

# ESTUDO DE CASO APLICADO A *MACHINE LEARNING* NAS INDÚSTRIAS: UMA BIBLIOMETRIA E REVISÃO SISTEMÁTICA DA LITERATURA

## Autores

Thiago Loretto Arico Martins<sup>1</sup>

Claudio Keiji Iwata<sup>2</sup>

Allan dos Anjos Pestana<sup>3</sup>

Marília Macorin de Azevedo<sup>4</sup>

Fabricio José Piacente<sup>5</sup>

## Resumo

Este artigo apresenta um estudo bibliométrico e uma revisão sistemática da literatura sobre as contribuições do método de estudo de caso em aplicações de *machine learning* no setor industrial no período de 2018 a 2022. A bibliometria é uma técnica de pesquisa que utiliza métodos estatísticos para analisar e mensurar diferentes aspectos da produção científica. Por outro lado, a revisão sistemática da literatura é um método de pesquisa rigoroso e estruturado que tem como objetivo sintetizar e analisar criticamente as evidências disponíveis sobre um tópico específico. Nesse contexto, a busca por artigos científicos foi realizada nas bases de dados *Scopus*, *IEEE*, *Web of Science* e *Google Scholar*, com critérios de inclusão de publicações no período mencionado. Utilizou-se a ferramenta Bibliometrix para compilar e analisar os dados obtidos, fornecendo informações sobre o período de busca, quantidade de fontes, autores, coautoria internacional, palavras-chave, referências e citações por documento. Nos resultados, verificou-se que a produção científica sobre o tema apresentou um forte crescimento anual no período analisado. Identificou-se também os países e fontes mais relevantes em número de publicações científicas sobre o tema. Na revisão sistemática da literatura, foi adotado um protocolo específico, auxiliado pela ferramenta Parsifal. Os estudos selecionados destacaram as contribuições do *machine learning* na indústria, principalmente na previsão de falhas e melhoria de desempenho dos processos produtivos. As aplicações de *machine learning* foram utilizadas para aprender com os dados e otimizar os processos sem a necessidade de programação explícita. Em conclusão, este artigo forneceu uma visão abrangente sobre as contribuições do método de estudo de caso em aplicações de *machine learning* na indústria. Os resultados obtidos podem servir como base para futuras pesquisas e fornecer *insights* para o avanço nessa área. Essa pesquisa adotou uma abordagem exploratória, combinando análises quantitativas e qualitativas.

**Palavras-chave:** *Machine learning*. Estudo de caso. Bibliometria. Revisão sistemática da literatura. Sistemas produtivos.

## **CASE STUDY APPLIED TO MACHINE LEARNING IN INDUSTRIES: A BIBLIOMETRY AND SYSTEMATIC REVIEW OF THE LITERATURE**

### **Abstract**

*This article presents a bibliometric study and a systematic literature review on the contributions of the case study method in machine learning applications in the industrial segment from 2018 to 2022. Bibliometrics is a discipline that uses statistical methods to analyze and measure different aspects of scientific production, while systematic literature review is a rigorous and structured research method aimed at synthesizing and critically analyzing the available evidence on a specific topic. In this*

<sup>1</sup> <https://orcid.org/0009-0001-7299-859X>

<sup>2</sup> <https://orcid.org/0009-0004-0767-9949>

<sup>3</sup> <https://orcid.org/0009-0000-3478-0867>

<sup>4</sup> <https://orcid.org/0000-0003-0225-8155>

<sup>5</sup> <https://orcid.org/0000-0001-8306-4541>

context, the search for peer-reviewed scientific articles was conducted in databases such as Scopus, IEEE, Web of Science, and Google Scholar, with inclusion criteria for publications during the mentioned period, where a relevant number of materials were obtained for analysis. The Bibliometrix tool was used to compile and analyze the data, providing information on the search period, quantity of sources, authors, international co-authorship, keywords, references, and citations per document. As a result, it was found that the scientific production on the subject showed a strong annual growth during the analyzed period, and the most relevant countries and sources with the highest number of scientific publications on the topic were identified. In the systematic literature review, a specific protocol was adopted, assisted by the Parsifal tool. The selected studies highlighted the contributions of machine learning in the industry, mainly in predicting failures and improving performance in production processes. Machine learning applications were used to learn from data and optimize processes without the need for explicit programming. In conclusion, this bibliometric study and systematic literature review provided a comprehensive view of the contributions of the case study method in machine learning applications in the industry. The obtained results can serve as a basis for future research and provide valuable insights for advancements in this area. The research adopted an exploratory approach, combining quantitative and qualitative analyses.

**Keywords:** Machine learning. Case study. Bibliometrics. Systematic literature review. Production systems.

## INTRODUÇÃO

A abordagem de estudo de caso é amplamente reconhecida e utilizada como um método de pesquisa para a investigação de fenômenos complexos e situações reais no mundo. Essa abordagem oferece uma oportunidade para a exploração detalhada de um caso específico, possibilitando a análise dentro de um contexto particular e, assim, alcançando uma compreensão aprofundada e abrangente de todos os seus aspectos envolvidos. Autores que são referências nessa área, como Yin (2001) e Stake (1995), tiveram contribuições fundamentais para o desenvolvimento e aprimoramento contínuo dessa abordagem.

Dentro do campo dos sistemas produtivos, a aplicação de estudos de caso tem se mostrado especialmente relevante no contexto do *machine learning*. O conceito de *machine learning* refere-se à capacidade de um sistema ou algoritmo aprender e aprimorar seu desempenho ao longo do tempo, tudo isso sem que seja necessária uma programação explícita para tal. Nesse sentido, o uso de modelos e algoritmos de aprendizado automático tem ganhado destaque, permitindo que previsões sejam realizadas com base em dados previamente conhecidos.

A aplicação de estudos de caso no contexto do *machine learning* tem proporcionado *insights* que contribuem para a compreensão e aprimoramento dos sistemas produtivos. Ao investigar casos específicos dentro desse campo, é possível realizar uma análise minuciosa e detalhada das estratégias de seleção de casos, dos métodos utilizados para a coleta e análise de dados, das estratégias implementadas para o uso do *machine learning* e da avaliação dos resultados obtidos. Essa abordagem proporciona, assim, uma compreensão contextualizada,

fornecendo um conhecimento mais abrangente e detalhado a respeito das complexidades e desafios enfrentados na aplicação do *machine learning* na indústria.

