RNF03:** (Produto). A suíte de serviços deve respeitar os princípios da flexibilidade. Flexibilidade é um atributo que reflete a gama de comportamentos que a solução existente pode ter configurado para abordar, por exemplo, processamento de dados de operação de máquinas, de processos de manutenção e outros dados relacionados aos processos produtos

da Indústria 5.0.

**RNF04:** (Sustentabilidade). A suíte de serviços deve considerar a sustentabilidade do *software*, através do treinamento contínuo dos modelos inteligentes (quando novos conjuntos forem apresentados), para manter os *insights* oferecidos pela mesma (por exemplo, predição de melhores soluções de algoritmos de aprendizagem de máquina para determinados conjuntos de dados) consistente e atualizado.

**RNF05:** (Usabilidade). A suíte de serviços e os DTs desenvolvidos devem ser fáceis de usar.

Conforme já dito, a DSR opera em ciclos e por meio deles artefatos e processos são desenvolvidos, avaliados e melhorados, de forma evolutiva. A cada ciclo, novos requisitos, problemas e oportunidades de melhoria podem surgir, criando, portanto, a oportunidade da execução de mais um ciclo. Assim, com base nos requisitos funcionais e não funcionais definidos, foi executado o primeiro ciclo de DSR, detalhado na próxima seção.

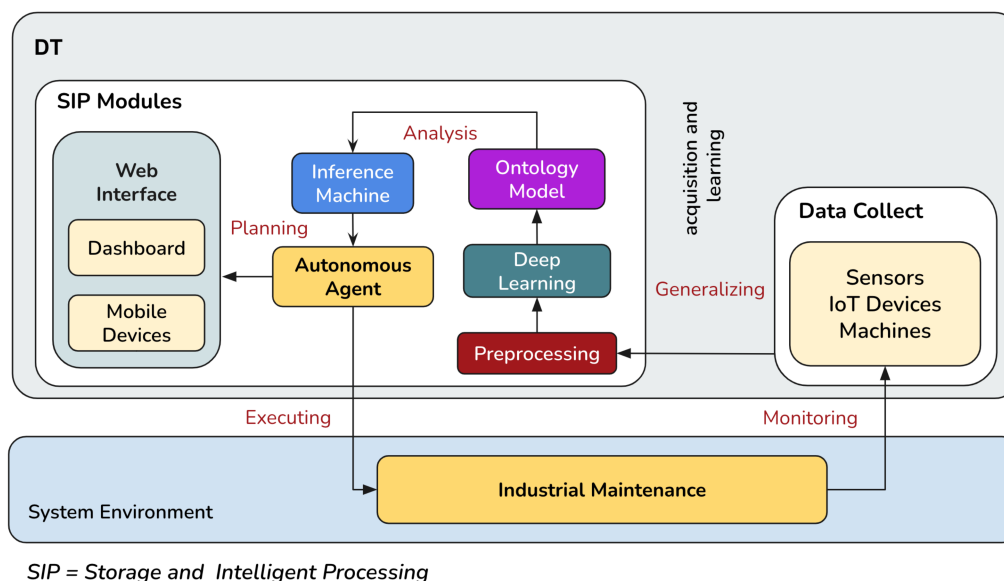
### 4.3 PRIMEIRO CICLO DA DSR

A partir dos resultados do MSL e dos Requisitos Funcionais e Não Funcionais definidos foi possível estabelecer as conjecturas para a primeira versão da plataforma. Assim, foi desenvolvida a primeira versão da DT-Create, baseada em módulos e organizada em módulos. O objetivo da suíte de serviços é auxiliar na análise de dados de processos industriais relacionados à manutenção preditiva, conforme detalha a Figura 11. A suíte de serviços é baseada em arquiteturas para o desenvolvimento de DTs (Eramo et al., 2021) considerando a divisão em Sistema Atual (*Actual System*) e Sistema de Gerenciamento, ou DT (*Managing System*).

A DT-Create foi especificada para auxiliar na análise de informações sobre eventos relacionados à necessidade de manutenção em equipamentos, a partir de dados de sensores, tanto de máquinas quanto do próprio ambiente operacional. Por meio da rede de comunicação de dados de máquinas e dispositivos IoT deve ser possível capturar, tratar e armazenar as informações em bancos de dados. Esses dados iniciais são enriquecidos com informações de especialistas a partir de análises manuais sobre quais dados são relacionados ou não a eventos críticos. A análise dos dados é realizada a partir do aprendizado de máquina supervisionado por meio de redes neurais. Um modelo preditivo é criado para classificação de eventos como críticos ou não. Assim, com a chegada de novos dados, o modelo é capaz de prever a classificação do evento, como crítico ou não.

De forma combinada, uma estrutura de processamento semântico, a partir de ontologias, foi também proposta para criar alertas a partir de regras lógicas relacionadas a ontologia. Neste sentido, a ontologia oferece uma estrutura de descoberta de conhecimento, a partir das informações existentes, permitindo obter *insights* além da capacidade preditiva

Figura 11 - DT-Create no primeiro ciclo DSR.



Fonte: Elaborada pelo autor (2023).

do modelo. O suporte à decisão é possível a partir do processamento de dados do eventos por um modelo preditivo baseado em redes neurais profundas e processamento ontológico. DT-Create permite a incorporação de alertas que podem ser exibidos a operadores em campo ou até mesmo sinalização de parada de equipamentos, por meio de mensagens enviadas diretamente aos equipamentos.

Portanto, este primeiro ciclo permitiu a construção de um artefato com funcionalidades para análise de dados combinadas com processamento ontológico para fornecer informações de suporte à decisão. A partir desta primeira versão da DT-Create, foi possível a especificação de um DT para o suporte a manutenção preditiva na indústria têxtil (da Silva et al, 2021). Neste primeiro ciclo, uma avaliação foi conduzida e foram identificados pontos de melhoria na DT-Create e DT derivado. Especificamente, identificamos as seguintes limitações: (1) De acordo com a natureza do conjunto de dados (*dataset*) analisado, a utilização de um único tipo de aprendizagem de máquina poderia tornar o processo menos efetivo, uma vez que não temos um modelo de aprendizagem que seja eficiente em sua capacidade preditiva em qualquer contexto, ou seja, com qualquer *dataset*. (2) A auto adaptação se mostrou como uma possibilidade de melhoria, uma vez que a partir dos dados de entrada, é possível selecionar a técnica de processamento inteligente mais adequada. Além disso, com a evolução dos equipamentos e soluções na Indústria 5.0, é possível realizar adaptações também no sistema produtivo, tais como desligamento de equipamentos, redirecionamento de produção para outro sistema produtivo, entre outros. Desta forma, identificou-se a oportunidade de evoluir a DT-Create para adotar novas características de auto adaptação e possibilidade de seleção do processamento inteligente de dados.

A partir da identificação destes pontos de melhoria, novos requisitos funcionais foram identificados e um novo ciclo de DSR foi conduzido (Seção 4.4).

**RF07:** A suíte de serviços deve utilizar diferentes técnicas de processamento inteligente, de acordo com os dados de entrada.

**RF08:** A suíte de serviços deve seguir novas funcionalidades relacionadas à auto adaptação.

#### 4.4 SEGUNDO CICLO DA DSR

Considerando os novos requisitos funcionais definidos RF07 e RF08, foi conduzido um segundo ciclo DSR com o objetivo evoluir o componente de processamento inteligente, permitindo a seleção de um modelo preditivo dentre um conjunto de modelos disponíveis que fosse mais assertivo considerando um dado conjunto de dados de entrada. Além disso, a suíte de serviços foi reprojetaada, reprojetaada, com o objetivo de incorporar novas características de auto adaptação. Assim, a DT-Create permite diferentes tipos de processamento inteligente, adaptável a uma maior variedade de cenários onde é necessário enriquecer informações para tomada de decisão na manutenção preditiva.

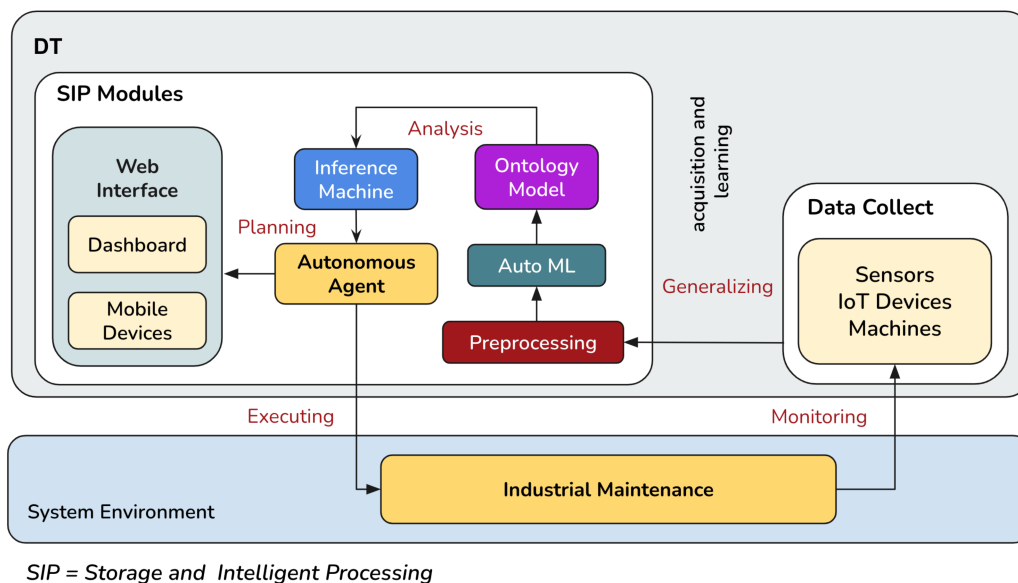
O componente *Deep Learning*, definido no primeiro ciclo DSR (Figura 11) foi o componente que sofreu maior mudança, sendo renomeado como Auto ML, de acordo com a Figura 12, refletindo as mudanças relacionadas a seleção de modelos de aprendizagem. No primeiro ciclo DSR a DT-Create utilizava redes neurais como modelo de aprendizagem e no segundo ciclo DSR foi substituído por um processo de treinamento e avaliação de um conjunto de modelos preditivos, com a seleção automática do melhor modelo disponível para utilização com o conjunto de dados de entrada.

O novo componente Auto ML executa, a partir da entrada de um *dataset*, um conjunto de etapas. A primeira etapa é a seleção dos atributos mais relevantes, que consiste em selecionar recursos no *dataset* que mais contribuem na previsão da variável alvo <sup>1</sup>, com ganho na redução de risco de ajuste excessivo (*over-fitting*) no modelo, melhorando sua precisão e diminuindo o tempo de treinamento. A segunda etapa, o treinamento de vários modelos, tem como objetivo encontrar o modelo mais adequado ao *dataset*. Nesta etapa, uma função treina e avalia o desempenho de todos os modelos dentro de um conjunto de modelos disponíveis, utilizando validação cruzada. A partir das pontuações médias de validação cruzada, é possível selecionar então o modelo mais adequado, que passará por ajustes de hiperparâmetros (*tuning*) para aumentar sua capacidade preditiva. Após a seleção do modelo e do processo de *tuning*, é realizada a etapa final deste componente, onde o modelo preditivo criado passa a ser o modelo de referência para avaliação de novos dados.

---

<sup>1</sup> Variável alvo ou variável dependente é o que um modelo de ML tenta prever ou classificar com base nas demais variáveis em um *dataset*.

Figura 12 - DT-Create no segundo ciclo DSR.



Fonte: Elaborada pelo autor (2023).

As seções a seguir detalham os principais componentes da DT-Create especificada após o segundo ciclo DSR e como os requisitos estabelecidos tanto no primeiro quanto no segundo ciclo DSR foram atendidos.

#### 4.5 DT-CREATE: PRINCIPAIS COMPONENTES

Conforme dito, no primeiro ciclo DSR foi especificada a primeira versão da DT-Create, capaz de gerar alertas a partir de processamento inteligente, tendo como base dados processados por um modelo preditivo específico baseado em redes neurais. No segundo ciclo, a DT-Create evoluiu de forma a considerar vários modelos de aprendizagem de máquina e selecionar de forma automática o melhor modelo de acordo com o *dataset* disponível, adquirindo desta forma, características de auto adaptação.

Assim, a suíte de serviços DT-Create, especificada no segundo ciclo de DSR consiste em uma suíte de serviços para especificação de DTs, capaz de aprender com dados processados através de modelos preditivos e raciocínio ontológico e ainda promover mudanças no sistema produtivo (*Actual System*) através de alertas ou até mesmo envio de mensagens a equipamentos de campo. A Figura 12 apresenta um modelo conceitual da DT-Create e seus principais componentes:

**System Environment:** Representa o espaço físico onde os dispositivos de produção estão instalados e, por meio de uma infraestrutura de redes de comunicação de dados, transmite dados de telemetria. Máquinas mais antigas geralmente não têm suporte à conectividade em redes industriais, mas podem ser conectadas através de dispositivos IoT. Uma vez conectadas, podem transmitir dados de telemetria, condição operacional atual,

status de falha, entre outras, por meio de soquetes de comunicação existentes. Assim, o monitoramento do ambiente consiste na captura de dados de telemetria de máquinas e dados de sensores em tempo real, direcionando-os até o componente *Data Collect*.

**Data Collect:** Oferece a integração dos dados vindos de diversos dispositivos através de protocolo de comunicação MQTT, permitindo o envio de informações de estado a um dispositivo central quando consultados. A DT-Create considera que os dispositivos de campo se conectam a um nó central chamado *Broker*, que adota o modelo de comunicação orientado à publicação e assinatura de tópicos. O *Broker* atua como despachante para os dispositivos IoT e demais equipamentos conectados, armazenando os dados coletados dos dispositivos em um servidor. Dessa forma, os dados obtidos dos dispositivos em campo são enviados ao *broker* e a partir dele armazenados em bancos de dados para serem utilizados pelo componente Auto ML (Figura 12), passando pelo componente *Preprocessing*, responsável pela limpeza e formatação dos dados.

**Preprocessing:** O pré-processamento é uma etapa importante, pois os dados recebidos dos dispositivos podem apresentar erros, valores ausentes, dados duplicados ou que necessitem de ajuste, o que pode comprometer ou inviabilizar o processo de aprendizagem de máquina. Após o pré-processamento, os dados já podem ser consumidos pelo componente Auto ML.

**AutoML:** O componente Auto ML é responsável pelo processamento dos algoritmos de aprendizagem de máquina e descoberta de conhecimento. Este componente treina e avalia vários modelos de aprendizagem de máquina e seleciona o modelo mais adequado para ser utilizado por aquele conjunto de dados (*dataset*). Este modelo é então utilizado para analisar novos dados oriundos dos equipamentos, dados de ambiente e demais dados históricos para estimar a probabilidade de ocorrência de determinadas classes, como, por exemplo, falhas nos equipamentos. A cada novo *dataset* de entrada, um novo modelo preditivo poderá ser selecionado e ajustado. Dessa forma é possível manter soluções geradas pela suíte de serviços para cada DT específico, em relação ao modelo preditivo utilizado. Os dados após a análise pelo componente Auto ML passam a ter um rótulo atribuído pelo algoritmo de inteligência artificial. As informações enriquecidas são processadas pelo componente *Ontology Model*.

**Ontology Model:** O modelo de ontologia modela as entidades e suas propriedades para um dado contexto. A partir das relações existentes entre as propriedades é possível inferir novas propriedades e enriquecer mais uma vez o *dataset*, através do componente *Inference Machine*.

**Inference Machine:** O componente *Inference Machine* é relacionado ao *Ontology Model*, pois processa os dados relacionados a eventos críticos instanciados na ontologia, juntamente com as informações contextuais adquiridas pelos sensores ambientais e, com base em regras lógicas, produz novas informações e relações que podem ser utilizadas para



ativar o componente *Autonomous Agent*.

**Autonomous Agent:** com base nas informações fornecidas pelos componentes Auto ML, *Ontology Model* e *Inference Machine* o agente autônomo pode disparar alertas ou modificar o *System Environment*. Assim, o *Autonomous Agent* seleciona os dispositivos disponíveis que devem receber os alertas e/ou processar alterações diretamente nos dispositivos ou no funcionamento do *Autonomous Agent*, como desligar uma máquina específica, por exemplo.

**Web Interface:** Componente que permite a apresentação, em um *dashboard*, das informações resultantes do processamento por inteligência artificial e processamento ontológico. Fornece acesso a informações sobre os últimos eventos críticos detectados e exibe mensagens de alertas que influenciam no planejamento da manutenção preditiva. Estabelece a comunicação com dispositivos móveis que acompanham operadores, para fazer chegar um alerta de intervenção necessária em um equipamento, por exemplo.

Dada a importância do processamento inteligente na DT-Create, detalhamos a seguir o suporte inteligente na tomada decisão, considerando processamento inteligente utilizando algoritmos de aprendizagem de máquina (Seção 4.6) e inferências ontológicas (seção 4.7) e sua combinação (seção 4.8), bem como as características de auto-adaptação (seção 4.9).

#### 4.6 USO DE ALGORITMOS DE APRENDIZAGEM DE MÁQUINA

Os cenários de operação industrial contribuem, mesmo que indiretamente, para explicar eventos que merecem atenção como, por exemplo, falhas. Portanto, o cenário e suas informações são úteis para ações preventivas em eventos futuros. Mais especificamente, dados de sensores como temperatura de máquinas, volume de produção de equipamentos, ruídos e vibrações produzidos por componentes, corrente e tensão elétrica entre outros podem fornecer evidências sobre determinados comportamentos e contribuem para a construção de um modelo que permita prever com certa precisão a ocorrência de outros eventos semelhantes ou classificá-los corretamente em uma classe de interesse. Também existem outras informações relacionadas ao processo de manutenção como tempo e custo de reparo que constituem esse conjunto diverso de informações, mas que reunidas enriquecem a construção de um modelo preditivo.