Ao longo deste artigo, serão apresentados de forma abrangente, além dos dados da bibliometria, os resultados e discussões resultados de uma revisão sistemática da literatura. Serão abordados diversos tópicos importantes, tais como a seleção criteriosa dos casos, os métodos de coleta e análise de dados, as estratégias criteriosas empregadas para a implementação do *machine learning* e a avaliação dos resultados alcançados. Além disso, serão discutidos os principais benefícios observados e as limitações identificadas no uso do estudo de caso nesse contexto específico, bem como serão sugeridas possíveis direções para futuras pesquisas.

Espera-se que este trabalho científico possa contribuir para entender o avanço do método de estudo de caso em aplicações de *machine learning* na indústria. Ao oferecer uma análise aprofundada das pesquisas existentes, almeja-se fornecer embasamento teórico sólido e subsídios práticos para os profissionais, pesquisadores e tomadores de decisão que desejam explorar plenamente o potencial do *machine learning* nos sistemas produtivos. A compreensão abrangente dos desafios, benefícios e melhores práticas relacionadas ao uso do *machine learning* em um contexto real é importante para impulsionar a eficiência, a inovação e o contínuo progresso na indústria como um todo.

Uma quantidade significativa de estudos de caso tem relatado ganhos notáveis que resultaram da aplicação do *machine learning* em diferentes setores industriais. Diversos pesquisadores têm relatado esses avanços, como é o caso dos estudos realizados por Fernandes *et al.* (2018), Canizo *et al.* (2019), Lorente-Leyva *et al.* (2020), Papachristou, Chrysopoulos e Bilalis (2021) e Behdinian *et al.* (2022). Essas investigações têm oferecido evidências concretas dos benefícios proporcionados pela adoção de técnicas de aprendizado de máquina nas indústrias. Diante do contexto apresentado, surge a questão de pesquisa: “De que forma o método científico de estudo de caso tem sido empregado nas pesquisas que envolvem o *machine learning* no segmento industrial?”.

## 1.2 Objetivo

Para responder à questão de pesquisa, estabeleceu-se como objetivo geral mapear as contribuições do método de estudo de caso aplicado ao *machine learning* no segmento industrial no período de 2018 a 2022 e, para apoiar esse objetivo, dois objetivos específicos foram estabelecidos:

- Por meio de uma bibliometria, apresentar o cenário de publicações sobre os temas agrupados: Estudo de Caso, *Machine learning* e Indústria;
- Apresentar os principais grupos de ganhos a partir de uma Revisão Sistemática da Literatura.

## 2 REFERENCIAL TEÓRICO

Neste referencial teórico, serão apresentados os temas *machine learning* e estudo de caso.

### 2.1 *Machine learning*

De acordo com a visão de Mahesh (2020), o *Machine learning* (ML) pode ser compreendido como um campo científico que se dedica ao estudo de algoritmos e modelos estatísticos utilizados por sistemas computacionais para realizar tarefas específicas sem a necessidade de programação explícita. A aplicação desses algoritmos de aprendizado é ampla e permeia diversas aplicações que fazem parte do nosso cotidiano. Um exemplo prático é o mecanismo de busca na *web*, como o *Google*, que apresenta um desempenho eficiente graças a um algoritmo de aprendizado que foi capaz de aprender a classificar as páginas da *web* de forma adequada. Além disso, esses algoritmos são empregados para finalidades variadas, como mineração de dados, processamento de imagens e análise preditiva, entre outras.

Mahesh (2020) destaca alguns dos principais modelos de algoritmos de *machine learning*, a saber:

*Supervised Learning*: Aprendizado supervisionado, no qual um algoritmo é treinado para mapear entradas em saídas com base em exemplos rotulados. Os algoritmos aprendem padrões a partir dos dados de treinamento e os aplicam para prever ou classificar novos dados.

*Decision Tree*: Árvore de decisão que representa escolhas e resultados de forma hierárquica. Cada nó da árvore representa atributos a serem classificados, e cada ramo representa um valor que o nó pode assumir.

Naive Bayes: Técnica de classificação baseada no Teorema de Bayes, que parte da suposição de independência entre os preditores. Essa técnica é amplamente empregada em classificação de texto e agrupamento.

*Support Vector Machine* (SVM): Técnica de aprendizado supervisionado utilizada para classificação e análise de regressão. Ela tem a capacidade de realizar classificação linear e não linear.

*Principal Component Analysis* (PCA): Procedimento estatístico que busca reduzir a dimensionalidade dos dados, encontrando combinações lineares de variáveis correlacionadas em componentes principais não correlacionados.

A partir dessas informações, nota-se que o campo do aprendizado de máquina abarca uma diversidade de modelos de algoritmos e é objeto de extensa exploração e pesquisa no período analisado.

## 2.2 Estudo de caso

Segundo Yin (2001), o estudo de caso foi considerado o método menos relevante nas ciências sociais, e os pesquisadores que o adotam são muitas vezes vistos como desviados do foco acadêmico. No entanto, Yin (2001) destaca que o método de estudo de caso continua sendo amplamente utilizado em várias áreas de estudo e é frequentemente adotado como modelo para teses e dissertações em disciplinas como engenharia e sistemas produtivos. O autor também apresenta um paradoxo: se o método de estudo de caso apresenta fragilidades, por que os pesquisadores continuam a utilizá-lo?

Yin (2001) propõe duas possíveis explicações para essa questão. Em primeiro lugar, ele sugere que os pesquisadores podem não ser treinados para trabalhar com métodos alternativos, o que limita suas opções de escolha. Em segundo lugar, Yin (2001) argumenta que renomados pesquisadores e líderes em suas áreas de estudo utilizam o estudo de caso como método principal de pesquisa, o que demonstra a importância prática desse enfoque.