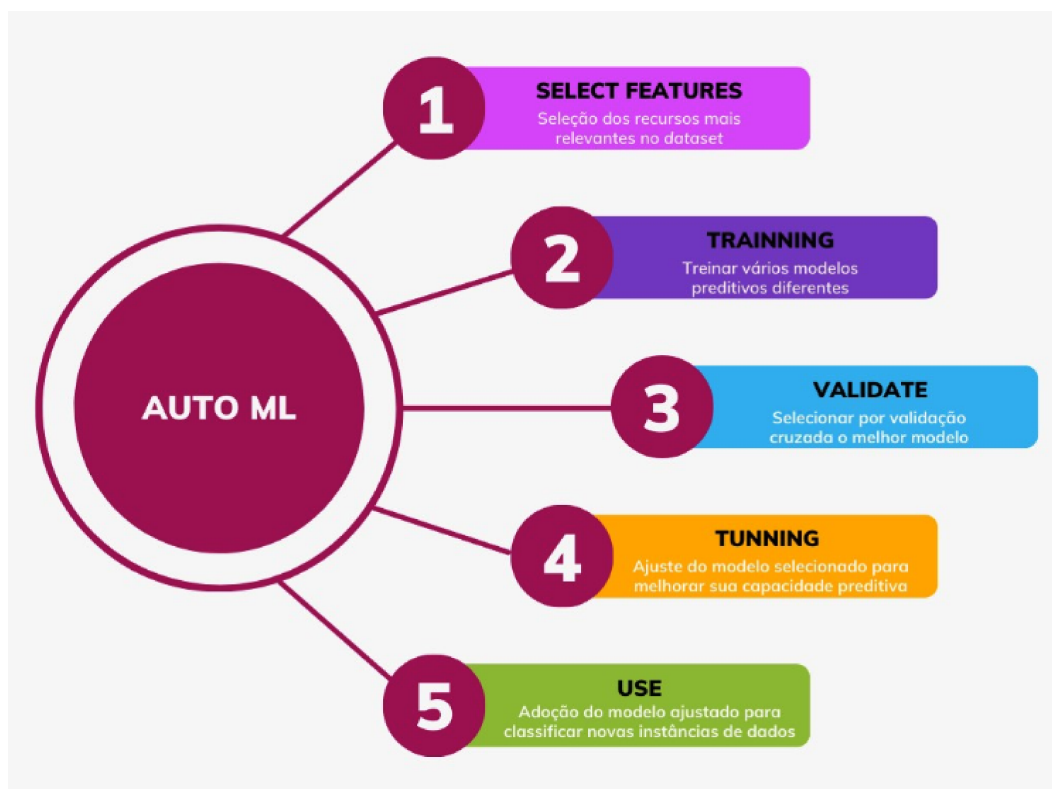
Mesmo que para especialistas as variáveis custo, tempo de reparo, tipo de evento sejam de fácil compreensão, não é viável modelar o cenário de classificação de dados sem o auxílio de ML. Uma das principais contribuições da aprendizagem de máquina é a capacidade de processar e analisar grandes volumes de dados, permitindo a identificação de padrões e tendências que seriam difíceis ou até mesmo impossíveis de serem detectadas por humanos (Mohammed et al., 2016). Portanto, o grande volume de dados dos processos industriais "esconde" conhecimento que pode ser aproveitado para a tomada de decisões a

partir destas técnicas.

Na DT-Create, ML é uma técnica importante considerando a aderência do modelo preditivo alinhado ao *dataset* de entrada. Um *dataset* pode ser obtido a partir de uma base de dados, onde especialistas selecionam ou até mesmo combinam dados dos processos (eventos de falhas, custo de reparo, tempo de reparo) resultando em um conjunto de atributos de interesse para treinamento de modelos de predição. No entanto, mesmo com o *dataset* previamente analisado por especialistas, é importante selecionar o modelo de ML mais adequado. Essa é uma tarefa complexa, que demanda conhecimento sobre o problema e os dados, além da necessidade de uma análise exploratória para compreensão das distribuições, correlações e possíveis problemas nos dados (valores ausentes ou *outliers*).

Portanto, a proposta é treinar e avaliar o desempenho de vários estimadores usando validação cruzada, obtendo o modelo preditivo mais adequado para um determinado *dataset*. Para isso, duas tarefas relacionadas ao ML são previstas: a seleção dos atributos mais relevantes em um *dataset* e a avaliação de modelos. A partir da obtenção do modelo com melhor precisão, dentre os disponíveis, na validação cruzada, o modelo é ajustado para ser mais preciso e colocado em uso. A partir desse momento novos eventos (por exemplo, falhas) podem ser avaliados pelo modelo e conhecimento é obtido, seja na extração de conhecimento a partir dos dados e do processamento por ML, seja nas informações de suporte a decisão que ficam registradas junto ao evento. Em outro momento, o modelo pode ser recalibrado, ou seja, novamente treinado, agora com mais dados enriquecidos a partir do histórico de predições. A Figura 13 mostra o fluxo de ML utilizando o DT-Create.

Figura 13 - Ciclo de AutoML.



Fonte: Elaborada pelo autor (2023).

#### 4.7 USO DE ONTOLOGIAS

A DT-Create integra metodologias de aprendizagem automática e de processamento ontológico para facilitar a extração de conhecimento a partir de dados tanto dos processos, máquinas e equipamentos conectados quanto do contexto onde as operações são realizadas. O objetivo é enriquecer os dados e contribuir nos processos de tomada de decisão.

Assim, após o processamento de dados pelo modelo criado pelo componente Auto ML uma ontologia relacionada ao cenário de aplicação do DT é instanciada pelo componente *Ontology Model* para processar semanticamente informações previamente processadas com dados de contexto. O processamento semântico, que contribui para processar um evento enriquecendo semanticamente os dados já processados, com a possibilidade de extrair conhecimentos novos. Para isso, o componente *Inference Machine*, a partir de regras SWRL (*Semantic Web Rule Language*) e processamento de algoritmos de inferência, pode estabelecer novos relacionamentos entre os dados, derivando novas informações, que servirão de entrada para execução de ações automáticas pelo componente *Autonomous Agent*.

## 4.8 ONTOLOGIA E APRENDIZAGEM DE MÁQUINA

Nas seções anteriores, discutimos como a ontologia pode ser usada para extrair conhecimento semântico a partir de dados, realizar inferências e derivar relacionamentos entre entidades e também detalhamos como o treinamento de modelos preditivos é conduzido na DT-Create. Embora essas abordagens possam contribuir de forma independente, o diferencial da proposta está na combinação do Auto ML e do processamento de regras ontológicas, que permite a combinação dos resultados do processamento por aprendizagem de máquina com enriquecimento semântico das informações. Dessa forma é possível um suporte à decisão baseado tanto na predição de estado de processos, máquinas e equipamentos e também de variáveis de ambiente que contribuem para estabelecer um nível operacional crítico, que demanda intervenção relacionada a manutenção preditiva.

## 4.9 AUTO ADAPTAÇÃO

A DT-Create oferece auto adaptação em dois pontos importantes no processo de suporte à manutenção preditiva. No primeiro ciclo DSR o artefato elaborado apresentou no componente Autonomous Agent a capacidade de enviar mensagens de alerta para usuários ou outros sistemas. A decisão de ação do agente é baseada nas saídas geradas pelos componentes AutoML e *Inference Machine*, respectivamente.

No entanto, no segundo ciclo DSR, percebemos que os conjuntos de dados que servem de entrada para toda a cadeia de processamento, dada à sua natureza, podem exigir a seleção de modelos de aprendizagem de máquina diferentes para cada tipo de *dataset* de entrada. Dessa forma, o componente AutoML é capaz de selecionar o modelo de aprendizagem de acordo com o *dataset* disponível. Além disso, o modelo preditivo resultante do processo pode ser novamente processado quando um novo *dataset* é selecionado para treinamento, ou de forma automática de acordo com as características do conjunto de dados onde o modelo preditivo é utilizado. As características de auto-adaptação neste segundo ciclo foram aprimoradas, considerando inclusive a possibilidade de enviar mensagens de parada de equipamentos em casos extremos ou desvio de produção para outros equipamentos, permitindo a auto-adaptação diretamente no sistema produtivo.

## 4.10 ESTRATÉGIAS DE IMPLEMENTAÇÃO

A suíte de serviços DT-create foi desenvolvida, considerando os requisitos funcionais e não funcionais ressaltados nas seções 4.2 e 4.3. Os RF01 e RF02 foram alcançados a partir do *Broker* especificado no componente *System Environment*. Os requisitos funcionais RF03 foi alcançado a partir dos componentes Auto ML, *Ontology Model* e *Inference Machine*. O RF04 foi atingido a partir do armazenamento dos dados em um banco de dados. O RF5 foi especificado a partir do componente *Web Interface*. O RF6 pode ser atingido a partir

de um melhor suporte a tomada de decisão, privilegiando assim o fator humano. Além disso, a ações do componente Autonomous Agent ajudam na questão da sustentabilidade, uma vez que contribui para uma maior sobrevida dos equipamentos, a partir da prevenção de falhas críticas. Os requisitos Funcionais RF07 e RF08 foram cumpridos a partir da evolução da DT-Create, com características de auto-adaptação e seleção de modelos de aprendizagem.

Por fim, os requisitos não funcionais RNF01 e RNF02 foram atingidos a partir do módulo *Broker* do componente *System Environment*. O RNF03 e RNF04 foi atingido a partir dos componente AutoML e o RNF05 a partir também do componente *Web Interface*.

A linguagem de programação *Python* foi utilizada para a especificação do DT-Create. A escolha da linguagem *Python* é justificada pela flexibilidade na instalação de bibliotecas essenciais ao pré-processamento e ao desenvolvimento de modelos preditivos. Bibliotecas como *Pandas*, que simplificam o carregamento de conjuntos de dados em formatos como CSV, texto ou planilhas, facilitam a análise e manipulação de dados. A biblioteca *PyCaret*, dá suporte à criação de modelos preditivos por meio de técnicas de aprendizado de máquina, e a biblioteca *Flask* oferece uma estrutura *Python* leve que permite o desenvolvimento rápido de aplicações web. Além disso, a manipulação de ontologias é conduzida de forma eficiente utilizando a biblioteca *OWL2Ready*. Consequentemente, esta uniformidade de linguagem de programação em todos os módulos da aplicação agiliza a manutenção e a integração de novas funcionalidades. O armazenamento de dados é gerenciado usando o *MySQL Database Management System* (DBMS) devido à sua ampla aceitação e facilidade de integração com outros sistemas predominantes em ambientes organizacionais, permitindo a entrada rápida de dados sem a necessidade de modificações iniciais. As linguagens HTML e *JavaScript* foram utilizadas no desenvolvimento do componente *Web Interface*. Para isso, projetou-se uma interface de usuário web, caracterizada pela sua capacidade responsiva, que pode ser utilizada sem a necessidade de instalação de *softwares* adicionais em dispositivos de campo, tais como *smartphones* e *tablets*, o que favorece a usabilidade, sendo compatível com a mais ampla gama de dispositivos possível.

#### 4.11 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Este capítulo apresentou os ciclos de DSR para a especificação da suíte de serviços DT-Create. Foram detalhados os componentes processamento inteligente da DT-Create, incluindo o uso de ontologias e técnicas de aprendizado de máquina. No Capítulo 5 apresentamos a avaliação da DT-Create a partir de dois estudos de caso.

## 5 AVALIAÇÃO DA DT-CREATE

Este capítulo descreve a avaliação da suíte de serviços DT-Create, a partir da especificação de DTs em dois cenários de aplicação diferentes, ambos relacionados a aplicações na indústria. O primeiro relacionado a falhas em equipamentos utilizados na indústria têxtil e o segundo relacionado a falhas em máquinas da indústria metalúrgica.

Conforme já dito no capítulo anterior, a DSR tem seu maior esforço concentrado no ciclo de *design*, ou seja, construção e avaliação dos artefatos obtidos (Hevner, 2007) e, por sua importância, esse processo deve ser bem detalhado. Portanto, tanto a construção (Capítulo 4) quanto a avaliação devem ser rigorosos. Nesse processo o rigor é traduzido pela aplicação do conjunto de fundamentos e metodologias necessárias à pesquisa. Como resultado, em cada um dos ciclos, conhecimento científico foi produzido, servindo de base para novas versões dos componentes que fazem parte da solução proposta. Essa construção de versões é um processo iterativo, onde se verifica a cada iteração se o artefato resolveu o problema proposto. Além disso, a cada iteração o artefato passa por refinamentos para produzir melhores resultados.

Para realizar o estudo, foram utilizados dados de sensores e equipamentos, que foram processados e analisados para verificar como esses dados enriquecidos apoiam a tomada de decisões em relação à manutenção preditiva. Nesta avaliação, adotamos os seguintes passos, definição do escopo (Seção 5.1), detalhamento dos cenários de avaliação (Seção 5.2), Especificação das Ontologias (seção 5.3), Estudo de Caso 1 (Seção 5.4), Estudo de Caso 2 (Seção 5.5).

### 5.1 DEFINIÇÃO DE ESCOPO

Para avaliar a solução proposta, definimos o escopo da avaliação através do GQM (Goal, Questions, Metrics) (Basili and Weiss, 1984): O objetivo é analisar o suporte a tomada de decisões relacionadas a manutenção preditiva, considerando dados de sensores e equipamentos, no contexto da Indústria 5.0. A Questão de Pesquisa a ser investigada nesta avaliação é: “Como apoiar a tomada de decisão relacionada a manutenção preditiva na Indústria 5.0, com o uso de DTs?” Portanto, na avaliação, DTs especificados a partir da suíte de serviços DT-Create devem ser capazes de gerar *insights* sobre possíveis falhas relacionadas aos equipamentos monitorados, considerando também o ambiente de operação e informações de contexto.

Esta avaliação teve como objetivo verificar se o artefato funciona, se as conjecturas teóricas estão alinhadas com as expectativas e se o artefato (DT-Create) auxilia na especificação de DTs para apoio na tomada de decisões de manutenção preditiva. A partir dos resultados da avaliação do artefato no primeiro ciclo e também no conhecimento científico adquirido, confrontamos as conjecturas teóricas levantadas durante a revisão

da literatura. Realizamos mais uma interação, melhoramos o artefato e o mesmo foi reavaliado. Por fim, verificamos se o artefato responde à questão de pesquisa (QP), levando em consideração análises mais complexas apoiadas em aprendizado de máquina e ontologias. Para comunicar resultados, publicamos artigos em conferências e *workshops* (Esteves et al., 2023; Silva et al., 2023; Da Silva et al., 2021; Gomes et al., 2023).

A Figura 14 ilustra os elementos de avaliação utilizados neste trabalho, apresentando a abordagem teórica da pesquisa (elementos do lado direito da figura), o artefato (elementos do lado esquerdo) e o contexto de aplicação (elementos no topo da figura). Esta separação evidencia a correlação entre o desenvolvimento tecnológico aplicado e o conhecimento teórico-científico.

De acordo com o Referencial Teórico apresentado na Figura 10, soluções encontradas no MSL suportam um domínio complexo cujo contexto pode mudar rapidamente. Na indústria e mais especificamente na Indústria 5.0, novos sensores podem ser instalados em máquinas e equipamentos como também novos dispositivos IoT podem ser adicionados. Ainda assim, as soluções devem continuar a funcionar adequadamente, apoiando os gestores de manutenção na tomada de decisões. Também a solução deve continuar oferecer auto adaptação do sistema produtivo a partir de do processamento lógico de uma ontologia, mesmo quando o cenário de aplicação mudar e, por sua vez, também o conjunto de dados. Assim, a solução deve permitir a especificação de DTs de suporte a manutenção preditiva, a despeito das variações de cenários de aplicação.

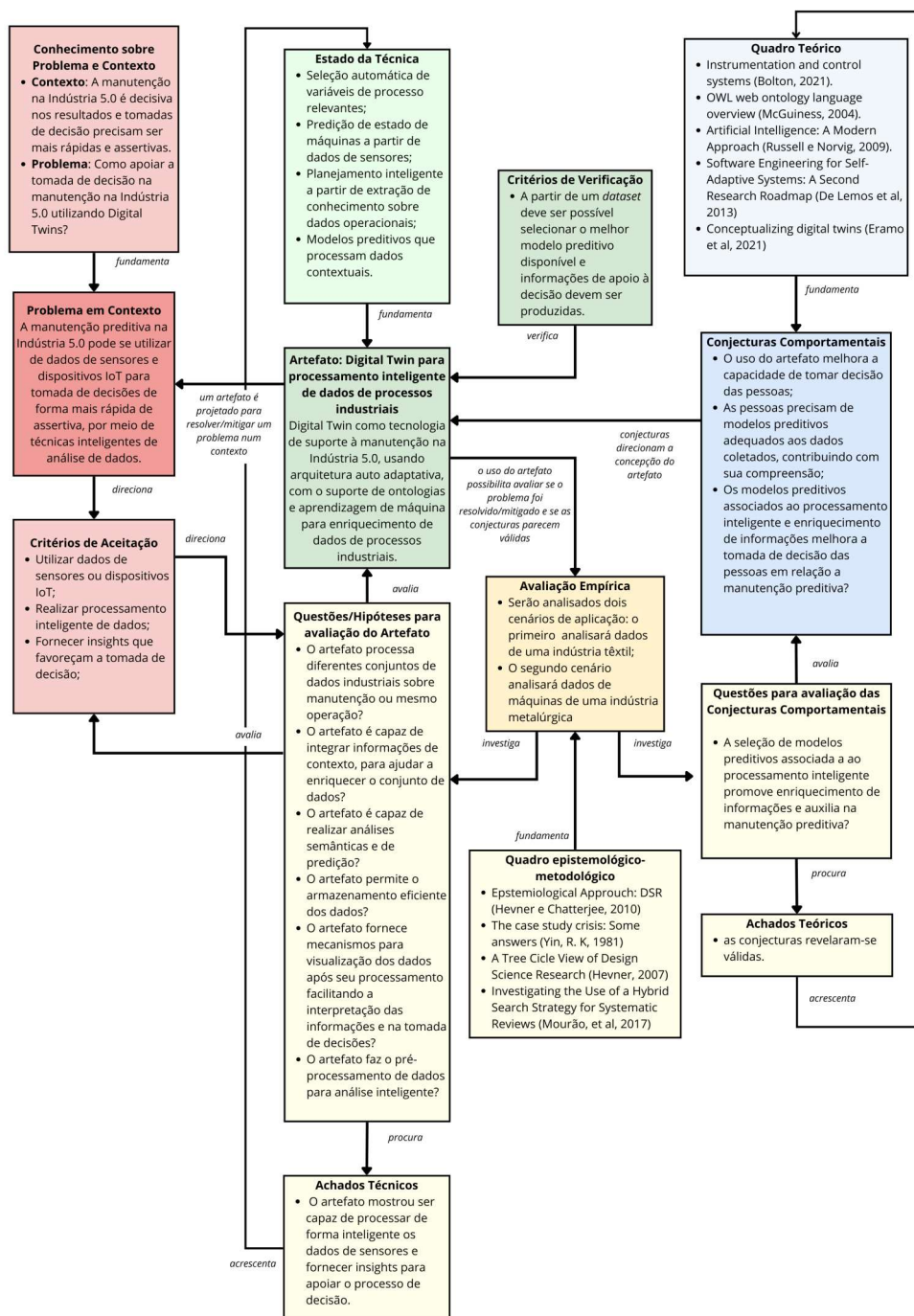
No primeiro ciclo de avaliação, especificamos a suíte de serviços DT-Create para processar dados de sensores de ambiente, dados de falhas e realizamos um estudo de caso para avaliar o artefato. No segundo ciclo, os componentes de processamento inteligente foram melhorados e conduzimos o segundo estudo de caso. Os estudos de caso referem-se a cenários onde a solução forneceu suporte à decisão na manutenção preditiva, a partir da aplicação de técnicas inteligentes sobre dados de falhas de equipamentos e de contexto. Em ambos os ciclos ontologias específicas foram instanciadas.

## 5.2 ESPECIFICAÇÃO DAS ONTOLOGIAS

Conforme já dito, para apoiar o desenvolvimento de DTs a partir da suíte de serviços DT-Create, é necessário o apoio de uma ontologia relacionada ao domínio de aplicação. No Estudo de Caso 1 foi especificada a ontologia *SmartMaintenace*. No Estudo de Caso 2 foi especificada a ontologia *SensorEquipment*.

Para desenvolver as ontologias foram executadas as seguintes fases (De Nicola e Missikoff, 2016), (De Nicola e Missikoff, 2016): (i) Especificação, (ii) Conceitualização, (iii) Formalização e (iv) Avaliação. Na fase de especificação, identifica-se o propósito, o escopo, a linguagem de implementação e as Questões de Competência (QC) pretendidas.

Figura 14 - Elementos DSR.



Fonte: (Pimentel et al., 2020) (adaptada).

Uma QC é uma sentença em linguagem natural que expressa um padrão para uma questão que as pessoas/aplicações computacionais esperam que uma ontologia responda. As QC para cada ontologia foram elaboradas, baseadas em entrevistas com especialistas e em documentos relacionados a manutenção preditiva. Utilizamos conceitos ontológicos, incluindo suas classes, relações e regras de inferência, para responder a cada QP. Portanto, para cada cenário de avaliação, ontologias específicas foram instanciadas, assim como as



questões de competência relacionadas.

A fase de conceituação focou na organização e estruturação do significado semântico dos dados. Esta fase baseou-se na especificação dos termos e nas relações definidas entre esses termos. A fase de formalização foi realizada utilizando a ferramenta *Protegé* e as linguagens OWL 2.0 e SWRL. Definimos também um conjunto de regras para apoiar o processamento semântico dos termos e descobrir novas relações entre eles. Por fim, a fase de avaliação consistiu na realização de verificação e validação. A etapa de verificação consiste em verificar a exatidão e validar a ontologia. A correção das ontologias foi feita através de um processo de verificação utilizando o plugin *Pellet reasoner*. Esta validação garante que a ontologia cumpra seu propósito por meio da resposta as Questões de Competência (QC). Com as ontologias formalizadas e validadas, as mesmas foram utilizadas para o processamento semântico de informações relacionadas a manutenção preditiva de equipamentos industriais. Especialistas brasileiros em manutenção preditiva de equipamentos industriais validaram a ontologia.

### 5.3 CENÁRIOS DE AVALIAÇÃO

Os estudos de caso foram conduzidos de acordo com as seguintes etapas (Runeson e Martin, 2021) (Runeson et al., 2012): (i) projeto do estudo de caso (preparação e planejamento para coleta de dados), (ii) execução (coleta de evidências), (iii) análise dos dados coletados e (iv) relatório dos resultados. Os cenários de estudo de caso são industriais, mais especificamente indústria têxtil na produção de tecidos (Estudo de Caso 1) e metalúrgica em processos de tratamento térmico (Estudo de Caso 2).

No Estudo de Caso 1 (CS1) (Ciclo 1 do DSR) foram utilizados dados históricos de falhas de máquinas da indústria têxtil, tais como: tipo de falha ocorrida, momento do evento de falha, tempo de reparo, custo, criticidade da falha, temperatura e umidade do ambiente no momento do evento de falha. Para facilitar a reprodutibilidade deste estudo de caso, os dados coletados estão disponíveis no GitHub.<sup>1</sup>

No segundo ciclo, realizamos o Estudo de Caso 2 (CS2) onde foram utilizados dados de estado de máquinas coletados de fornos de tratamento térmico de uma indústria. As máquinas possuem sensores coletando informações sobre: vibração global de máquina, colisão, desbalanceamento, rede elétrica, rolamentos 1, 2 e 3. Durante a condução do estudo, tivemos contato direto com os especialistas da indústria metalúrgica. Entrevistas técnicas foram realizadas, considerando que o foco era entender como os dados poderiam apoiar a tomada de decisões e quais cenários operacionais críticos precisavam ser mapeados. Os dados coletados não foram disponibilizados por questões de confidencialidade.

---

<sup>1</sup> *Dataset* com dados de falhas.

## 5.4 ESTUDO DE CASO 1 (PRIMEIRO CICLO DSR)

### 5.4.1 CONDUÇÃO

A indústria têxtil é um dos grandes *players* da Indústria 4.0 no Brasil. A manutenção da produtividade na indústria têxtil se deve à automação de máquinas e processos industriais, entre outros fatores. É estratégico utilizar técnicas que aceleram a tomada de decisões deste setor para aumentar a produtividade. Porém, sua continuidade ainda depende de estratégias que garantam sua constante inovação, incluindo sensores para coleta de dados e o uso crescente de técnicas de análise computacional de informações.

Neste primeiro estudo de caso, utilizamos a suíte de serviços em um contexto de sistema de produção têxtil. Os resultados iniciais apoiam a detecção preditiva de falhas nas máquinas e a consequente redução de custos operacionais. Portanto, a partir da análise dos dados, a arquitetura pode disparar alertas ou processar alterações diretamente nos dispositivos, fornecendo uma suíte de serviços inteligente e contextual para a Indústria 4.0.

As máquinas da indústria têxtil, principalmente do setor de “remalhação”, apresentam falhas durante o seu funcionamento que normalmente são resolvidas rapidamente, retornando a máquina ao estado operacional. Porém, no médio prazo, essas falhas podem indicar ineficiência operacional. Assim, avaliar a criticidade das falhas por meio de análise histórica de eventos e dados ambientais pode auxiliar na manutenção preditiva dos equipamentos.

No contexto produtivo, uma falha tolerada pode, no curto prazo, ser responsável por uma falha crítica ou até mesmo pela parada da máquina. Por exemplo, a sucessão de quebras de agulhas em uma máquina pode contribuir diretamente para a quebra de um componente mecânico e comprometer um lote de produção.

Associadas aos eventos de falha, temos informações do ambiente que contribuem indiretamente para a ocorrência de falhas críticas. Por exemplo, a temperatura e a umidade do ambiente de produção interferem nas propriedades dos fios e afetam as condições de funcionamento das máquinas. Portanto, informações contextuais são importantes neste cenário.

Geralmente, a temperatura e a umidade do ambiente influenciam substancialmente a produtividade do equipamento. Além disso, a proximidade e a experiência de um determinado funcionário podem ser cruciais na detecção da probabilidade de uma falha ou na análise do ambiente para a ocorrência de uma falha.

Para este estudo um *dataset* com 10.000 registros de falhas e outras informações técnicas e gerenciais foi utilizado. O *dataset* possui como variáveis a identificação da máquina, o tipo de falha, timestamp do evento de falha, time de reparo, custo do reparo, umidade e temperatura no momento da falha. Sobre os dados para supervisão temos a classe em que a falha foi classificada de acordo com a percepção da própria indústria: classe 1 para

falhas críticas e 0 para falhas não críticas. Os registros de falhas para treinamento já vêm acompanhados de sinalização de criticidade, o que favorece o aprendizado supervisionado do modelo. A Figura 15 mostra o *dataset* de entrada para aprendizagem supervisionada do modelo de classificação.

Figura 15 - *Dataset* com dados de falhas.

	machine_id	type_of_failure	timestamp	time_repair	cost	criticality	humid	temp	label
0	81	7	1583193600	0.363742	0.522	0.499	61	97	0
1	66	5	1593388800	0.297000	0.485	0.307	63	112	0
2	32	9	1600128000	0.383106	-0.075	0.798	6	40	1
3	43	4	1597708800	0.300732	0.553	0.531	48	141	0
4	62	2	1585699200	0.127000	0.335	0.367	74	148	0
...	...	...	...	...	...	...	...	...	...
9994	34	3	1599350400	0.391921	0.518	0.508	6	56	0
9995	78	3	1598918400	0.102000	0.494	0.272	56	88	0
9996	81	9	1595376000	0.339469	0.277	0.456	80	142	0
9997	97	7	1600905600	0.277222	0.457	0.424	61	99	0
9998	44	9	1581206400	0.250000	0.309	0.282	42	92	0

Fonte: Elaborada pelo autor (2023).

Um modelo de classificação de criticidade de falhas foi implementado em *Python* e utilizado neste estudo de caso. O objetivo era que as falhas fossem automaticamente classificadas como críticas ou não, de acordo com o aprendizado supervisionado. Utilizando a biblioteca *Scikit-Learn*, o *dataset* com 10.000 registros de falhas foi organizado em duas partes, uma com 75% dos dados para treinamento do modelo e os 25% restantes para avaliar a precisão do modelo nas previsões de classe. O *dataset* passou pelo pré-processamento para garantir a normalização dos dados e separação de dados para treino e teste, conforme mostra a Figura16:

Figura 16 - Preparação dos dados antes da criação do modelo preditivo.

```
#Pré-processamento
from sklearn import preprocessing
scaler = preprocessing.MinMaxScaler()
normalizados=pd.DataFrame(scaler.fit_transform(dados), columns=dados.columns)

# Divisão do dataset em dados de treino e teste
X_treino=dataframe_dados[:7500]
X_teste=dataframe_dados[7501:]

# Adição da coluna label do DataFrame
dataframe_labels = pd.DataFrame()
dataframe_labels['label'] = normalizados['label']

# Divisão do dataset em dados de treino e teste
y_treino = dataframe_labels[:7500]
y_teste = dataframe_labels[7501:]
```

Fonte: Elaborada pelo autor (2023).

O aprendizado de máquina foi feito através de uma rede neural baseada no classificador multicamadas *Percetron* (Rosenblatt, 1962) que possui uma ou várias camadas ocultas com um determinado número de neurônios e é treinada através de um algoritmo de retro propagação. Desta forma, os novos registros de falhas podem ser classificados automaticamente pela rede neural e são instanciados no modelo ontológico. Para avaliar a assertividade do modelo criado realizamos a validação cruzada, que obteve como resultado a precisão de 99%, como mostra a Figura 17:

Figura 17 - Validação cruzada do modelo preditivo.

```

from sklearn.model_selection import cross_val_score

# Realizar a validação cruzada
scores = cross_val_score(clf, y_teste, y_aux, cv=8) # cv é o número de folds

# Exibir os resultados
print("Accuracy scores for each fold:", scores)
print("Mean cross-validation score:", scores.mean())

Accuracy scores for each fold: [0.99680511 1.          0.99679487 0.99679487 1.          0.99679487
 0.99679487 0.99679487]
Mean cross-validation score: 0.9975974338494307

```

Fonte: Elaborada pelo autor (2023).

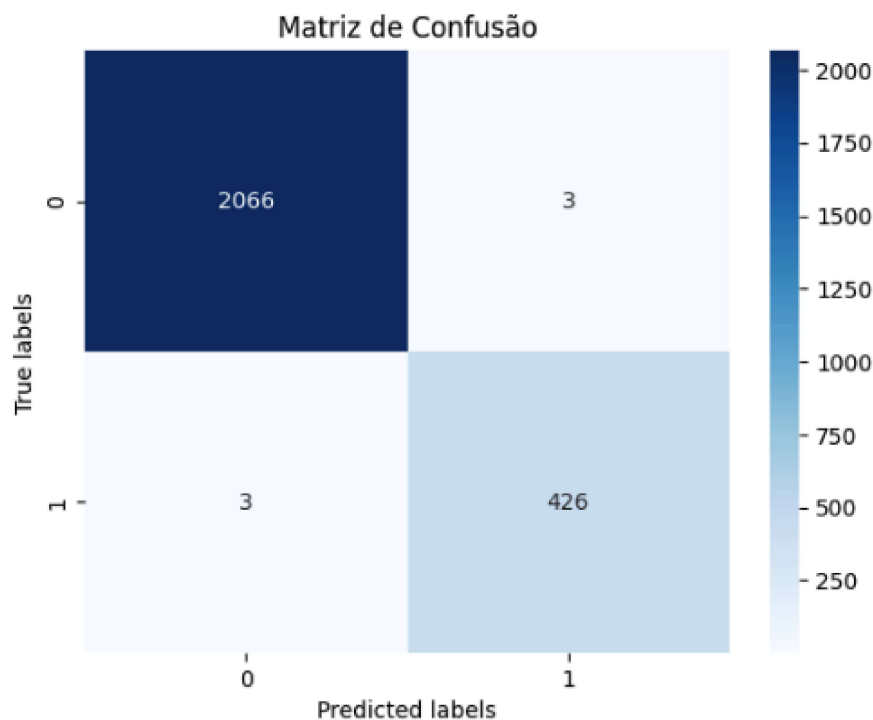
De forma adicional, avaliamos o desempenho do modelo de classificação em relação à forma com que as classes de interesse são determinadas. Após o teste uma matriz de confusão foi gerada e mostrada na Figura 18. Uma matriz de confusão permite verificar como um modelo de classificação define os rótulos das classes de interesse, através da comparação entre o que foi previsto pelo modelo e o que de fato está definido no conjunto de dados de teste. A ocorrência de falsos positivos e falsos negativos pode comprometer a precisão do modelo preditivo e a matriz de confusão permite justamente observar a distribuição das predições dentro das quatro categorias: verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos.

Analisando a distribuição das classes entre os rótulos previstos e rótulos reais foi possível perceber que o modelo tem elevada precisão na predição tanto para classes positivas quanto negativas. A elevada precisão do modelo por ser explicada por fatores como tamanho do *dataset* disponível para teste do modelo e desequilíbrio entre as classes positiva e negativa.

O resultado do processamento via rede neural foi um novo *dataset* com classes definidas. Após utilizar a rede neural para processar o *dataset*, o *dataset* resultante com os valores de criticidade atribuídos, foi processado pela ontologia *SmartMaintenance*, detalhada a seguir.

A arquitetura auto adaptativa favorece essa possibilidade por meio da disponibilização de dados oriundos de aprendizagem de máquina, combinados com dados de contexto, e processados por meio de ontologias e uso de agentes de *software* autônomos. Assim, esse

Figura 18 - Matriz de confusão para análise do modelo preditivo.



Fonte: Elaborada pelo autor (2023).

cenário considera três importantes elementos para gerenciamento e manutenção de falhas: construção de conhecimento sobre falhas, monitoramento de falhas críticas iminentes e autocontrole de equipamentos de produção através de agentes de *software* autônomos.

#### 5.4.2 ONTOLOGIA SMARTMAINTENANCE

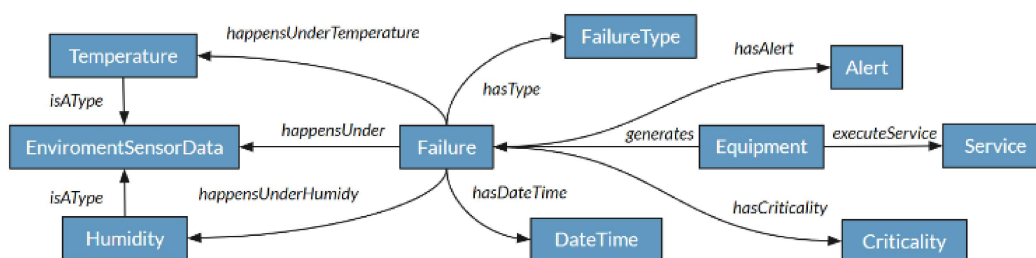
Conforme ressaltado na seção, uma ontologia específica foi criada para representar o cenário de ocorrência de falhas no ambiente têxtil e produzir conhecimento. Assim, a *SmartMaintenance Ontology* foi desenvolvida para apoiar o processamento semântico de informações na manutenção preditiva de falhas em equipamentos textéis (finalidade) e o escopo foi definido considerando o domínio de manutenção preditiva (escopo). As linguagens OWL 2.0 e SWRL foram utilizadas como linguagens de implementação e biblioteca *OWLReady2* utilizada para importação dos dados de um banco de dados relacional para a ontologia. Os usuários são tomadores de decisão relacionadas a manutenção preditiva na indústria.