Yin (2001) é amplamente reconhecido como um dos principais pesquisadores e defensores do estudo de caso. Em seus trabalhos, ele enfatiza a importância de conduzir uma pesquisa cuidadosa, sistemática e rigorosa para obter uma compreensão profunda de um caso. Yin (2001) destaca a necessidade de coletar e analisar dados de múltiplas fontes, empregando diferentes métodos e técnicas para abordar questões de pesquisa complexas.

Outra figura proeminente no campo do estudo de caso é Stake (1995), que propôs uma abordagem qualitativa para a realização de estudos de caso. Stake (1995) enfatiza a importância de compreender e interpretar o significado das experiências dos participantes envolvidos no caso, buscando obter uma visão aprofundada dos fenômenos estudados. Sua abordagem valoriza a riqueza e a complexidade dos dados qualitativos, permitindo uma compreensão mais completa do caso em questão.

Além de Yin (2001) e Stake (1995), Merriam (1998) é outra autora essencial no campo do estudo de caso. Suas contribuições têm sido fundamentais para a compreensão e o desenvolvimento dessa abordagem de pesquisa. Merriam (1998) enfoca a importância do estudo de caso como uma estratégia de pesquisa qualitativa, permitindo uma exploração aprofundada e holística de fenômenos complexos.

Existem pontos de convergência significativos entre o protocolo proposto por Merriam (1998) e as abordagens de Yin (2001) e Stake (1995). Em primeiro lugar, todos os três autores valorizam a compreensão contextualizada do caso estudado. Eles reconhecem a importância de analisar o caso em seu ambiente real e considerar os fatores e influências que moldam o fenômeno em questão.

Outro ponto de convergência entre Merriam (1998), Yin (2001) e Stake (1995) é a ênfase na interpretação e no significado dos dados. Todos eles reconhecem que a análise de dados em estudos de caso não se limita apenas a números e estatísticas, mas também envolve uma compreensão profunda do contexto e da experiência dos participantes envolvidos. Eles enfatizam a importância de interpretar os dados de forma reflexiva e holística, buscando identificar padrões, temas e *insights* que possam contribuir para a compreensão do fenômeno estudado.

Portanto, os protocolos propostos por Merriam (1998), Yin (2001) e Stake (1995) convergem na valorização da compreensão contextualizada, na coleta e análise de dados múltiplos e variados, bem como na interpretação reflexiva dos dados. Essas abordagens combinadas oferecem um conjunto robusto de diretrizes para conduzir estudos de caso de alta qualidade, permitindo uma exploração aprofundada de fenômenos complexos e contribuindo para o avanço contínuo do campo do estudo de caso.

Analisando a aplicabilidade dos estudos de caso, Yin (2001) destaca que eles são preferencialmente utilizados para responder a questões de pesquisa do tipo "como" e "por que". Além disso, Yin (2001) sugere que os estudos de caso explanatórios podem ser complementados por estudos exploratórios e descritivos. Yin (2001) propõe o seguinte roteiro a ser seguido para a realização de estudos de caso:

- Projeto do estudo de caso, onde são abordadas questões gerais e critérios para se julgar a qualidade dos projetos de pesquisa;
- Preparação para a coleta de dados, onde o pesquisador coleta dados a partir de múltiplas fontes, incluindo entrevistas, documentos, observações e artefatos. Nesse ponto, define-se também o protocolo de pesquisa, que deve conter uma visão geral

do projeto do estudo de caso, procedimentos de campo, questões do estudo de caso e guia para o relatório de estudo de caso;

- Coleta de evidências, onde são abordados os princípios para coleta de evidências e fontes de evidências;
- Análise das evidências do estudo de caso, onde o autor demonstra as estratégias analíticas gerais e os métodos principais e secundários de análise e;
- Composição de um relatório de estudo de caso, abordando estrutura e procedimento para se compor um relatório de estudo de caso.

Com base nesses pontos, pode-se concluir que o protocolo descrito é um guia para a condução de estudos de caso em pesquisa. O protocolo abrange diversas etapas, desde o projeto do estudo de caso até a composição do relatório final. Ele destaca a importância de estabelecer critérios de qualidade para os projetos de pesquisa, bem como a necessidade de coletar evidências de múltiplas fontes, como entrevistas, documentos, observações e artefatos.

A coleta de dados é vista como uma etapa fundamental, na qual o pesquisador deve seguir procedimentos de campo específicos. O protocolo também enfatiza a análise das evidências do estudo de caso, apresentando estratégias analíticas e métodos de análise. Por fim, é importante destacar a importância da composição de um relatório de estudo de caso estruturado, que deve seguir um procedimento específico.

### **3 MÉTODO**

Essa seção será composta dos dois métodos utilizados para o desenvolvimento deste artigo, que são Bibliometria e RSL – Revisão Sistemática da Literatura.

#### **3.1 Bibliometria**

Segundo Glänzel (1996), a bibliometria é uma técnica de pesquisa que se dedica ao estudo quantitativo da informação contida em documentos bibliográficos, com o objetivo de compreender a comunicação científica e a produção acadêmica. É uma abordagem que utiliza métodos estatísticos para analisar e mensurar diferentes aspectos da produção científica, como a quantidade de publicações, o impacto das citações, a colaboração entre autores e instituições, entre outros.

A fim de obter material atual e relevante, para a busca inicial, foram adotados os seguintes critérios:

- Somente artigos científicos revisados por pares, a fim de garantir a qualidade das publicações.
- Dentro do período de 2018 a 2022, para ter material atual para análise.

Para a realização do estudo bibliométrico, utilizou-se a principal base em número de publicações a partir da *string*("estudo de caso" OR "*case study*") AND (industr\* OR *manufactur\**) AND ("*machine learning*" OR "algoritmo\* de aprendizagem" OR "aprendizado de máquina" OR "reconhecimento de padrões"). A base *Scopus* retornou 2479 artigos, enquanto a *IEEE* retornou 1218, a *Web of Science* 915, e a base *Google Scholar* 100. Nesta última base, a quantidade de 100 artigos representa a busca até a 10ª página. A partir da 11ª página na base *Google Scholar*, a contagem de palavras-chaves tornou-se menor que 3, e decidiu-se interromper a busca nesse ponto. Utilizou-se a ferramenta *Bibliometrix* para a compilação e análise dos dados obtidos a partir do arquivo .bib retornado a partir da busca dos 2479 artigos da base *Scopus*. A Figura 1 demonstra uma compilação dos principais resultados da busca na base *Scopus*.