Questões de Competência foram assim definidas, com a ajuda de especialistas: (i) Qual é a temperatura média no momento da ocorrência das falhas? (ii) As falhas ocorrem mais frequentemente em condições de alta umidade comparadas a condições de baixa umidade? (iii) Qual tipo de falha é mais comum em determinadas condições ambientais? (iv) Existem diferenças na ocorrência de falhas entre os turnos de trabalho? (v) Qual é a criticidade média das falhas que ocorrem em condições de alta temperatura versus baixa

temperatura?

Para responder as QCs, foram especificadas regras lógicas relacionadas aos termos da ontologia. As principais classes e relacionamentos (*objectProperties*) da ontologia especificada são apresentadas na Figura 19.

Figura 19 - Principais classes e associações da ontologia *SmartMaintenance*.



Fonte: Elaborada pelo autor (2023).

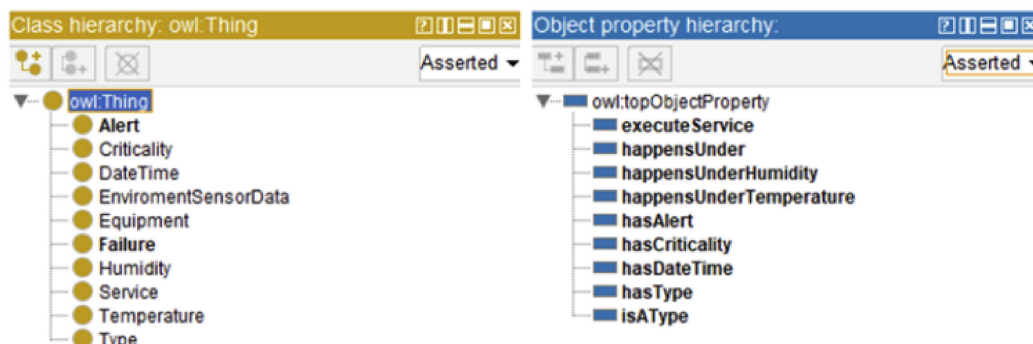
A classe **Failure** representa o evento de falha que contém informações como o código de falha do equipamento, o setor ao qual o equipamento pertence e também o código que identifica o evento de falha. **FailureType** é uma subclasse de **Failure**, representando o tipo de falha. Associada à classe **Failure** temos a classe **DateTime** que fornece o *timestamp* do evento de falha. As falhas registradas ocorrem em ambiente de produção onde são controladas as variáveis de umidade e temperatura, que são propriedades da classe **Failure**.

A classe **Criticality** representa o grau de criticidade estimado de uma falha pelo componente de processamento inteligente. A classe **Alert** está associada à classe **Failure** através da *objectProperty* **hasAlert**, que é definida dinamicamente a partir de regras SWRL que consideram características específicas das ocorrências de falhas, como a quantidade de determinado evento classificado como crítico, e também dados ambientais como temperatura e umidade. A Figura 20 mostra as principais classes e suas *objectProperties* na ferramenta *Protégé*.

As falhas registradas ocorrem em ambiente de produção onde são controladas as variáveis de umidade e temperatura, que são propriedades da classe **Failure**. A classe **Criticality** representa o grau de criticidade de uma falha estimada pelo modelo preditivo. A criticidade de uma falha pode ser projetada a partir de dados históricos, custos, tempo de reparo e outros.

Informações de contexto em tempo real enriquecem tais informações e aproveitam o conhecimento dos supervisores dos processos quando são estabelecidas regras lógicas para processamento ontológico. Essa combinação é um diferencial na gestão dos processos de manutenção e oferece suporte à decisão nas intervenções de manutenção preditiva. A Tabela 6 mostra as principais regras SWRL especificadas, de acordo com parâmetros

Figura 20 - Classes e propriedades da ontologia SmartMaintenance.



Fonte: Elaborada pelo autor (2023).

operacionais definidos por especialistas, que consideram os índices de temperatura e umidade do ambiente.

Tabela 6 – Avaliando temperatura e umidade por regras SWRL

#	Regras SWRL	Descrição
01	<p><b>Antecedente</b>            Failure(?f), Alert(?a), Humidity(?h),            humidityValue(?h, ?humid),            Temperature(?t),            temperatureValue(?t, ?temp),            lessThanOrEqual(?humid, 25),            greaterThanOrEqual(?temp, 35)</p> <p><b>Consequente</b>            -&gt; alertCode(?f, 100)</p>	<p>Se uma falha ocorrer e for detectada uma condição em que a umidade é menor ou igual a 25% e a temperatura é maior ou igual a 35°C, então associe o código de alerta 100 a essa falha".</p>
02	<p><b>Antecedente</b>            Temperature(?t), Failure(?f),            TemperatureValue(?t, ?temp),            Failure(?f), equal(?type, 4),            numberOfOccurrences(?f, ?num),            typeOfFailure(?f, ?type),            greaterThanOrEqual(?num, 4)</p> <p><b>Consequente</b>            -&gt; alertCode(?f, 100)</p>	<p>Se existe uma temperatura (Temperature) associada a um valor (temperatureValue) que é maior ou igual a 30, e existe uma falha (Failure) com um tipo específico (typeOfFailure) igual a 4, e essa falha ocorreu pelo menos 4 vezes (numberOfOccurrences maior ou igual a 4), então atribua o código de alerta 200 (alertCode) a essa falha.</p>

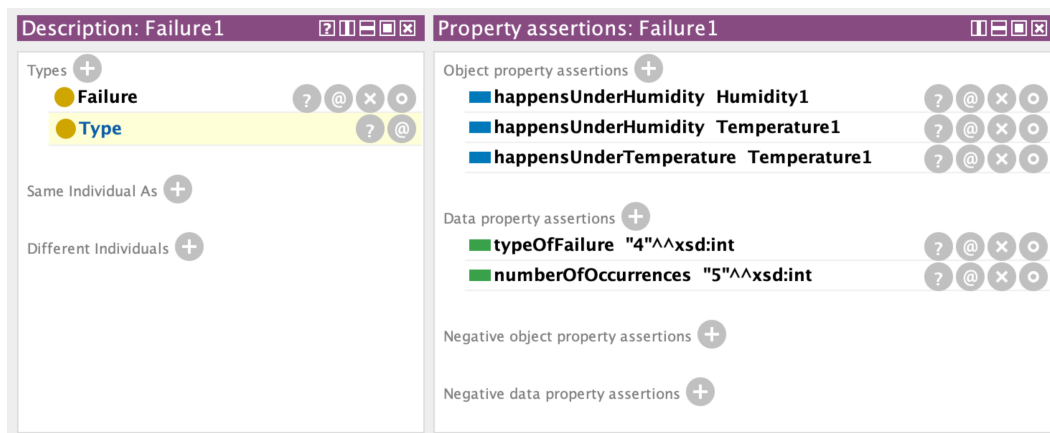
Fonte: Elaborada pelo autor (2023).

Considerando o conjunto de dados deste Estudo de Caso 1, a ontologia processou a combinação dos dados de falha já classificados pela rede neural com os dados ambientais e métricas de falha qualitativas e quantitativas que devem ser observadas. A Figura 21 detalha um exemplo de como a ontologia processou a inferência de propriedades da



instância “Failure1” com acumulado de falhas e atributos de tipo de falha definidos respectivamente como "4" e "5".

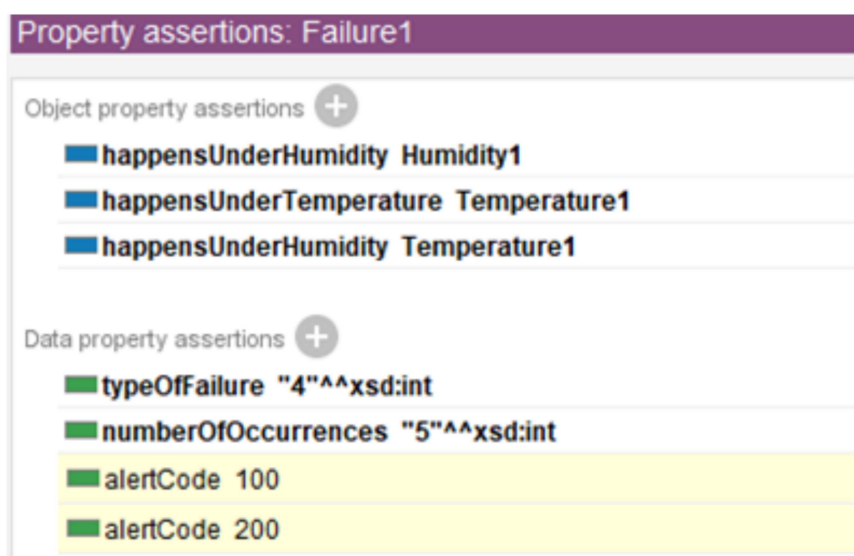
Figura 21 - Propriedades inferidas pela ontologia (*software* Protégé)



Fonte: Elaborada pelo autor (2023).

Considerando também as informações ambientais (temperatura e umidade), as regras do SWRL avaliaram os cenários onde uma intervenção técnica era urgente para mitigar o risco. A regra 01 da Tabela 6 gera um código de alerta com valor 100 quando a temperatura ambiente foi superior a 35°C e a umidade ambiente foi inferior a 25%. Em outro exemplo de uso de regra SWRL, processada agora pela regra 02 da Tabela 6, tem-se instância "Fault1", definida com os seguintes atributos e valores: temperatura: 45°C, umidade: 20%, tipo de falha: 4, e o número de ocorrências de falhas: 5. Após a execução da inferência, a ontologia processou dois alertas com os códigos “100” e “200”, conforme mostrado na Figura 22.

Figura 22 - Regra SWRL para geração de alerta representada no *software* Protégé.



Fonte: Elaborada pelo autor (2023).



### 5.4.3 ANÁLISE

Portanto, a partir dos dados utilizados no estudo de caso 1, i.e., dados de máquinas têxteis e dados ambientais, processados por uma rede neural e posteriormente instanciados na ontologia *SmartMaintenance* e processados pelas regras lógicas, foi possível obter *insights* relacionados a manutenção preditiva. Por exemplo, a maioria das falhas críticas possuíam relação direta com a umidade do ambiente, ou seja, os alertas gerados pela ontologia que envolviam elevada condição de umidade incidiam sobre eventos de falhas classificadas como críticas pelo modelo preditivo. As inferências produzidas sinalizam a necessidade de envio de alertas a partir do agente autônomo aos operadores próximos para verificar itens específicos da máquina, como, por exemplo, em um caso específico, agulhas que poderiam quebrar. A solução foi a interrupção antecipada, antes que falhas mais críticas pudessem comprometer o funcionamento do equipamento por mais tempo.

Assim, os resultados iniciais apontam para o apoio a detecção preditiva de falhas nas máquinas e a consequente redução de custos operacionais. Portanto, a partir da análise dos dados, o DT criado pôde disparar alertas, fornecendo uma solução inteligente e contextual para a Indústria 5.0, onde o fator humano e a sustentabilidade são destacados.

A partir da necessidade de interrupção por parte do operador, foi possível vislumbrar uma possibilidade de evolução da solução, no sentido do próprio DT acionar a interrupção do equipamento, sem a necessidade do operador. Neste último caso, mensagens de alerta podem ser enviadas diretamente às máquinas ou dispositivos IoT responsáveis pelas máquinas através do protocolo de troca de mensagens MQTT, de acordo com o alerta gerado, para preservar o equipamento ou mesmo garantir condições de segurança para a operação. Outro ponto crítico é o uso exclusivo da rede neural. Dependendo do tipo de dados oriundo de sensores e das variáveis utilizadas, uma rede neural pode não ser a melhor técnica de aprendizado de máquina a ser utilizada, dadas as características e especificidades de cada *dataset*. Assim, para o próximo ciclo de desenvolvimento do artefato vislumbrou-se a possibilidade de escolha da técnica de aprendizagem de máquina mais adequada a ser utilizada.

### 5.4.4 RESULTADOS

Desta forma, a solução proposta no primeiro ciclo DSR, orientada a dados de falhas de máquinas, processamento de redes neurais e modelos semânticos, foi capaz de avaliar conjuntamente a criticidade das falhas associadas aos dados ambientais e as características qualitativas e quantitativas das falhas enriquecendo a informação. Também produziu dados que permitem a ação do agente autônomo por meio do envio de mensagens de alerta via protocolo MQTT aos operadores das máquinas. Portanto, foi possível indicar que o uso combinado de processamento semântico e de aprendizagem de máquina permitiu que agentes de *software* autônomos detectassem falhas relacionadas à manutenção. Contudo,

perceamos que era necessário incrementar o artefato no que diz respeito à seleção de um modelo preditivo para processamento de dados e, a partir da constatação desse ponto de melhoria, um novo ciclo de DSR foi conduzido.

## 5.5 ESTUDO DE CASO 2 (SEGUNDO CICLO DSR)

### 5.5.1 CONDUÇÃO

A indústria metalúrgica utiliza em seus processos de fabricação mecânica fornos de tratamento térmico, equipamentos industriais capazes de alterar as propriedades físicas e, em alguns casos, químicas dos materiais através da aplicação controlada de calor. Esses fornos operam em uma ampla gama de temperaturas e são fundamentais em processos como têmpera, recozimento, endurecimento, normalização, entre outros, aplicados principalmente em metais e ligas para melhorar sua resistência, ductilidade, tenacidade, e outras características essenciais para a aplicação específica do material.

Considerando a manutenção Baseada em Condição (CBM), a análise de vibração é uma das técnicas utilizadas e é baseada no monitoramento e análise das vibrações emitidas pelos equipamentos em funcionamento. Cada máquina possui um padrão de vibração considerado "normal" quando está operando corretamente. Desvios desse padrão podem indicar a presença de falhas ou desgastes em componentes, como rolamentos, engrenagens, eixos e até mesmo problemas estruturais no próprio forno ou em seus sistemas de apoio e movimentação.

Recentemente, a área de manutenção preditiva da planta industrial de uma indústria metalúrgica passou por uma digitalização de processos e hoje conta com o uso de sensores em grande parte dos ativos de produção, ou seja, um grande volume de dados é coletado diariamente e há uma oportunidade de planejamento mais eficiente das intervenções de manutenção. Mesmo já existindo sistemas de monitoramento contínuo de máquinas a antecipação de falhas, ou seja, a manutenção preditiva considerando a condição das máquinas é um dos objetivos. A partir disso, vislumbra-se também um melhor suporte aos operadores dos fornos, a partir de *dashboards* mais informativos que apoiem a tomada de decisão de maneira mais facilitada, como também a possibilidade de desligamento automático de equipamentos de forma preventiva. Outro ponto de destaque é a questão da sustentabilidade industrial, uma vez que os fornos são equipamentos importantes na cadeia de produção e suas ações de parada para reparos não programados dispense considerável quantidade de recursos, mobilização de pessoal especializado, além de comprometer o cronograma de produção.

Considerando todas estas questões, foi disponibilizado um *dataset* com dados dos sensores relacionados a fornos específicos. Para este estudo a indústria metalúrgica forneceu dados da ordem de 1.000.000 de registros de três máquinas diferentes contendo informações

de sete sensores, que indicam modos de falha específicos, instalados em fornos de tratamento térmico de peças metálicas, sendo: vibração global, colisão, desbalanceamento, rede, rolamento 1, rolamento 2, rolamento 3. O *dataset* representa um intervalo operacional de 8 dias. O conjunto de sensores é utilizado para monitoramento de vibração do equipamento durante seu uso. Para cada sensor a empresa estabelece, a partir do conhecimento de seus especialistas, faixas de valores para cada sensor, que representam intervalos considerados aceitáveis ou normais, como mostra a Tabela 7:

Tabela 7 – Faixa considerada normal para dados capturados por sensores.

Sensor	Valor mínimo	Valor máximo
Vibração global (S1)	0	2.5
Colisão (S2)	0	5000
Desbalanceamento (S3)	0	2.5
Rede (S4)	0	5
Rolamento 1 (S5)	0	50
Rolamento 2 (S6)	0	60
Rolamento 3 (S7)	0	360

Fonte: Elaborada pelo autor (2023).

No primeiro ciclo da DSR, a DT-Create suportava o uso de dados previamente rotulados com classes que permitiam o aprendizado supervisionado, de forma que um modelo fosse treinado para rotular automaticamente novos dados. No entanto, os dados utilizados neste Estudo de Caso 2 não apresentavam informações de classes que permitissem ao algoritmo de aprendizado prever a que classe um dado está associado.

Com o novo componente Auto ML da DT-Create, outros tipos de algoritmos de aprendizagem de máquina podem ser considerados e selecionados. No caso específico dos dados indústria metalúrgica, algoritmos não supervisionados foram utilizados com o objetivo de detectar padrões nos dados e, considerando as características do *dataset* disponibilizado, não foi possível extrair informações que favorecessem a tomada decisão nas atividades de manutenção preditiva. Dessa forma, antes de iniciar as tarefas de aprendizagem de máquina, foi realizado um procedimento para estimar a condição de falha da máquina a partir das condições de falha de sensor, um pré-processamento especial, no caso do *dataset* desta indústria metalúrgica. Dessa forma, os dados disponíveis poderiam ser classificados e então um modelo preditivo poderia ser treinado para classificação de novos dados.

A partir dos dados disponíveis e com o uso de procedimentos estatísticos, foi possível estimar se o valor observado pelo sensor indicava a ocorrência de um modo de

falha (Bussab e Morettin, 2010). Para tal, seguiu-se o seguinte fluxo de trabalho para rotulagem de dados:

1. Calcular o Intervalo de Confiança (IC) de cada observação de cada sensor;
2. Estimar a condição de falha de cada sensor pela comparação do IC calculado com os valores de referência da Tabela 7;
3. Determinar o valor esperado da saída de cada tupla de sensores;
4. Comparar o valor esperado com o limite gerencial de manutenção.

Levamos em conta o tamanho amostral e desconsiderando o efeito de *outliers* assume-se a normalidade dos dados (Fischer, 2011), que não gera perda de generalização. Segundo (Moore, 1996), as distribuições normais são boas aproximações dos resultados de muitos tipos de resultados aleatórios, sendo aplicada a uma ampla gama de fenômenos no mundo real, incluindo medições de erro em processos de manufatura. Também muitos procedimentos de inferência estatística baseados em distribuições normais funcionam bem para outras distribuições aproximadamente simétricas. Portanto, adotamos a análise do intervalo de confiança para estimar, com nível de confiança de 95%, a classificação de cada saída de cada sensor como condição de falha (1) ou condição normal (0), calculado da seguinte forma:

$$IC = x_i \pm z \frac{s}{\sqrt{n}} \quad (1)$$

Onde:

- $x_i$  é a média da amostra;
- $z$  é o valor crítico da distribuição normal padrão correspondente ao nível de confiança desejado (por exemplo, 1,96 para um IC de 95%);
- $s$  é o desvio padrão da amostra;
- $n$  é o tamanho da amostra.

Uma vez conhecidos os intervalos de confiança para cada observação de cada sensor, foi possível verificar se o intervalo de confiança estimado estava contido no intervalo definido pela Tabela 7, produzindo o rótulo 0 que indica condição normal. Caso contrário a observação recebe o rótulo 1, que indica condição de falha. A Tabela 8 mostra um conjunto de sensores rotulados gerado por uma dada observação. Em destaque estão as observações que estão fora do intervalo normal de operação (sensores S4 e S7).

Tabela 8 – Rótulos definidos após a verificação de intervalos de confiança e parâmetros

Sensor	Valor observado	Rótulo
Vibração global (S1)	0.093809	0
Colisão (S2)	8.133628	0
Desbalanceamento (S3)	0.050851	0
<b>Rede (S4)</b>	<b>0.069705</b>	<b>1</b>
Rolamento 1 (S5)	4.365503	0
Rolamento 2 (S6)	5.017564	0
<b>Rolamento 3 (S7)</b>	<b>49.76829</b>	<b>1</b>

Fonte: Elaborada pelo autor (2023).

Após essa etapa temos o conjunto de valores observados e o respectivo conjunto de rótulos. A partir do conhecimento operacional sobre a relevância de cada modo de falha (sensores) sobre a condição de falha da máquina, foram atribuídos pesos para cada um dos sensores como mostra a Tabela 9, satisfazendo:

$$\sum_{P_i} = 1; P_i \geq 0 \quad (2)$$

onde  $P_i$  é a relevância do sensor para a definição da condição de falha da máquina, com valores estimados no intervalo  $[0,1]$ .

Tabela 9 – Relevância de cada sensor na determinação da condição de uma máquina.

Sensor	Peso
Vibração global (S1)	0.10
Colisão (S2)	0.03
Desbalanceamento (S3)	0.10
Rede (S4)	0.02
Rolamento 1 (S5)	0.25
Rolamento 2 (S6)	0.30
Rolamento 3 (S7)	0.20

Fonte: Elaborada pelo autor (2023).

Combinando a tupla de rótulos de cada sensor com os respectivos pesos foi possível calcular o valor esperado para a condição de falha da máquina, exemplificado na Tabela

10. O valor esperado é obtido por:

$$E = \sum_{Si.Pi} \quad (3)$$

onde  $Si$  representa o rótulo (0 ou 1) associado um sensor e  $Pi$  o peso associado ao sensor.

Tabela 10 – Combinação entre peso rótulo para estimar a relevância.

Sensor	Peso	Rótulo	Peso x Rótulo
Vibração global (S1)	0.10	0	0
Colisão (S2)	0.03	0	0
Desbalanceamento (S3)	0.10	0	0
Rede (S4)	0.02	1	0.02
Rolamento 1 (S5)	0.25	0	0
Rolamento 2 (S6)	0.30	0	0
Rolamento 3 (S7)	0.20	1	0.20

Fonte: Elaborada pelo autor (2023).

A amostra apresentada na Tabela 10 destaca um valor esperado igual a 0,22. O valor esperado pode ser usado como referência para tomada de decisão sobre intervenções de manutenção de acordo com a máquina em si ou da demanda de produção, a partir de cenários. Por exemplo: considerando uma máquina com elevado histórico de manutenções, ou seja, uma máquina mais sensível à falhas, o valor esperado pode ser ajustado para próximo de zero, ou seja, na ocorrência valores esperados próximos de zero a condição de falha já é sinalizada. Por outro lado, uma máquina recentemente adquirida e com baixo índice de manutenções pode ser operada em condições menos conservadoras, ou seja, a sinalização de alerta considera modos de falha com maior valor ponderado. A Tabela 11 mostra três exemplos de cenários para três máquinas diferentes.

Tabela 11 – Cenários de ajuste de valores esperados.

Condição de manutenção	Estilo de gerenciamento	Valor esperado (VE)
Alta	Conservador	$0 < VE < 0.3$
Média	Moderado	$0.3 \leq VE < 0.6$
Baixa	Agressivo	$VE \geq 0.6$

Fonte: Elaborada pelo autor (2023).

Assim, executadas estas tarefas de pré-processamento de dados, o *dataset* de sensores foi enriquecido, o que permitiu inferir a condição de falhas de máquinas, como mostra o exemplo da Tabela 12. Se considerados os parâmetros gerenciais da Tabela 11 e, a partir dos dados da Tabela 12, a condição de falha da máquina indica a necessidade de intervenção no cenário conservador.

Tabela 12 – Estimativa de condição de máquina a partir de dados de sensores.

S1	S2	S3	S4	S5	S6	S7	Condição de falha da máquina
0.09	8.13	0.05	0.07	4.37	5.01	49.7	0,2

Fonte: Elaborada pelo autor (2023).

Assim, os dados dos sensores foram enriquecidos com rótulos 0 ou 1 indicando, de acordo com os parâmetros de estilo de gerenciamento mostrados na Tabela 11 a necessidade de intervenção para uma determinada máquina. A Figura 23 mostra como ficou o *dataset* após o pré-processamento estatístico, com uma nova coluna acrescida à direita, nomeada como "label", apresentando valores 1 ou 0, indicando a necessidade intervenção (1) ou não (0).

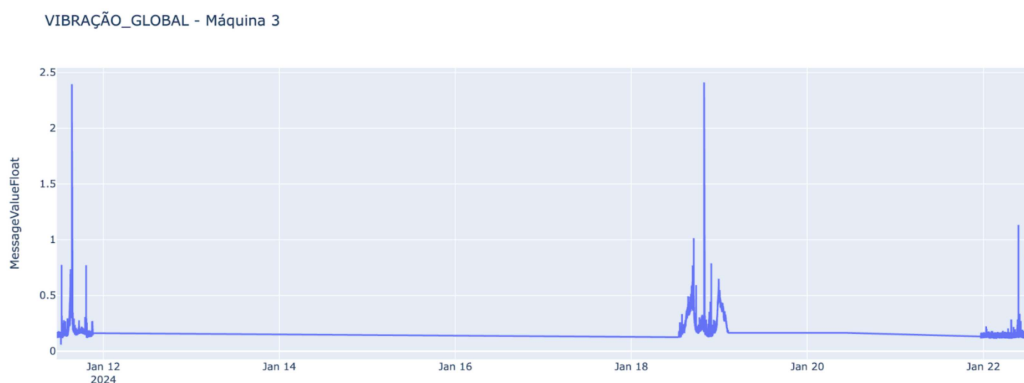
Figura 23 - Adição de coluna ao *dataset* após processamento estatístico.

	2F03_COLISÃO_S3	2F03_VIBRAÇÃO_GLOBAL_S3	2F03_DESBALANCEAMENTO_S3	label
<b>63023</b>	15.892150	0.245994	0.065820	1
<b>63271</b>	6.978260	0.239681	0.065876	1
<b>248179</b>	11.011000	0.352170	0.057822	1
<b>248180</b>	10.790450	0.352170	0.057822	1
<b>248181</b>	15.189050	0.352170	0.057822	1
...	...	...	...	...
<b>239509</b>	11.859380	0.150089	0.061494	0
<b>239510</b>	10.462910	0.150089	0.061494	0
<b>239511</b>	17.304460	0.150089	0.061494	0
<b>239512</b>	13.316650	0.150089	0.061494	0
<b>313731</b>	9.948064	0.136900	0.059801	0

Fonte: Elaborada pelo autor (2023).

Com o *dataset* devidamente rotulado e antes da aprendizagem de máquina foi possível perceber que o *dataset* apresenta um grande desbalanceamento entre classes, ou seja, há uma grande disparidade entre dados com rótulo 0 e dados com rótulo 1. A Figura 24 mostra que no intervalo de tempo amostral a classe 0 é predominante sobre a classe 1 (anormalidade).

Figura 24 - Amostra de dados de um sensor.



Fonte: Elaborada pelo autor (2023).

O desbalanceamento de classes é uma questão comum em problemas de classificação considerando a aprendizagem de máquina, ocorrendo quando as classes alvo (as categorias que o modelo está tentando prever) estão representadas de forma desigual nos dados. Isto significa que uma ou mais classes têm um número muito maior de instâncias (exemplos) do que outras no *dataset*.

O desequilíbrio entre as classes pode levar a vários problemas durante o treinamento e a avaliação de modelos de aprendizagem de máquina, influenciando negativamente o desempenho do modelo, especialmente na sua capacidade de prever as classes menos representadas. No caso específico do *dataset* utilizado neste estudo, a classe desbalanceada é justamente a de maior interesse (He and Garcia, 2009). Para isso, adotamos a subamostragem da classe majoritária (*under-sampling*) que reduz o número de instâncias na classe majoritária para equilibrar a distribuição das classes (Chawla et al., 2002).

Com o *dataset* pré-processado, foi possível submetê-lo a suíte de serviços DT-Create, para a seleção do modelo de aprendizado e posterior especificação do DT utilizado neste segundo estudo de caso. Esta seleção automática elegeu um dentre um conjunto de modelos de aprendizagem de máquina disponíveis a partir do desempenho de cada um deles sobre o *dataset*. Esta seleção automática é realizada na DT-Create a partir da biblioteca *PyCaret*.

*PyCaret* é uma biblioteca de aprendizado de máquina em Python que automatiza o fluxo de trabalho de ML. Ela oferece tarefas de pré-processamento de dados, seleção e treinamento de modelos, otimização de hiperparâmetros, análise e comparação de modelos, e a preparação do modelo para o ambiente de produção (Ali, 2020).

Ao fim do processo de aprendizagem de máquina espera-se ter como saída um modelo ajustado para predição a partir da entrada de novos dados de condição de máquinas. O processo iniciou com a divisão do *dataset* em dados de treino e teste.

```

1 from pycaret.classification import *
2 test=data.sample(frac=0.10, random_state=1)
3 train=data.drop(test.index)

```



Uma vez organizado o *dataset* em dois subconjuntos (treino e teste), a criação de modelo é parametrizada e já na parametrização o conjunto de dados é tratado em relação à remoção de *outliers*, normalização de dados e seleção das *features* mais relevantes.

```

1 reg = setup(data = train, target = 'label', remove_outliers= True
2           , normalize = True, feature_selection = True, session_id=1)
  best = compare_models()

```

Em seguida, a comparação entre modelos foi processada e o resultado da comparação é mostrado na Figura 25. A comparação entre modelos é uma função que avalia a performance tanto em precisão preditiva quanto em custo computacional e como saída é gerada uma lista ordenada de modelos de classificação.

Figura 25 - Tabela de comparação entre modelos de aprendizagem de máquina.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>dt</b>	Decision Tree Classifier	0.9526	0.9898	0.9514	0.9813	0.9659	0.8883	0.8902	0.6240
<b>et</b>	Extra Trees Classifier	0.9526	0.9898	0.9514	0.9813	0.9659	0.8883	0.8902	0.7740
<b>lightgbm</b>	Light Gradient Boosting Machine	0.9480	0.9919	0.9460	0.9802	0.9625	0.8782	0.8808	0.8290
<b>rf</b>	Random Forest Classifier	0.9473	0.9908	0.9481	0.9770	0.9620	0.8757	0.8781	0.8340
<b>knn</b>	K Neighbors Classifier	0.9427	0.9854	0.9481	0.9707	0.9588	0.8646	0.8672	0.4450
<b>gbc</b>	Gradient Boosting Classifier	0.9404	0.9894	0.9416	0.9734	0.9570	0.8598	0.8620	0.5220
<b>xgboost</b>	Extreme Gradient Boosting	0.9373	0.9883	0.9286	0.9821	0.9542	0.8551	0.8600	0.4570
<b>ada</b>	Ada Boost Classifier	0.9045	0.9624	0.9112	0.9522	0.9309	0.7763	0.7796	0.7010
<b>nb</b>	Naive Bayes	0.6792	0.6946	0.7489	0.7882	0.7673	0.2492	0.2515	0.4160
<b>qda</b>	Quadratic Discriminant Analysis	0.6792	0.6946	0.7489	0.7882	0.7673	0.2492	0.2515	0.4100
<b>svm</b>	SVM - Linear Kernel	0.6646	0.0000	0.7620	0.7737	0.7532	0.1856	0.2052	0.4060
<b>ridge</b>	Ridge Classifier	0.6158	0.0000	0.6504	0.7710	0.7048	0.1634	0.1695	0.4060
<b>lda</b>	Linear Discriminant Analysis	0.6158	0.6558	0.6504	0.7710	0.7048	0.1634	0.1695	0.4050
<b>lr</b>	Logistic Regression	0.6127	0.6558	0.6461	0.7699	0.7018	0.1590	0.1653	0.5910
<b>dummy</b>	Dummy Classifier	0.2941	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4100

Fonte: Elaborada pelo autor (2023).

Após a comparação entre modelos o modelo mais adequado foi a Árvore de Decisão, ainda que próximo aos demais modelos ele se destaca pelas métricas de acurácia, recall, F1-Score, Kapp e MCC. A acurácia avalia o desempenho geral de um modelo de classificação mostrando a proporção de previsões corretas (tanto positivas quanto negativas) em relação ao total de previsões feitas. O *recall*, também conhecido como sensibilidade ou taxa de verdadeiros positivos, avalia o desempenho de um modelo de classificação, especialmente em contextos onde as classes são desbalanceadas ou quando a detecção de classes minoritárias é particularmente importante, medindo a proporção de instâncias positivas reais que foram corretamente identificadas pelo modelo, ou seja, a fração de verdadeiros positivos em relação ao total de instâncias que são efetivamente positivas (a soma de verdadeiros positivos e

falsos negativos). O *F1-Score* avalia o desempenho de modelos, especialmente em situações onde existem classes desbalanceadas ou quando a precisão e o recall são igualmente importantes. Ele combina precisão e *recall* em uma única métrica através de sua média harmônica, proporcionando um balanço entre essas duas medidas. O coeficiente *Kappa* é uma métrica utilizada para medir a concordância entre duas avaliações, quantificando quanto a classificação feita por um modelo de aprendizagem de máquina concorda com a classificação verdadeira (ou com outra avaliação) além do que seria esperado por acaso. O Coeficiente de Correlação de *Matthews* (MCC), avalia a qualidade de modelos binários de classificação. Pode ser usada mesmo quando as classes são de tamanhos muito diferentes, tornando-se particularmente útil em situações de conjuntos de dados desbalanceados.

Encontrado o modelo de classificação mais adequado, o próximo passo foi definir o modelo de classificação e fazer o ajuste dos hiperparâmetros. O parâmetro *fold* especifica o número de subconjuntos (*folds*) em que o conjunto de dados será dividido durante o processo de validação cruzada.

```

1 dt = create_model('dt', fold=3)
2 tuned_dt = tune_model(dt, optimize="Accuracy")
3 tuned_model_custom = tune_model(dt)
4 predict_model(dt)

```

Após a criação o modelo foi ajustado pela função `tune_model`. Aqui o ajuste é feito de forma a aumentar sua capacidade de generalização. Em seguida o modelo ajustado (`dt`) foi definido como o modelo de predição (`predict_model`). Como saída temos o modelo de predição ajustado, como mostra a Figura 26. O modelo ajustado aparentemente apresenta desempenho inferior ao modelo escolhido inicialmente. Isto ocorre porque o ajuste reduz o *overfitting* do modelo, que significa que o modelo tem ótimo desempenho com dados de treinamento, mas não tem o mesmo desempenho ao prever resultados em dados não vistos ou de teste. Portanto o modelo após a ajuste mesmo sendo menos preciso se comparado ao modelo inicial é mais preciso na predição sobre novos dados.

Figura 26 - Modelo de aprendizagem de máquina ajustado.

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	Decision Tree Classifier	0.9199	0.9776	0.9093	0.9757	0.9413	0.8158	0.8208

Fonte: Elaborada pelo autor (2023).

A partir da escolha do modelo de aprendizagem de máquina mais adequado, feito a partir da DT-Create, o modelo foi utilizado na especificação do DT de suporte a tomada de decisão na indústria metalúrgica, com o objetivo de apoiar os procedimentos de manutenção preditiva.