**Figura 1** Compilação dos principais resultados da busca na base *Scopus*

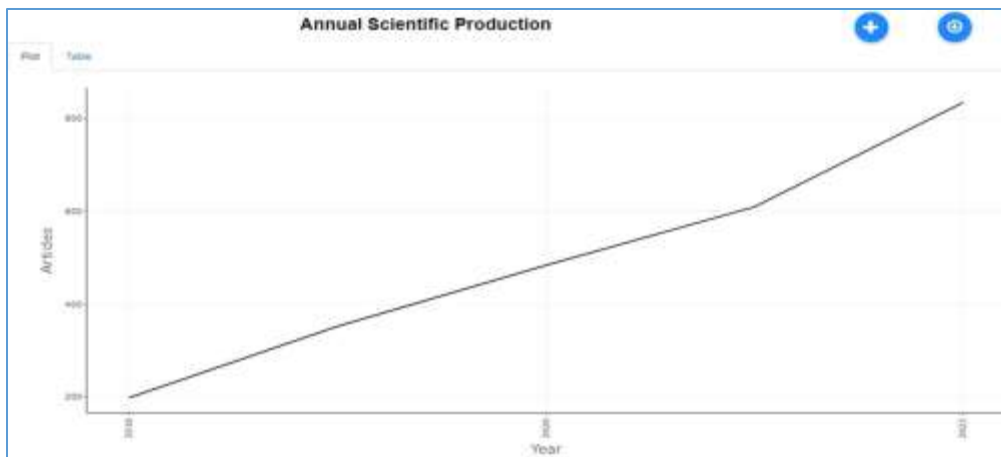


**Fonte:** os autores (2023)

Com base na Figura 3, pode-se concluir que o período de busca foi de 2018 a 2022, resultando em 1235 fontes (artigos) encontradas. Dos 1235 artigos, 2479 foram retornados a partir dessas fontes, e a taxa anual de crescimento das publicações é de 43,3%. Além disso, observa-se que 7350 autores foram referenciados, sendo que apenas 97 autores publicaram artigos sem a participação de outros autores. A taxa de coautoria internacional é de 23,07%, e a média de coautoria é de 3,85 autores por documento. Também é notável que os autores utilizaram 5876 palavras-chave, referenciaram em média 1 fonte por documento, com uma média de idade dos documentos de 2,38 anos e uma média de citações por documento igual a 9,096. O Gráfico 1 demonstra a produção científica por ano dentro do recorte de 2018 a 2022.



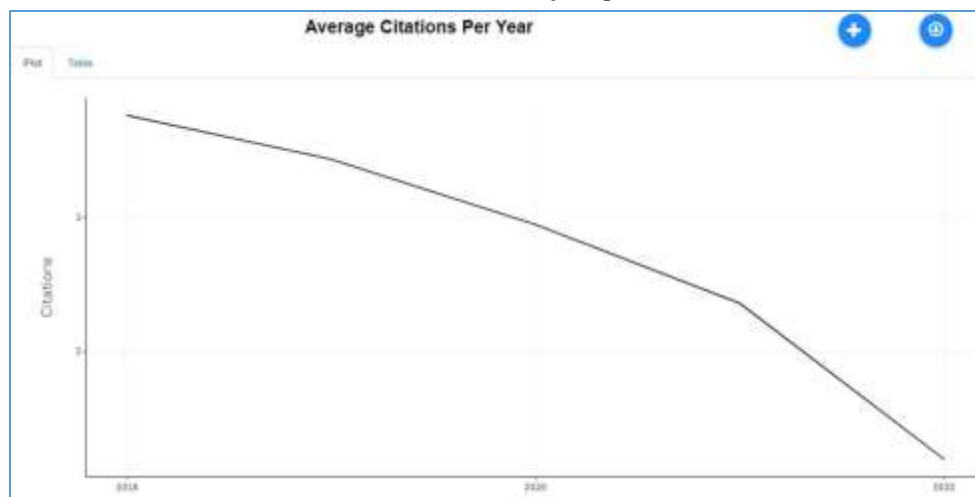
**Gráfico 1** Produção científica por ano



**Fonte:** os autores (2023)

Pode-se concluir que a produção científica sobre o tema buscado está em ascensão no período de 2018 a 2022. Em contrapartida, o gráfico 2 demonstra uma queda no número de citações por ano.

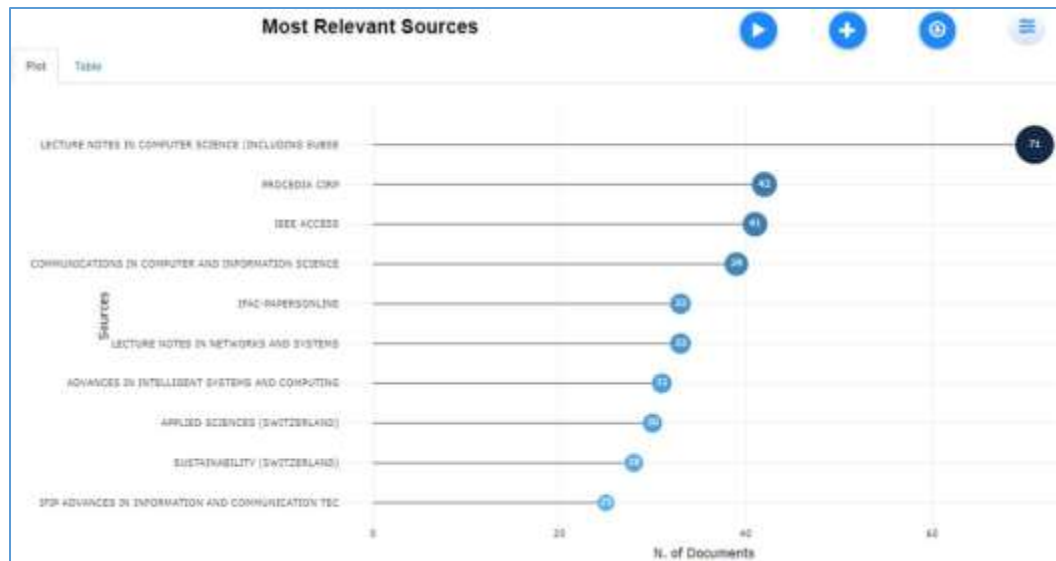
**Gráfico 2** Citações por ano



**Fonte:** os autores (2023)

O motivo da queda no número de citações por ano não será explorado neste estudo, mas é possível campo para exploração em outras pesquisas. O gráfico 3 demonstra as fontes mais relevantes em número de publicações para o período de 2018 a 2022.