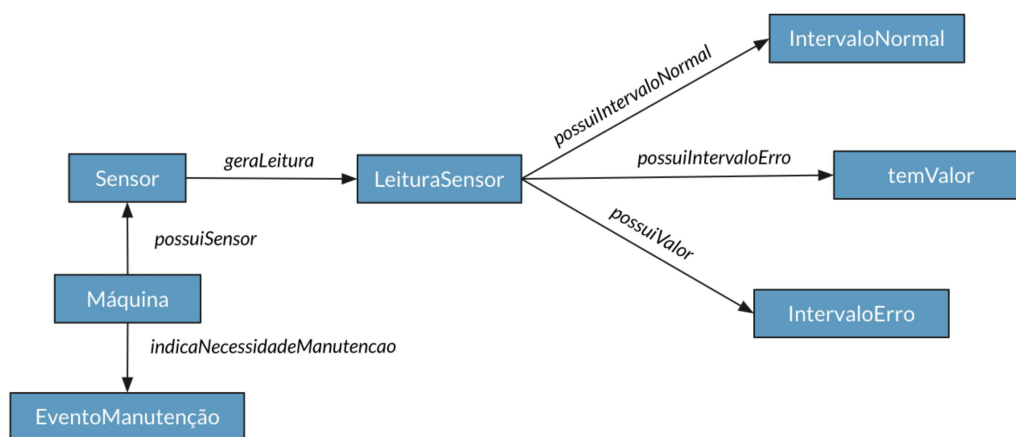
Para o processamento semântico, ontologia foi criada especificamente para o domínio de aplicação do CS2, a partir das questões de competência.

### 5.5.2 ONTOLOGIA SENSOREQUIPAMENT

A ontologia, denominada *SensorEquipament*, utilizada neste estudo de caso, é direcionada à manutenção preditiva de máquinas com base na análise de condições de operação indicadas por sensores de forma mais genérica. Nesse sentido, é uma ontologia mais genérica do que a utilizada no Estudo de Caso 1.

Para sua especificação, Questões de Competência foram elaboradas, com a ajuda de especialistas: (i) Quais máquinas estão atualmente indicando uma condição de erro? (ii) Quais sensores indicaram condições de erro em uma máquina específica? (iii) Existem padrões de erro que ocorrem simultaneamente em várias máquinas? (iv) Quais foram as últimas leituras de cada sensor em uma máquina específica antes de um evento de manutenção? (v) Quantos eventos de manutenção foram iniciados devido a condições de erro nos últimos seis meses? (vi) Quais máquinas não apresentaram nenhum indicativo de erro nos últimos 12 meses? (vii) Qual é a frequência média de erros para cada tipo de sensor em diferentes máquinas?

Figura 27 - Ontologia para representação da condição operacional de máquinas.



Fonte: Elaborada pelo autor (2023).

A Figura 27 apresenta as principais classes da ontologia. O domínio é formado por máquinas e informações de sensores que representam a condição de operação da máquina, possuindo as seguintes classes:

- **Máquina:** Representa as máquinas (genericamente).
- **LeituraSensor:** Valores retornados pelos sensores.

- **IntervaloNormal**: Faixas de valores considerados normais para cada tipo de leitura de sensor.
- **IntervaloErro**: Faixas de valores que indicam uma condição de erro.
- **EventoManutencao**: Eventos de manutenção que são planejados e realizados com base nas análises das leituras dos sensores.

A partir das informações das classes que compõem a ontologia foram descritas as propriedades que mostram os relacionamentos entre as classes indicadas na Figura 27:

- **possuiSensor** (Maquina -> Sensor): Indica que uma máquina possui um ou mais sensores.
- **geraLeitura** (Sensor -> LeituraSensor): Relaciona um sensor às suas leituras.
- **temIntervaloNormal** (Sensor -> IntervaloNormal): Associa um sensor ao seu intervalo de operação normal.
- **temIntervaloErro** (Sensor -> IntervaloErro): Associa um sensor ao seu intervalo que indica condição de erro.
- **temValor** (LeituraSensor -> xsd:int): O valor numérico de uma leitura de sensor, podendo ser 0 (normal) ou 1 (erro).
- **indicaNecessidadeManutencao** (Maquina -> EventoManutencao): Indica que uma máquina requer manutenção.

A partir das propriedades já definidas na ontologia e dos dados pre-processados pelo modelo de aprendizagem de máquina, os mesmos foram instanciados na ontologia e o processamento de regras SWRL permitirá o enriquecimento das informações. As seguintes regras SWRL foram especificadas para descoberta de *insights* em relação às oportunidades de antecipação de ações de manutenção a partir de dados dos sensores. De forma ilustrativa, para inferir a necessidade de manutenção a partir dos dados dos sensores foram processadas as regras que analisam as leituras dos sensores rotulados como 0 (valor medido dentro do intervalo normal de operação) ou 1 (fora do intervalo de operação normal), como mostra a Tabela 13:

Tabela 13 – Avaliando a condição operacional de máquinas por regras SWRL.

#	Regras SWRL	Descrição
01	<p><b>Antecedente</b>  Sensor(?sensor) ^  geraLeitura(?sensor, ?leitura) ^  temValor(?leitura, 1)</p> <p><b>Consequente</b>  -&gt; indicaNecessidadeManutencao(?maquina)</p>	Se um sensor gera uma leitura cujo valor é 1 (indicativo de erro), então a máquina associada a esse sensor requer manutenção.
02	<p><b>Antecedente</b>  Sensor(?sensor1) ^  geraLeitura(?sensor1, ?leitura1) ^  temValor(?leitura1, 1) ^  Sensor(?sensor3) ^  geraLeitura(?sensor3, ?leitura3) ^  temValor(?leitura3, 1) ^  differentFrom(?sensor1, ?sensor3)</p> <p><b>Consequente</b>  -&gt; indicaNecessidadeManutencao(?maquina)</p>	Caso o conjunto dos rolamentos que equipam uma máquina, monitorados pelos sensores S5, S6 e S7, apresentem sinalização de falha simultaneamente um alerta é gerado.

Fonte: Elaborada pelo autor (2023).

A regra 01 da Tabela 13 considera a ocorrência de uma falha associada ao sensor S1 (vibração global). Nesse caso, o sensor S1 é importante no contexto funcional da máquina e precisa ser monitorado. Já a regra 02 da Tabela 13 avalia a ocorrência de falhas em conjunto dos rolamentos que equipam uma máquina, representados pelos sensores S5, S6 e S7. Especificamente, os três modos de falhas simultâneos sinalizam a necessidade de avaliação pela equipe de manutenção.

Assim, a partir de uma consulta SPARQL (detalhada abaixo) foi possível recuperar a quantidade de eventos de falha indicado por um grupo de sensores.

```

1 SELECT ?sensor (COUNT(?evento) AS ?quantidadeFalhas)
2 WHERE {
3   ?sensor ex:IntervaloErro ?evento .
4   FILTER (?sensor IN (S5, S6, S7))
5 }
6 GROUP BY ?sensor

```

Dessa forma, o processamento ontológico no CS2 permitiu extrair conhecimento a partir do processamento semântico e da verificação da ocorrência conjunta de eventos que são importantes do ponto de vista da manutenção preditiva e que não faziam parte

do *dataset* original e que também não foram determinados durante o processamento no componente Auto ML.

### 5.5.3 ANÁLISE

Considerando a condução do estudo de caso é possível analisar os resultados alcançados. Com o modelo ajustado e pronto para uso foi possível verificar seu desempenho. A Figura 28 mostra o modelo preditivo sendo processado considerando dados novos, oriundos de sensores.

Figura 28 - Processamento do modelo preditivo com um novo conjunto de dados.

	2F03_COLISÃO_S3	2F03_VIBRAÇÃO_GLOBAL_S3	...	2F03_DESBALANCEAMENTO_S3	label	prediction_label	prediction_score
726	8.522325	0.241362	...	0.066188	1	1	0.7298
348	7.289566	0.129020	...	0.060549	0	0	0.6406
102	11.989620	0.135731	...	0.058455	1	1	0.9944
1272	12.550910	0.133094	...	0.061983	1	1	0.9985
1994	17.922621	0.388702	...	0.060862	1	1	0.9979
...	...	...	...	...	...	...	...
1104	13.603150	0.146172	...	0.061615	0	0	0.9555
1179	11.143620	0.135574	...	0.062383	1	1	0.9983
1654	9.865036	0.136189	...	0.062684	1	1	0.9722
368	10.959950	0.142156	...	0.060549	0	0	0.6406
859	21.858290	0.253324	...	0.065797	1	1	0.9901

Fonte: Elaborada pelo autor (2023).

Como observado na Figura 26 o modelo apresenta um desempenho excepcionalmente alto em todas as métricas avaliadas, indicado pelas altas pontuações em AUC, Kappa, e MCC. No entanto, o conjunto de dados utilizado para a criação do modelo é relativamente pequeno em relação ao período de tempo amostral. Espera-se que os dados de entrada sejam equilibrados na sinalização dos eventos de falha e não falha de máquinas e essa proporção não foi equilibrada, o que nos levou a aplicar subamostragem da classe majoritária (*undersampling*). Assim, consideramos que para utilização do modelo preditivo em produção é necessário o treinamento com uma base de dados maior e que seja equilibrada nas duas classes de interesse.

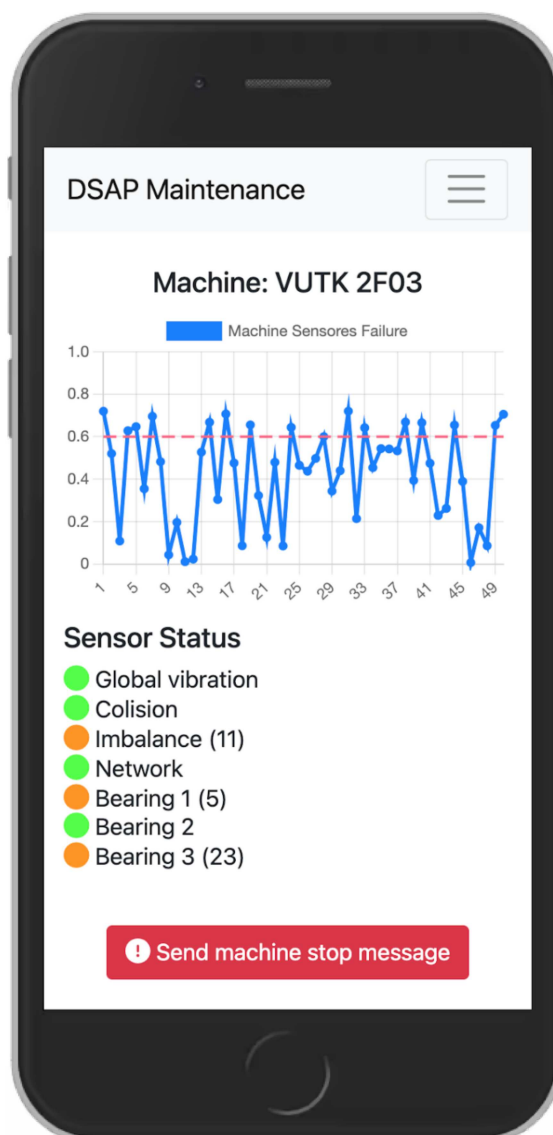
Considerando este segundo ciclo de DSR o foco foi na oportunidade de melhoria do modelo de aprendizagem de máquina, visto que no primeiro ciclo adotamos um único modelo como referência. Portanto, buscamos definir de forma automática o modelo de aprendizagem de máquina mais adequado ao conjunto de dados de interesse e os resultados obtidos foram positivos, uma vez que foi possível estabelecer um modelo de aprendizagem de máquina ajustado para o conjunto de dados e capaz de gerar previsões. Esse modelo foi selecionado de forma automática e ajustado de forma a ter um desempenho adequado para o conjunto de dados de treinamento.

Também foi possível perceber que para um conjunto de dados novo o modelo foi capaz de fazer previsões com elevada precisão. No entanto, destacamos que o conjunto

de dados inicial não estava adequadamente preparado para o processo de aprendizagem de máquina e foi dedicado esforço para o pré-processamento dos dados e também do processamento estatístico dos dados para inferência de classes de forma a favorecer o aprendizado de máquina. Da mesma forma, o volume de dados pode ser incrementado de forma a apresentar um maior volume de eventos da classe alvo para melhoria do processo de aprendizagem de máquina.

A partir do processamento ontológico e regras específicas, foi possível monitorar a condição das máquinas através de gráficos a partir de um aplicativo mobile. No aplicativo, os gestores de manutenção puderam acompanhar a evolução da condição das máquinas em intervalo de horas, receber notificações sobre falhas críticas de cada um dos sensores e também o estado do conjunto como um todo através do gráfico principal. A Figura 29 apresenta o *dashboard* do Aplicativo Móvel, gerado a partir da suíte de serviços DT-Create.

Figura 29 - Dashboard da Aplicação Mobile.



Fonte: Elaborada pelo autor (2023).

Na Figura 29 é possível observar que o cenário de operação está configurado como moderado, representado pela linha tracejada em cor vermelha. Nesta configuração falhas do conjunto são toleradas no intervalo  $[0, 0.6]$ . Em alguns momentos os sensores reportam valores acima do ideal e o *dashboard* acumula esses eventos dentro da janela de observação, sinalizando quais sensores apresentaram sinal de falha. Dessa forma, dada a condição geral da máquina e a condição de cada um dos sensores é possível enviar um sinal de parada de máquina em casos críticos. As oscilações do gráfico representam as sinalizações de falhas dos sensores individualmente, que em conjunto representam a condição atual da máquina.

## 5.6 RESULTADOS

Como resultado final, avaliando os dois ciclos de avaliação realizados, ou seja, Estudo de Caso 1 e Estudo de Caso 2, podemos considerar que os resultados foram satisfatórios, com algumas ressalvas, relatadas nesta seção. A questão de pesquisa: “Como apoiar a tomada de decisão relacionada a manutenção preditiva na Indústria 5.0, com o uso de DTs?”, pode ser respondida, considerando que em ambos os estudos de caso, um DT foi desenvolvido para apoiar problemas relacionados a manutenção preditiva em diferentes cenários da Indústria 5.0. O fator humano foi considerado, a partir do desenvolvimento de *dashboards* amigáveis, com informações geradas a partir de processamento inteligente, cujo objetivo foi facilitar a tomada de decisão. A questão da sustentabilidade também foi avaliada, considerando que as medidas tomadas a partir das recomendações geradas pelo processamento inteligente permitem uma vida útil mais longa para os equipamentos.

De maneira específica, podemos destacar que o novo componente Auto ML avaliado no segundo estudo de caso, gerou um modelo preditivo capaz de fazer previsões com elevada precisão para um *dataset* desconhecido (*dataset* de teste). No entanto, destacamos que o *dataset* inicial não estava adequadamente preparado para o processo de aprendizagem de máquina e foi dedicado esforço para o pré-processamento dos dados a partir do uso de procedimentos estatísticos sobre os dados o que permitiu a especificação de classes de forma a favorecer o aprendizado de máquina. Desta forma, pudemos verificar que mesmo com a seleção automática de modelos de aprendizagem, tempo considerável foi dedicado ao pré-processamento, o que reforça que não é trivial utilizar novo conjunto de dados, mesmo com a possibilidade de seleção de técnicas de aprendizagem. Da mesma forma, destacamos a necessidade de um grande volume de dados de forma a apresentar um conjunto maior de eventos da classe alvo, para melhoria do processo de aprendizagem de máquina.

O componente *Ontology Model*, uma vez instanciado, demonstrou assertividade na representação das informações concernentes aos eventos de falha indicados pelos sensores instalados nas máquinas. Ademais, este componente foi capaz de extrair conhecimento a partir dos dados fornecidos pelo componente de Auto ML, enriquecendo os dados por meio de processamento semântico adequado.



Desta forma, a solução proposta no segundo ciclo DSR, orientada a dados das condições operacionais de máquinas, processamento de dados por meio de modelos preditivos automatizados e modelos semânticos, foi capaz de avaliar conjuntamente a criticidade de falhas apontadas por sensores associadas a processamentos semânticos e as características qualitativas e quantitativas das falhas enriquecendo a informação. Também produziu dados que permitem a ação do agente autônomo por meio da exibição de mensagens de alerta em um *dashboard* aos operadores das máquinas e gestores. Dessa forma, foi possível indicar que o uso combinado de técnicas inteligentes e de aprendizagem de máquina, permitiu que agentes de *software* autônomos pudessem detectar falhas de equipamentos e necessidades de manutenção.

## 5.7 TRIANGULAÇÃO DOS RESULTADOS

Os dois estudos de caso apresentaram contribuições distintas para o aprimoramento do artefato. No primeiro ciclo DSR o uso de ontologias trouxe enriquecimento dos dados disponíveis através do processamento de inferência de ontologias e dados de contexto. No segundo ciclo DRS avançamos no processo de aprendizagem de máquina, criando uma estrutura onde é possível automatizar via código de programação as etapas de seleção, criação, ajuste de distribuição do modelo de aprendizagem de máquina. Este incremento é importante para o auto-ajuste do modelo de predição, uma vez que podemos ter um agente atuando como um serviço que pode calibrar o modelo em intervalos de tempo.

A partir dos resultados do estudo de caso 2, confrontamos as conjecturas teóricas sobre o uso de aprendizado de máquina e ontologia levantadas após o estudo de caso 1 e analisamos se, após essa interação, o artefato pode apoiar a tomada de decisão com o processo de aprendizagem de máquina mais completo e preciso.

O ambiente industrial onde o processo de manutenção acontece é complexo e multifacetado, com muitas variáveis e fatores que podem afetar a qualidade da manutenção preditiva. Essa complexidade pode tornar desafiador o desenvolvimento de modelos de aprendizado de máquina que podem capturar com precisão os relacionamentos e padrões nos dados e fazer previsões ou recomendações confiáveis.

Além disso, modelos de aprendizado de máquina para manutenção preditiva podem ser necessários para lidar com uma grande variedade de tipos de dados e fontes, como dados de sensores, registros e outros tipos de dados, o que pode aumentar a complexidade do processo de desenvolvimento do modelo. Para mitigar essa dificuldade, usamos ontologia. No geral, a combinação de ontologia e aprendizado de máquina pode melhorar significativamente os *insights* obtidos com sistemas de manutenção preditiva e permitir uma tomada de decisão mais orientada por dados nas operações de manutenção.

Ainda sobre o primeiro ciclo DSR o uso de redes neurais foi selecionado como referência de modelo de aprendizagem sobre os dados, mas não havia uma seleção de

modelos de forma automatizada que considerasse a aderência do modelo ao conjunto de dados em questão. Avançamos no segundo ciclo DRS, a partir do uso de uma biblioteca de aprendizagem de máquina para seleção automática de um modelo de aprendizagem de máquina. O artefato também foi melhorado no pré-processamento de informações, como uso de processos automáticos de remoção de outliers, normalização de dados, seleção de recursos mais relevantes e testes de qualidade do modelo.

Dessa forma, o artefato no segundo ciclo DRS é capaz de oferecer suporte à decisão nos processos de manutenção preditiva, seja por enriquecimento de dados via processamento que combina ontologias e dados contexto, seja por processamento estatístico e auto-seleção de modelos de aprendizagem.

## 5.8 AMEAÇAS À VALIDADE

Esta seção discute ameaças à validade que podem afetar ou limitar a validade dos resultados. As limitações desta pesquisa estão relacionadas ao nível da qualidade dos dados de entrada. No primeiro ciclo DSR os dados utilizados já possuíam rótulos das classes de interesse (criticidade), o que facilitou o processamento inteligente. No segundo ciclo DSR, os dados foram utilizados sem pré-processamento, o que pode ocorrer em outros cenários de aplicação. A ausência de pré-processamento pode influenciar a interpretação dos resultados e a capacidade de generalizar as conclusões para outros contextos ou sistemas. No contexto específico do segundo estudo de caso, a falta de pré-processamento invalidaria o uso da suíte de serviços DT-Create.

**Validade Interna.** Os modelos podem se tornar excessivamente complexos (sobreajuste ou *overfitting*, capturando ruídos ou padrões aleatórios dos dados de treinamento que não se generalizam de forma adequada para dados não vistos. Isso prejudica a capacidade de inferir causas precisas dos resultados. Se os dados utilizados para treinar e testar o modelo não forem representativos de todas as condições operacionais dos sensores industriais, os resultados podem ser enviesados (viés de seleção de dados). Também inconsistências ou erros nos dados dos sensores podem levar a conclusões errôneas sobre o desempenho do modelo (viés de medição).

**Validade Externa.** A capacidade de generalizar os resultados para outros ambientes industriais ou tipos de sensores pode ser limitada se o conjunto de dados utilizado for muito específico ou não abrangente. Essa limitação é conhecida como generalização. Além disso, o uso de uma ontologia de domínio inadequada pode invalidar os resultados finais. Da mesma forma, as condições sob as quais os dados foram coletados podem não refletir todas as configurações operacionais possíveis, o que limita a aplicabilidade do modelo em diferentes contextos, um problema conhecido como configurações de operação.

**Validade de Construção:** A validade de construção é crucial para garantir que um modelo de aprendizado de máquina represente com precisão o problema que se pretende

solucionar. Essa validação envolve dois aspectos principais:

1. Representatividade: Os modelos devem capturar com precisão as relações entre os dados dos sensores e os fenômenos físicos ou processos industriais subjacentes. Caso contrário, os modelos podem apresentar resultados inconsistentes ou irrelevantes;
2. Adequação das *features*: A seleção e a engenharia *features* adequadas são essenciais para que os modelos aprendam os padrões relevantes nos dados. *Features* mal escolhidas podem levar a modelos que não representam corretamente o problema, resultando em baixa performance e previsões imprecisas.

**Validade de Critério:** A validade preditiva dos modelos, ou seja, sua capacidade de prever resultados futuros, pode ser comprometida em duas situações:

1. Dificuldade em lidar com situações não previstas: O modelos podem ter sido treinados em um conjunto de dados específico e, por isso, apresentar dificuldades em prever resultados em situações que divergem significativamente do contexto de treinamento;
2. Falta de comparação com *benchmarks*: A avaliação da performance dos modelos é incompleta sem compará-lo com outros modelos ou *benchmarks* da indústria. Essa comparação permite contextualizar o desempenho dos modelos propostos e determinar se eles realmente apresentam um avanço significativo em relação às soluções existentes.

## 5.9 CONSIDERAÇÕES FINAIS DO CAPÍTULO

Considerando as dificuldades de tomada de decisão na Indústria 5.0, este trabalho procurou apoiar as decisões no domínio da manutenção preditiva, por meio de uma suíte de serviços. Para isso, a DT-Create foi proposto. a DT-Create especifica DTs capazes de realizar a coleta, processamento, armazenamento e enriquecimento de dados coletados por sensores, sejam instalados em máquinas ou no ambiente operacional. Para verificar a validade dos resultados, dois estudos de caso foram realizados e detalhados nesse capítulo. Evidências foram apresentadas sobre a viabilidade de instanciar a DT-Create para apoiar a tomada de decisões no contexto da indústria têxtil. As decisões neste domínio devem ser tomadas considerando a diversidade de informações e dispositivos em diferentes contextos. Além disso, as informações de contexto muitas vezes não são usadas na tomada de decisões devido à complexidade de gerenciar um alto volume de dados heterogêneos. decisões. O primeiro estudo de caso discutiu a viabilidade de uso da DT-Create para especificar um DT capaz de auxiliar na tomada de decisões na Indústria Têxtil. No segundo estudo de caso, foi utiliza uma nova versão da DT-Create, onde o processamento inteligente foi aprimorado. A condução do estudo de caso permitiu verificar a viabilidade da solução

proposta também no domínio de fornos industriais, permitindo responder a questão de pesquisa proposta.

## 6 CONCLUSÃO

Este trabalho estabelece uma estrutura teórica que consiste em conceitos-chave pertencentes à IoT, modelos semânticos e sistemas inteligentes de apoio à decisão. O trabalho identificou, a partir de um MSL da literatura, os desafios relacionados à utilização de dados de sensores para a tomada de decisões nos processos de manutenção.

Dadas as dificuldades associadas à tomada de decisões de manutenção na Indústria 4.0 e 5.0, o trabalho busca soluções propondo uma abordagem suíte de serviços para especificação de DTs, contribuição para responder à Questão de Pesquisa: “Como apoiar a tomada de decisão relacionada à manutenção preditiva na Indústria 5.0 com o uso de DTs?”.

A suíte de serviços proposta, a DT-Create, engloba um conjunto de componentes para coleta, processamento, armazenamento e enriquecimento de dados obtidos por meio de sensores, que podem ser instalados em máquinas e equipamentos ou ainda no ambiente de produção. Para avaliar a eficácia da arquitetura proposta, dois ciclos DSR foram executados e dois estudos de caso foram conduzidos para gerar evidências sobre a viabilidade da proposta no apoio à tomada de decisões no domínio da manutenção preditiva.

As técnicas de ML podem melhorar o uso do sistema IoT em atividades de manutenção preditiva. Algoritmos de aprendizagem de máquina foram usados para analisar dados coletados por sensores e identificar padrões e tendências que podem não ser imediatamente aparentes para os seres humanos. Esta análise pode ajudar os gestores de manutenção a decidir como gerenciar os ativos de produção e melhorar a eficiência geral e mitigar falhas de alto grau crítico a partir da antecipação de ações de manutenção. Por outro lado, as ontologias foram utilizadas para o processamento semântico das informações geradas pelos modelos de aprendizagem de máquina. No contexto da manutenção preditiva, a ontologia ajudou a organizar e classificar os dados coletados pelos sensores, produzindo alertas sobre cenários críticos que demandam intervenção.

Consideramos que os resultados apresentados foram consistentes, pois abordam demandas cada vez mais importantes, relacionadas à Indústria 4.0 e também da Indústria 5.0. Pretendemos reforçar e priorizar atributos de qualidade, como flexibilidade, extensibilidade e escalabilidade. Novas regras semânticas também podem ser definidas para apoiar o enriquecimento de dados na ontologia e sua integração com outras ontologias específicas de domínio para aumentar a capacidade de extração de conhecimento. Também novas técnicas de pré-processamento dos conjuntos e dados podem ser utilizadas aumentando a capacidade de generalização da DT-Create e sua aplicação em vários domínios diferentes. Finalmente, seria útil realizar novos experimentos em outros subdomínios de manutenção da Indústria 5.0 para avaliar o suporte oferecido pela DT-Create em diferentes subdomínios de aplicações.

Com o objetivo de cumprir a extensibilidade, a DT-Create poderia acomodar novos e diferentes dispositivos de comunicação e fornecer suporte a novos módulos de processamento inteligente, provando ser consistente e escalável. Em relação à flexibilidade, a DT-Create permitiu a seleção automática de modelos de aprendizagem de máquina adequados, exigindo apenas a entrada de dados no padrão esperado para sua operação. Portanto, novos modelos de aprendizagem de máquina disponíveis trariam impacto positivo no funcionamento da DT-Create. Além disso, à medida que foram utilizados novos algoritmos inteligentes e mais recursos, observamos que a solução permaneceu estável e funcional durante todo o processo de treinamento e avaliação dos algoritmos, mostrando seu potencial para atender aos atributos de qualidade estabelecidos.

DT-Create é uma suíte de serviços para apoiar a tomada de decisões em processos de manutenção preditiva na Indústria 5.0. Ao avaliar o artefato em domínios das indústrias, evidenciamos que podemos monitorar a saúde de máquinas e equipamentos. Isso foi possível usando uma solução inteligente capaz de enriquecer as decisões por meio dos serviços oferecidos. A partir desta solução, os gestores de manutenção tomar decisões e mudar estratégias de manutenção.

A partir dos resultados obtidos evidencia-se o potencial de enriquecimento das decisões complexas de manutenção com informações já disponíveis, geradas pelos ativos de produção, mas ainda não processadas por meio de técnicas inteligentes. Essa perspectiva abre portas para pesquisas futuras, especialmente em relação à predição de falhas a partir da análise de equipamentos, possibilitando o agendamento automático de tarefas de manutenção. Adicionalmente, pesquisas sobre a proveniência dos dados se tornam importantes, considerando o contexto dos dados da IoT para enriquecer e aprimorar a tomada de decisões na área de manutenção.

A relevância desta proposta reside na sua capacidade de integrar semântica e IA, oferecendo à comunidade científica uma suíte de serviços robusta para o aprimoramento da tomada de decisões nesse contexto. Também combina inferências semânticas com previsões de aprendizado de máquina para enriquecer dados e otimizar a tomada de decisões em manutenção industrial. Através da visualização de resultados preditivos e informações sobre o estado das máquinas em dispositivos móveis, a solução oferece aos gestores uma ferramenta relevante para aumentar a eficiência e a produtividade das operações.

As decisões na manutenção preditiva precisam ser tomadas considerando a diversidade de informações e dispositivos presentes em diferentes contextos. Além disso, as informações de contexto geralmente não são usadas no processo de tomada de decisão devido à complexidade de gerenciar dados heterogêneos. Este trabalho apresentou uma suíte de serviços que visava resolver os problemas de coleta, processamento e visualização de dados para apoiar a tomada de decisões.

Nossos resultados apresentam relevância por combinarem tecnologias que impul-

sionam o progresso da manutenção preditiva na Indústria 5.0, promovendo o aumento da eficiência, da produtividade e da competitividade das indústrias. Apoiada na coleta e análise de dados em tempo real, a DT-Create subsidia decisões que previnam ou minimizem paralisações de produção decorrentes de falhas críticas em máquinas. O aprimoramento da suíte de serviços proporciona benefícios adicionais, como:

- Aumento da disponibilidade dos equipamentos: A previsão de falhas permite o planejamento de intervenções de manutenção antes que estas ocorram, maximizando a produtividade e a vida útil dos equipamentos;
- Otimização dos recursos em manutenção: A diretividade do uso de mão de obra e peças de reposição para intervenções realmente necessárias gera economia e eficiência;
- Aumento da segurança na operação: A minimização de riscos de acidentes e paradas não planejadas garante um ambiente de trabalho mais seguro para os colaboradores.

Em trabalhos futuros, pretendemos gerar novas instâncias da DT-Create integrá-las a um ecossistema de *software*. Essa iniciativa permitirá explorar aspectos de colaboração, comunicação e integração entre diferentes áreas de negócio, com o objetivo de aprimorar o processo de tomada de decisões. Acreditamos que a combinação e o processamento de fontes de dados e sensores adicionais podem gerar resultados mais precisos, reduzir custos e contribuir para a sustentabilidade da indústria.

## REFERÊNCIAS

- Mobley R. Keith. **An Introduction to Predictive Maintenance**. Elsevier, 2002.
- Hansen, R. C. **Overall Equipment Effectiveness: A Powerful Production/Maintenance Tool for Increased Profits**. Industrial Press Inc., 2001.
- Jardine, Andrew KS e Tsang, Albert HC. **Maintenance, Replacement, and Reliability: Theory and Applications**. CRC Press, 2005.
- Wang K. **Intelligent Predictive Maintenance (IPdM) System–Industry 4.0 Scenario**. WIT Transactions on Engineering Sciences. 113, 2016, 259-268.
- Gomes, J., Esteves, I., Neto, V.V.G., David, J.M.N., Braga, R., Arbex, W., Kassab, M., and de Oliveira, R.F. **A Scientific Software Ecosystem Architecture for the Livestock Domain**. Information and Software Technology, 160, 2023, 107240; Elsevier.
- Esteves, Izaque, Braga, Regina, David, José Maria N., and Stroele, Victor. **A Self-adaptative Architecture to Support Maintenance Decisions in Industry 4.0**. In Proceedings of the International Conference on Advanced Information Networking and Applications, pages 274–285, Springer, 2023.
- Silva, Izaque, Gomes, Jonas, Braga, Regina, David, José Maria N., Stroele, Victor, Soares, Rômulo, Oliveira, Vitória, and De Oliveira, André Luiz. **AI-Based development on software ecosystem platforms**. In Proceedings of the XXXVII Brazilian Symposium on Software Engineering, pages 148–153, 2023.
- Da Silva, Izaque Esteves, Braga, Regina, David, José Maria N., and Stroele, Victor. **Anomalies Detection in records of operational failures using IoT devices and data mining**. In Anais Estendidos do XVII Simpósio Brasileiro de Sistemas de Informação, pages 209–216, SBC, 2021.
- Rosenblatt, Frank. **Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms**. Spartan Books, 1962.
- Runeson, P., Host, M., Rainer, A., Regnell, B. **Case study research in software engineering: Guidelines and examples**. John Wiley & Sons, 2012.
- De Nicola, Antonio and Missikoff, Michele. **A Lightweight Methodology for Rapid Ontology Engineering**. Communications of the ACM, 59(3), 2016, 79–86.
- Feilmayr, Christina and Wöß, Wolfram. **An analysis of ontologies and their success factors for application to business**. Data & Knowledge Engineering, 101, 2016, pages 1–23, Elsevier.
- Mohammed, Mohssen, Khan, Muhammad Badruddin, and Bashier, Eihab Bashier Mohammed. **Machine Learning: Algorithms and Applications**. CRC Press, 2016.
- Jean-Baptiste, Lamy. **Ontologies with Python: Programming OWL 2.0 Ontologies with Python and Owlready2**. Springer, 2021.
- Soloman, Sabrie. **Sensors Handbook**. McGraw-Hill, Inc., 2009.



- Cheng C., Guelfirat T., Messinger C., Schmitt J.O., Schnelte M., Weber P. **Semantic Degrees for Industrie 4.0 Engineering: Deciding on the Degree of Semantic Formalization to Select Appropriate Technologies**. Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering, 1010–1013, 2015.
- Dorst W., Glohr C., Han T., Knafla F., Loewen U., Rosen R., Schiemann T., Vollmar F., Winterhalter C. **Implementation Strategy Industrie 4.0. Report on the Results of the Industrie 4.0 Platform**. Bitkom/VDMA/ZVEI, 13, 2016.
- Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., Hoffmann, M. **Industry 4.0**. Business & Information Systems Engineering, 6, 239–242, Springer, 2014.
- Santos, B.P., Silva, L.A., Celes, C.S.F.S., Borges, J.B., Neto, B.S. Peres, Vieira, M.A.M., Vieira, L.F.M., Goussevskaia, O.N., Loureiro, A. **Internet das Coisas: da Teoria à Prática**. Minicursos SBRC-Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, 31, 16, 2016.
- Breque, M., De Nul, L., Petridis, A., et al. **Industry 5.0: Towards a Sustainable, Human-Centric and Resilient European Industry**. Luxembourg, LU: European Commission, Directorate-General for Research and Innovation, 2021.
- Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., Adda, M. **On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges**. Applied Sciences, 12(16), 8081, MDPI, 2022.
- Weyns, D. **An Introduction to Self-Adaptive Systems: A Contemporary Software Engineering Perspective**. John Wiley & Sons, 2020.
- Eramo, R., Bordeleau, F., Combemale, B., van Den Brand, M., Wimmer, M., Wortmann, A. **Conceptualizing Digital Twins**. IEEE Software, 39(2), 39–46, IEEE, 2021.
- Qi, Q., Tao, F. **Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison**. IEEE Access, 6, 3585–3593, IEEE, 2018.
- Negri, E., Fumagalli, L. **A Review of the Roles of Digital Twin in CPS-Based Production Systems**. Procedia Manufacturing, 11, 939–948, Elsevier, 2017.
- Simon, H.A. **The Sciences of the Artificial**. MIT Press, 1996.
- Pimentel, M., Filippo, D., dos Santos, T.M. **Design Science Research: Pesquisa Científica Atrelada ao Design de Artefatos**. RE@D-Revista de Educação a Distância e eLearning, 3(1), 37–61, 2020.
- Bolton, W. **Instrumentation and Control Systems**. Newnes, 2021.
- SANTOS, M.M.D., Leme, M.O., JUNIOR, S.L.S. **Indústria 4.0: Fundamentos, Perspectivas e Aplicações**. Saraiva Educação SA, 2018.
- Cornago, S., Tan, Y.S., Brondi, C., Ramakrishna, S., Low, J.S.C. **Systematic Literature Review on Dynamic Life Cycle Inventory: Towards Industry 4.0 Applications**. Sustainability, 14(11), 6464, MDPI, 2022.
- Fan, Y.J., Yin, Y.H., Da Xu, L., Zeng, Y., Wu, F. **IoT-Based Smart Rehabilitation System**. IEEE Transactions on Industrial Informatics, 10(2), 1568–1577, IEEE, 2014.