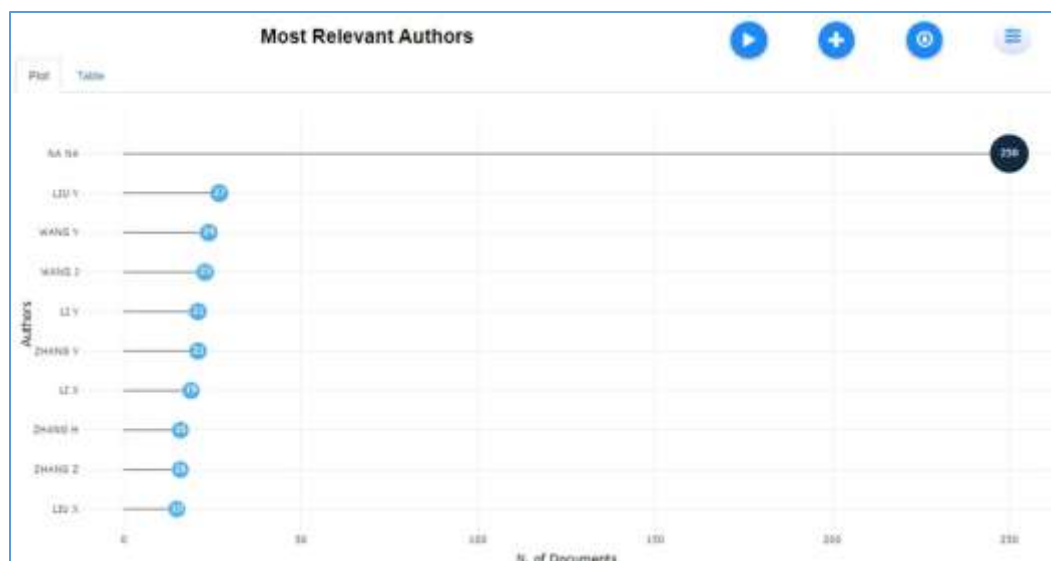
Gráfico 3 Fontes mais relevantes



Fonte: os autores (2023)

Pode-se concluir que a maior parte das publicações partiu da fonte *Lecture Notes in Computer Science* (71 publicações), seguida de *Procedia CIRP* (42 publicações), *IEEE Access* (41 publicações) e *Communications in Computer and Information Science* (39 publicações). O gráfico 4 demonstra os autores mais relevantes em número de publicações.

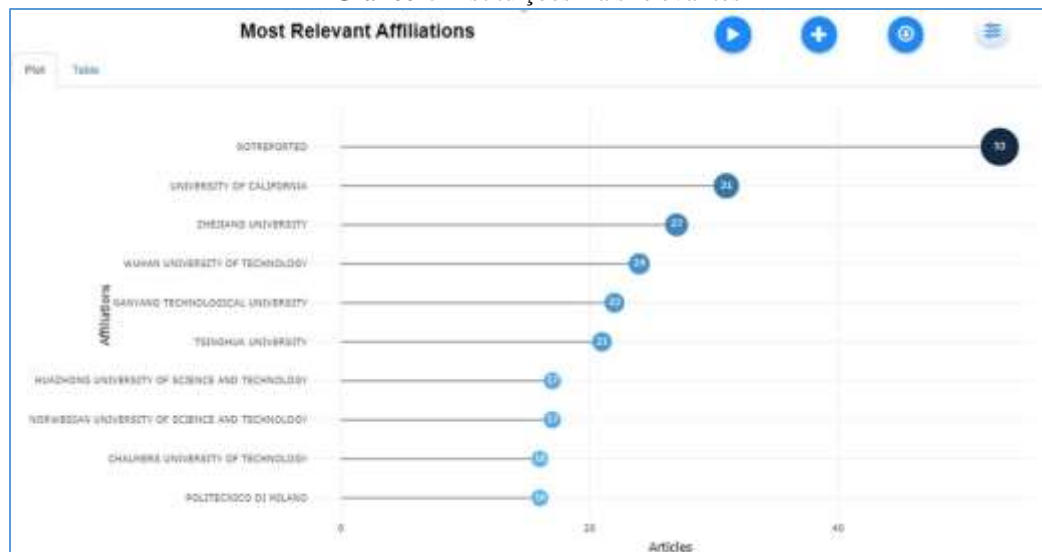
Gráfico 4 Autores mais relevantes



Fonte: os autores (2023)

Pode-se concluir que grande parte dos dados não foram apresentados no arquivo extraído da base *Scopus* (250 publicações). Daqueles onde apresentam-se os dados de autoria no arquivo *.bib*, pode-se concluir que Liu Y. apresenta a maior parte das publicações (27 publicações), seguido de Wang Y. (24 publicações) e Wang J. (23 publicações). O gráfico 5 demonstra as instituições com maior número de afiliações dos autores dos artigos dessa busca na base *Scopus*.

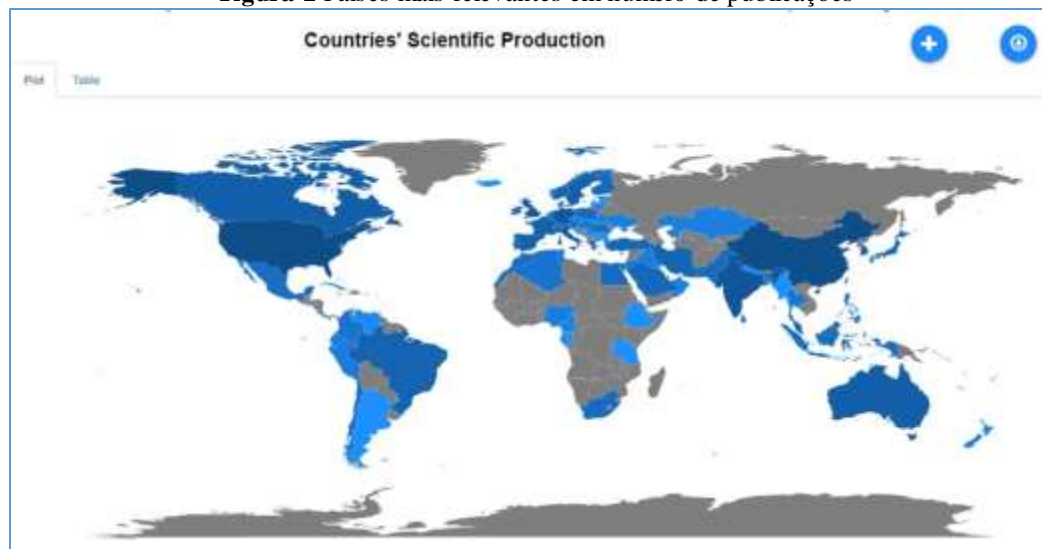
**Gráfico 5** Instituições mais relevantes



**Fonte:** os autores (2023)

Pode-se concluir que grande parte dos dados (53) não foram retornados nessa busca na base *Scopus*. Adiante, a *University of California* apresenta a maior parte das afiliações (31 afiliações), seguida da *Zhejiang University* (27 afiliações) e *Wuhan University of Technology* (24 afiliações). A figura 2 demonstra em escala de azul (quanto mais escuro, maior o número de publicações) os países mais relevantes em número de publicações.

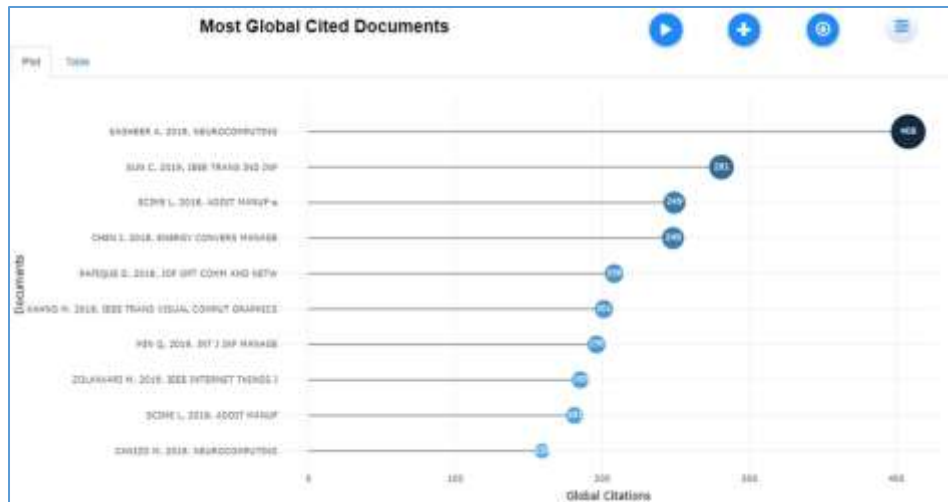
**Figura 2** Países mais relevantes em número de publicações



**Fonte:** os autores (2023)

Pode-se concluir que grande os principais países que publicaram sobre o tema entre 2018 e 2022 são China, Estados Unidos, Índia, Austrália, Inglaterra, Alemanha, Brasil e Canadá. O gráfico 6 demonstra os mais citados documentos em âmbito global.

Gráfico 6 Documentos mais citados em âmbito global



Fonte: os autores (2023)

Pode-se concluir a maior parte dos documentos citados provém da fonte *Sagheer A. Neurocomputing* (408 citações), seguida de *Sun C. IEEE Trans Ind Inf* (281 citações), *Scime L. Addit Manuf-a* (249 citações) e *Chen J. Energy Convers Manage* (248 citações). A figura 3 demonstra a núvem de palavras retornada pela ferramenta Bibliometrix para o arquivo .bib extraído da base *Scopus*.

Figura 3 Documentos mais citados em âmbito global



Fonte: os autores (2023)

Pode-se concluir que palavras como *Machine learning*, *Deep learning*, *Case Studies*, *Learning Systems*, *Decision Trees*, *Manufacture* e outras foram relevantes nos trabalhos retornados da base *Scopus* e forneceram uma nova visão para compor uma *string* de busca, qual foi adotada na RSL. A bibliometria realizada e aqui apresentada foi fundamental para a compreensão do cenário de pesquisa e forneceu embasamento para seguir-se com a revisão sistemática da literatura.

### 3.2 Revisão Sistemática da Literatura

De acordo com Fink (2005), a Revisão Sistemática da Literatura (RSL) é um processo sistemático, claro, amplo e replicável, com o objetivo de identificar, avaliar e resumir a totalidade dos trabalhos completos e registrados elaborados por pesquisadores, acadêmicos e profissionais. No entanto, conforme apontado por Okoli (2019), é importante ressaltar que o livro de Fink (2005), embora seja um guia para realizar revisões sistemáticas, tem um foco nas ciências da saúde e uma abordagem superficial em relação aos estudos qualitativos, o que limita sua aplicabilidade para pesquisadores nas áreas de administração e ciências sociais. De acordo com Okoli (2019), uma revisão sistemática da literatura deve ser abrangente e independente, seguindo uma metodologia clara e explícita para descrever os procedimentos utilizados. Ela deve incluir todo o material relevante e ser passível de replicação por outros pesquisadores que desejem adotar a mesma abordagem na revisão do tema. Conforme Okoli (2019), um protocolo de revisão sistemática da literatura deve contemplar oito estágios, que são apresentados no Quadro 1.