- Greengard, S. **The Internet of Things**. MIT Press, 2021.
- Frank, A.G., Dalenogare, L.S., Ayala, N.F. **Industry 4.0 Technologies: Implementation Patterns in Manufacturing Companies**. International Journal of Production Economics, 210, 15–26, Elsevier, 2019.
- Drath, R., Horch, A. **Industrie 4.0: Hit or Hype? [Industry Forum]**. IEEE Industrial Electronics Magazine, 8(2), 56–58, IEEE, 2014.
- Hillar, G.C. **MQTT Essentials-A Lightweight IoT Protocol**. Packt Publishing Ltd, 2017.
- Soni, D., Makwana, A. **A Survey on MQTT: A Protocol of Internet of Things (IoT)**. International Conference on Telecommunication, Power Analysis and Computing Techniques (ICTPACT-2017), 20, 173–177, 2017.
- Gruber, T.R. **Toward Principles for the Design of Ontologies Used for Knowledge Sharing?** International Journal of Human-Computer Studies, 43(5-6), 907–928, Elsevier, 1995.
- Uschold, M. and Gruninger, M. **Ontologies: Principles, Methods and Applications**. The Knowledge Engineering Review, 11(2), 93–136, Cambridge University Press, 1996.
- McGuinness, D. and STANFORD UNIV CA DEPT OF COMPUTER SCIENCE. **Development of Semantic Web-Markup Languages, Web Services, Rules, Explanation, Querying, Proof and Reasoning**, 2008.
- Musen, M.A. **The Protégé Project: A Look Back and A Look Forward**. AI Matters, 1(4), 4–12, ACM New York, NY, USA, 2015.
- Horridge, M. and Bechhofer, S. **The OWL API: A Java API for OWL Ontologies**. Semantic Web, 2(1), 11–21, IOS Press, 2011.
- Carroll, J.J., Dickinson, I., Dollin, C., Reynolds, D., Seaborne, A., Wilkinson, K. **Jena: Implementing the Semantic Web Recommendations**. Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters, 74–83, 2004.
- Noy, N.F. and McGuinness, D.L. **Ontology Development 101: A Guide to Creating Your First Ontology**, Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and... , 2001.
- Steinmetz, C. **Uma Abordagem para a Integração de Sistemas Industriais Aplicando o Conceito de Internet das Coisas e de Modelos Semânticos no Contexto da Indústria 4.0**, 2018.
- Russell, S.J. and Norvig, P. **Artificial Intelligence: A Modern Approach**. London, 2010.
- Negnevitsky, M. **Artificial Intelligence: A Guide to Intelligent Systems**. Pearson Education, 2005.
- Min, H. **Artificial Intelligence in Supply Chain Management: Theory and Applications**. International Journal of Logistics: Research and Applications, 13(1), 13–39, Taylor & Francis, 2010.

- Chouchene, A., Carvalho, A., Lima, T.M., Charrua-Santos, F., Osório, G.J., Barhoumi, W. **Artificial Intelligence for Product Quality Inspection Toward Smart Industries: Quality Control of Vehicle Non-Conformities**. 2020 9th International Conference on Industrial Technology and Management (ICITM), 127–131, IEEE, 2020.
- Zhong, R.Y., Xu, X., Klotz, E., Newman, S.T. **Intelligent Manufacturing in the Context of Industry 4.0: A Review**. *Engineering*, 3(5), 616–630, Elsevier, 2017.
- Hinduja, H., Kekkar, S., Chourasia, S., Chakrapani, H.B. **Industry 4.0: Digital Twin and Its Industrial Applications**. *Int. J. Sci. Eng. Technol*, 8(4), 2020.
- Ahmad, K., Abdelrazek, M., Arora, C., Bano, M., Grundy, J. **Requirements Engineering for Artificial Intelligence Systems: A Systematic Mapping Study**. *Information and Software Technology*, 107176, Elsevier, 2023.
- De Lemos, R., Giese, H., Müller, H.A., Shaw, M., Andersson, J., Litoiu, M., Schmerl, B., Tamura, G., Villegas, N.M., Vogel, T., et al. **Software Engineering for Self-Adaptive Systems: A Second Research Roadmap**. *Software Engineering for Self-Adaptive Systems II: International Seminar, Dagstuhl Castle, Germany, October 24-29, 2010 Revised Selected and Invited Papers*, 1–32, Springer, 2013.
- Aivaliotis, P., Georgoulas, K., Chryssolouris, G. **The Use of Digital Twin for Predictive Maintenance in Manufacturing**. *International Journal of Computer Integrated Manufacturing*, 32(11), 1067–1080, Taylor & Francis, 2019.
- Hevner, A.R., March, S.T., Park, J., Ram, S. **Design Science in Information Systems Research**. *Management Information Systems Quarterly*, 28(1), 6, 2008.
- Hevner, A.R. **A Three Cycle View of Design Science Research**. *Scandinavian Journal of Information Systems*, 19(2), 4, 2007.
- Lee, J., Bagheri, B., Kao, H.-A. **A Cyber-Physical Systems Architecture for Industry 4.0-Based Manufacturing Systems**. *Manufacturing Letters*, 3, 18–23, Elsevier, 2015.
- Neves, P.C., Schmerl, B., Câmara, J., Bernardino, J. **Big Data in Cloud Computing: Features and Issues**. *International Conference on Internet of Things and Big Data*, 2, 307–314, SCITEPRESS, 2016.
- Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., Khan, S.U. **The Rise of “Big Data” on Cloud Computing: Review and Open Research Issues**. *Information Systems*, 47, 98–115, Elsevier, 2015.
- Kitchenham, B., Brereton, P. **A Systematic Review of Systematic Review Process Research in Software Engineering**. *Information and Software Technology*, 55(12), 2049–2075, Elsevier, 2013.
- Mourão, E., Pimentel, J.F., Murta, L., Kalinowski, M., Mendes, E., Wohlin, C. **On the Performance of Hybrid Search Strategies for Systematic Literature Reviews in Software Engineering**. *Information and Software Technology*, 123, 106294, Elsevier, 2020.
- Petticrew, M., Roberts, H. **Systematic Reviews in the Social Sciences: A Practical Guide**. John Wiley & Sons, 2008.

- Magee, J. and Kramer, J., **Dynamic Structure in Software Architectures**. ACM SIGSOFT Software Engineering Notes,
- Provost, Foster and Fawcett, Tom. **Data Science for Business: What you need to know about data mining and data-analytic thinking**. "O'Reilly Media, Inc.", 2013.
- MOHAMMED, Mohssen; KHAN, Muhammad Badruddin; BASHIER, Eihab Bashier Mohammed. **Machine learning: algorithms and applications**. Crc Press, 2016.
- Basili, Victor R and Weiss, David M. **A Methodology for Collecting Valid Software Engineering Data**. IEEE Transactions on Software Engineering, (6), 1984, 728–738.
- Bussab, Wilton de O and Morettin, Pedro A. **Estatística básica**. In: Estatística Básica, 2010, xvi–540.
- Moore, Leslie M. **The Basic Practice of Statistics**. Taylor & Francis, 1996.
- Fischer, Hans. **A History of the Central Limit Theorem: From Classical to Modern Probability Theory**. Springer, Volume 4, 2011.
- He, Haibo and Garcia, Edwardo A. **Learning from imbalanced data**. IEEE Transactions on knowledge and data engineering, 21(9), 2009, 1263–1284.
- Chawla, Nitesh V., Bowyer, Kevin W., Hall, Lawrence O., and Kegelmeyer, W. Philip. **SMOTE: synthetic minority over-sampling technique**. Journal of artificial intelligence research, 16, 2002, 321–357.
- Ali, M. **PyCaret: An open source, low-code machine learning library in Python**. PyCaret version, 2, 2020.
- Womack, James P., Jones, Daniel T., and Roos, Daniel. **The Machine That Changed the World: The Story of Lean Production—Toyota's Secret Weapon in the Global Car Wars That Is Now Revolutionizing World Industry**. Simon and Schuster, 2007.
- Ashton, Kevin et al. **That 'internet of things' thing**. RFID Journal. 22(7), 2009, 97–114.
- van Dinter, R., Tekinerdogan, B., and Catal, C. **Predictive Maintenance Using Digital Twins: A Systematic Literature Review**. Information and Software Technology, 2022, 107008; Elsevier.
- Amershi, Saleema, Begel, Andrew, Bird, Christian, DeLine, Robert, Gall, Harald, Kamar, Ece, Nagappan, Nachiappan, Nushi, Besmira, and Zimmermann, Thomas. **Software engineering for machine learning: A case study**. In Proceedings of the 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP), pages 291–300, IEEE, 2019.
- Vallim Filho A.R.A., Moraes D.F., Vallim M.V.B.A., Silva L.S., Silva L.A. **A Machine Learning Modeling Framework for Predictive Maintenance Based on Equipment Load Cycle: An Application in a Real World Case**. Energies. 15(10), 2022, 3724; <https://doi.org/10.3390/en15103724>.

Karaiskos V., Zinas N., Gkamas T., Karolos I.A., Pikridas C., Vrettos N., Tsioukas V., Kontogiannis S. **Proposed Industry 4.0 Maintenance Framework for Critical and Demanding Infrastructures and Processes**. In Proceedings of the 2022 7th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), 2022, pages 1–5; IEEE.

Leontaris L., Mitsiaki A., Charalampous P., Dimitriou N., Leivaditou E., Karamanidis A., Margetis G., Apostolakis K.C., Pantoja S., Stephanidis C., et al. **A Blockchain-Enabled Deep Residual Architecture for Accountable, In-Situ Quality Control in Industry 4.0 with Minimal Latency**. Computers in Industry. 149, 2023, 103919; <https://doi.org/10.1016/j.compind.2022.103919>.

Naqvi S.M.R., Ghufuran M., Meraghni S., Varnier C., Nicod J.-M., Zerhouni N. **Human Knowledge Centered Maintenance Decision Support in Digital Twin Environment**. Journal of Manufacturing Systems. 65, 2022, pages 528–537; <https://doi.org/10.1016/j.jmsy.2022.04.013>.

Bousdekis A., Lepenioti K., Apostolou D., Mentzas G. **A review of data-driven decision-making methods for industry 4.0 maintenance applications**. Electronics.10(7), 2021, 828; <https://doi.org/10.3390/electronics10070828>.

Yıldız G.B., Soylu B. **Integrating preventive and predictive maintenance policies with system dynamics: A decision table approach**. Advanced Engineering Informatics.56, 2023, 101952; <https://doi.org/10.1016/j.aei.2023.101952>.

Rojek I., Jasiulewicz-Kaczmarek M., Piechowski M., Mikołajewski D. **An artificial intelligence approach for improving maintenance to supervise machine failures and support their repair**. Applied Sciences.13(8), 2023, 4971; <https://doi.org/10.3390/app13084971>.

Ortiz G., Caravaca J.A., García-de-Prado A., Boubeta-Puig J., et al. **Real-time context-aware microservice architecture for predictive analytics and smart decision-making**. IEEE Access.7, 2019, 183177–183194; <https://doi.org/10.1109/ACCESS.2019.2953491>.

Kontogiannis S., Gkamas T., Pikridas C. **Deep Learning Stranded Neural Network Model for the Detection of Sensory Triggered Events**. Algorithms. 16(4), 2023, 202; <https://doi.org/10.3390/algorithms16040202>.

Rodrigues J.A., Farinha J.T., Mendes M., Mateus R.J.G., Cardoso A.J.M. **Comparison of Different Features and Neural Networks for Predicting Industrial Paper Press Condition**. Energies. 15(17), 2022, 6308; <https://doi.org/10.3390/en15176308>.

Rosati R., Romeo L., Cecchini G., Tonetto F., Viti P., Mancini A., Frontoni E. **From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0**. Journal of Intelligent Manufacturing. 34(1), 2023, 107–121; <https://doi.org/10.1007/s10845-021-01812-0>.

Ferreira L., Pilastrri A., Romano F., Cortez P. **Using supervised and one-class automated machine learning for predictive maintenance**. Applied Soft Computing. 131, 2022, 109820; <https://doi.org/10.1016/j.asoc.2022.109820>.

Chen C., Wang C., Lu N., Jiang B., Xing Y. **A Data-Driven Predictive Maintenance Strategy Based on Accurate Failure Prognostics**. *Eksploatacja i Niezawodność*. 23(2), 2021, 387–394.

de Carvalho Chrysostomo G.G., de Aguiar Vallim M.V.B., Da Silva L.S., Silva L.A., de Aguiar Vallim Filho A.R. **A Framework for Big Data Analytical Process and Mapping—BAProm: Description of an Application in an Industrial Environment**. *Energies*. 13(22), 2020, 6014.

Cavaliere S., Salafia M.G. **A Model for Predictive Maintenance Based on Asset Administration Shell**. *Sensors*. 20(21), 2020, 6028.

Arena S., Florian E., Zennaro I., Orrù P.F., Sgarbossa F. **A Novel Decision Support System for Managing Predictive Maintenance Strategies Based on Machine Learning Approaches**. *Safety Science*. 146, 2022, 105529.

Katona, A., Panfilov, P. **Building Predictive Maintenance Framework for Smart Environment Application Systems**. In *Proceedings of the Annals of DAAAM and International DAAAM Symposium*, 2018, pp. 0460–0470.

Naqvi, S.M.R., Ghufuran, M., Meraghni, S., Varnier, C., Nicod, J.-M., Zerhouni, N. **CBR-based Decision Support System for Maintenance Text Using NLP for an Aviation Case Study**. In *Proceedings of the 2022 Prognostics and Health Management Conference (PHM-2022 London)*, IEEE, 2022, pp. 344–349.

Chang, R.-I., Lee, C.-Y., Hung, Y.-H. **Cloud-based Analytics Module for Predictive Maintenance of the Textile Manufacturing Process**. *Applied Sciences*, 11(21), 2021, 9945; <https://doi.org/10.3390/app11219945>.

Strauß, P., Schmitz, M., Wöstmann, R., Deuse, J. **Enabling of Predictive Maintenance in the Brownfield Through Low-cost Sensors, an IIoT-Architecture and Machine Learning**. In *Proceedings of the 2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018, pp. 1474–1483.

Peres, R.S., Rocha, A.D., Leitao, P., Barata, J. **IDARTS—Towards Intelligent Data Analysis and Real-time Supervision for Industry 4.0**. *Computers in Industry*, 101, 2018, pp. 138–146.

Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., Loncarski, J. **Machine Learning Approach for Predictive Maintenance in Industry 4.0**. In *Proceedings of the 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, IEEE, 2018, pp. 1–6.

Kanawaday, A. and Sane, A. **Machine learning for predictive maintenance of industrial machines using IoT sensor data**. In *Proceedings of the 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pages 87–90, 2017.

Cachada, A. et al. **Maintenance 4.0: Intelligent and predictive maintenance system architecture**. In *Proceedings of the 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, volume 1, pages 139–146, 2018.

Achouch M., Dimitrova M., Dhouib R., Ibrahim H., Adda M., Sattarpanah Karganroudi S., Ziane K., Aminzadeh A. **Predictive Maintenance and Fault Monitoring Enabled by Machine Learning: Experimental Analysis of a TA-48 Multistage Centrifugal Plant Compressor**. Applied Sciences. 13(3), 2023, 1790; <https://doi.org/10.3390/app13031790>.

Gohel, H.A. et al. **Predictive maintenance architecture development for nuclear infrastructure using machine learning**. Nuclear Engineering and Technology, 52(7): 1436–1442, 2020.

Ayvaz, S. and Alpay, K. **Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time**. Expert Systems with Applications, 173: 114598, 2021.

Calabrese M., Cimmino M., Fiume F., Manfrin M., Romeo L., Ceccacci S., Paolanti M., Toscano G., Ciandrini G., Carrotta A., et al. **SOPHIA: An Event-Based IoT and Machine Learning Architecture for Predictive Maintenance in Industry 4.0**. Information. 11(4), 2020, 202; <https://doi.org/10.3390/info11040202>.

Rojek I., Jasiulewicz-Kaczmarek M., Piechowski M., Mikołajewski D. **An Artificial Intelligence Approach for Improving Maintenance to Supervise Machine Failures and Support Their Repair**. Applied Sciences. 13(8), 2023, 4971; <https://doi.org/10.3390/app13084971>.

Aboshosha A., Haggag A., George N., Hamad H.A. **IoT-based Data-driven Predictive Maintenance Relying on Fuzzy System and Artificial Neural Networks**. Scientific Reports. 13(1), 2023, 12186; <https://doi.org/10.1038/s41598-023-28625-8>.

Pejić Bach M., Topalović A., Krstić Ž., Iveć A. **Predictive Maintenance in Industry 4.0 for the SMEs: A Decision Support System Case Study Using Open-Source Software**. Designs. 7(4), 2023, 98; <https://doi.org/10.3390/designs7040098>.

Neto A.A., Carrijo B.S., Romanzini J.G.B., Deschamps F., de Lima E.P. **Digital Twin-Driven Decision Support System for Opportunistic Preventive Maintenance Scheduling in Manufacturing**. Procedia Manufacturing. 55, 2021, 439–446; <https://doi.org/10.1016/j.promfg.2021.10.059>.

Gupta V., Mitra R., Koenig F., Kumar M., Tiwari M.K. **Predictive Maintenance of Baggage Handling Conveyors Using IoT**. Computers & Industrial Engineering. 177, 2023, 109033; <https://doi.org/10.1016/j.cie.2022.109033>.

Chuang S.-Y., Sahoo N., Lin H.-W., Chang Y.-H. **Predictive Maintenance with Sensor Data Analytics on a Raspberry Pi-Based Experimental Platform**. Sensors. 19(18), 2019, 3884; <https://doi.org/10.3390/s19183884>.