**Quadro 1** Estágios do protocolo de RSL

<b>Estágios</b>	<b>Descrição</b>
1. Identificação do objetivo	É fundamental que os revisores estabeleçam de forma clara o propósito e os objetivos da revisão, tornando-os explícitos para os leitores.
2. Planejamento do protocolo e treinamento da equipe	É necessário elaborar um protocolo detalhado e fornecer treinamento aos revisores, garantindo consistência na condução da revisão.
3. Aplicação de uma seleção prática	Os revisores devem explicitar quais estudos foram considerados e quais não foram justificando suas decisões de exclusão.
4. Busca da bibliografia	Os revisores devem descrever detalhadamente os procedimentos da busca bibliográfica e justificar a abrangência da pesquisa conduzida.
5. Extração os dados	Os revisores devem de forma metódica e organizada obter as informações relevantes de cada estudo que está sendo analisado na revisão.
6. Avaliação da qualidade	Os revisores devem estabelecer critérios que permitam avaliar a qualidade dos artigos selecionados, classificando-os de acordo com as metodologias de pesquisa adotadas.
7. Síntese dos estudos	Os revisores devem integrar de maneira apropriada os dados extraídos dos estudos, aplicando técnicas adequadas tanto quantitativas quanto qualitativas.
8. <i>Report</i> da revisão	A revisão deve ser redigida de maneira detalhada o bastante para que outros pesquisadores possam reproduzir os resultados de forma autônoma.

**Fonte:** adaptado de Okoli (2019)

A análise do quadro 1 demonstra a importância fundamental dos oito passos propostos pelo autor para a execução de uma revisão sistemática da literatura com rigor metodológico. Eles constituem uma estrutura sólida que orienta o processo de forma abrangente, garantindo a consistência e a confiabilidade dos resultados obtidos.

### 3.3 Planejamento

#### 3.3.1 Planejamento do protocolo

Para realizar a Revisão Sistemática da Literatura (RSL), adotou-se um protocolo específico disponibilizado na plataforma online Parsifal. No Parsifal, o protocolo é dividido em quatro etapas: revisão, planejamento, condução e declaração. Pode-se observar que os oito estágios propostos por Okoli para a RSL estão abrangidos nas etapas do Parsifal.

Na etapa de revisão, foram inseridos na ferramenta três dados iniciais: o título do trabalho, que coincide com o título deste artigo; a descrição, que consiste no resumo do artigo; e os nomes dos autores. No planejamento, a primeira subetapa diz respeito à definição do protocolo, na qual foi formalizado o objetivo geral: mapear as contribuições do método de estudo de caso aplicado ao *machine learning* no segmento industrial no período de 2018 a 2022. Como objetivos específicos, formalizou-se:

- Por meio de uma bibliometria, apresentar o cenário de publicações sobre os temas agrupados: Estudo de Caso, *Machine learning* e Indústria;
- Apresentar os principais grupos de ganhos a partir de uma Revisão Sistemática da Literatura.

Utilizou-se o método PICOC (População, Intervenção, Comparação, Saídas e Contexto) para estruturar os dados na ferramenta. O Quadro 2 apresenta os dados formalizados em cada campo.

**Quadro 2** Formalização dos dados do PICOC

Método PICOC	Dados formalizados
<i>Population</i>	Artigos científicos publicados em revistas e eventos acadêmicos
<i>Intervention</i>	Coleta dos dados das aplicações de <i>machine learning</i> na indústria
<i>Comparison</i>	Diferentes aplicações
<i>Outcome</i>	<i>Report</i> sobre as principais aplicações observadas
<i>Context</i>	Artigos científicos publicados em revistas e eventos acadêmicos sobre estudos realizados em indústrias

**Fonte:** os autores (2023)

Na plataforma Parsifal, foi inserida a seguinte questão de pesquisa: como o método científico estudo de caso tem sido empregado em pesquisas que envolvem *machine learning* no segmento industrial?

Considerando a natureza dos dados pesquisados e a núvem de palavras demonstrada na Figura 3, três palavras-chave e seus sinônimos, relacionados diretamente com a intervenção proposta pelo método PICOC, foram selecionados. O Quadro 3 apresenta as palavras-chave que foram selecionadas e seus respectivos sinônimos.

**Quadro 3** Palavras-chave e seus sinônimos

Palavra-chave	Sinônimos	Relacionada a
Estudo de caso	<i>Case study</i> <i>Study of case</i>	Intervenção
Indústria	Industrial <i>Industry</i> <i>Manufacturing</i>	Intervenção
<i>Machine learning</i>	Algoritmos de Aprendizagem Aprendizado de Máquina <i>Deep learning</i> Reconhecimento de Padrões Árvores de Decisão <i>Decision Tree</i>	População

**Fonte:** os autores (2023)

Para formular a *string* de pesquisa, utilizou-se:

"Estudo de caso" juntamente com seus sinônimos relacionados, utilizando o operador booleano *OR* entre elas;

"Indústria" juntamente com seus sinônimos relacionados, utilizando o operador booleano *OR* entre elas;

*"Machine learning"* juntamente com seus sinônimos relacionados, utilizando o operador booleano *OR* entre elas;

Entre cada grupo de sinônimos das palavras-chave, foi utilizado o operador booleano *AND*. Dessa maneira, a *string* resultante foi a seguinte:

("estudo de caso" *OR* "*case study*" *OR* "*study case*") *AND* (industr\* *OR* manufactur\*) *AND* ("*machine learning*" *OR* "algoritmo\* de aprendizagem" *OR* "aprendizado de

máquina” OR “*deep learn\**” OR “reconhecimento de padrões” OR “árvore\* de decis\*” OR “*Decision Tree*”).

Para realizar a busca, foram selecionadas as duas principais bases de informação em número de artigos, conforme apresentado no Quadro 4, considerando que essas bases forneceriam um número significativo de artigos disponíveis para consulta..

**Quadro 4** Fontes utilizadas para busca de artigos

Fonte	URL
Scopus	Scopus.com
IEEE	ieeexplore.ieee.org

**Fonte:** os autores (2023)

Ao realizar a pesquisa nas bases de dados, foram aplicados critérios de inclusão e exclusão com o objetivo de controlar o número de resultados e garantir uma base sólida para a avaliação posterior da qualidade dos artigos. O quadro 5 apresenta os critérios de inclusão e exclusão adotados durante o processo de busca.

**Quadro 5** Critérios de inclusão e exclusão

Critérios de inclusão	Critérios de exclusão
Dentro do período de 2018 a 2022	Não é artigo científico
Escrito em português ou inglês	

**Fonte:** os autores (2023)

Entende-se que tais critérios adotados para inclusão e exclusão realizariam filtro relevante, utilizando somente material atual, em linguagem compreensível pelos autores e descartando-se materiais onde há a possibilidade de não revisão por pares.

### 3.3.2 Planejamento de avaliação da qualidade

Com o intuito de avaliar a qualidade dos artigos para seleção, foi planejado o uso do método PRISMA (MOHER, 2009). O método PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) é um guia para relatar revisões sistemáticas e metanálises. O método envolve a identificação de estudos relevantes, a extração e síntese dos dados e a avaliação da qualidade dos estudos incluídos.

## 3.4 Condução

### 3.4.1 Busca e importação de estudos



A plataforma Parsifal permite a realização de buscas diretamente nas bases de dados *Scopus* e *IEEE*. No entanto, devido a restrições de acesso em rede privada, as buscas foram realizadas separadamente no laboratório de informática do Centro Estadual de Educação Tecnológica Paula Souza e posteriormente importadas para o Parsifal. Os resultados das buscas em cada base estão apresentados na Tabela 1.

**Tabela 1** Artigos obtidos

Base	Artigos encontrados
<i>Scopus</i>	2479
IEEE	1218

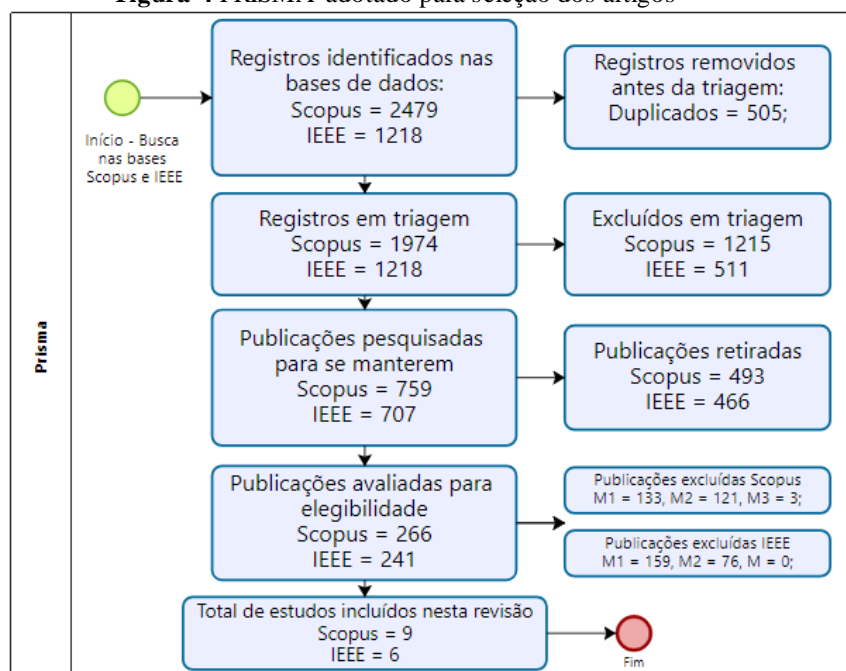
Fonte: os autores (2023)

Foi observado que a maioria dos artigos obtidos foi proveniente da base de dados *Scopus* e a segunda maior foi a *IEEE*.

### 3.4.2 Seleção dos artigos

Apesar de a plataforma Parsifal apresentar seu protocolo para seleção de artigos por meio de perguntas, optou-se por utilizar a ferramenta Prisma para tornar clara a significativa redução e refinamento dos artigos buscados nas bases *Scopus* e *IEEE*, que foram necessários para chegar aos artigos selecionados para esta Revisão Sistemática da Literatura (RSL). A Figura 4 demonstra o caminho adotado conforme a ferramenta Prisma para a seleção dos artigos.

**Figura 4** PRISMA adotado para seleção dos artigos



Fonte: os autores (2023)

A Figura 4 demonstra que, inicialmente, as buscas foram realizadas nas bases *Scopus* e IEEE, resultando em 2.479 artigos na base *Scopus* e 1.218 na base IEEE. Desses, 505 foram identificados como duplicatas entre as bases, sendo removidos. Em seguida, durante a triagem, 1.215 artigos da base *Scopus* e 511 da base IEEE foram excluídos, com base na leitura de títulos e resumos, por não estarem diretamente relacionados ao tema ou por não serem estudos de caso adequados para ajudar a responder à questão de pesquisa.

Posteriormente, 493 publicações da base *Scopus* e 466 da base IEEE foram removidas, uma vez que estavam restritas para consulta gratuita, redirecionavam para páginas onde não havia *links* para *download* ou o *download* demorava excessivamente para ser concluído. Em seguida, os 266 artigos da base *Scopus* e 241 da base IEEE foram divididos entre os autores para leitura do texto integral. No entanto, devido ao grande número de publicações que chegaram a essa etapa no Prisma e considerando o tempo disponível para a elaboração do artigo, optou-se por realizar um refinamento adicional, o que resultou na exclusão dos artigos:

- que não apresentaram resultados claros, chamados de motivo 1 (M1, na figura 4);
- que estavam mais inclinados à área computacional do que à aplicação na indústria, chamados de motivo 2 (M2, na figura 4);

Além disso, por causa do motivo 3 (M3, na Figura 4), optou-se por excluir três artigos da base *Scopus*, uma vez que apresentavam partes do texto ausentes. A Tabela 2 fornece um resumo dos artigos selecionados por base.

**Tabela 2** Artigos selecionados

Base	Artigos selecionados
<i>Scopus</i>	9
IEEE	6

Fonte: os autores (2023)

Pode-se concluir que foram selecionados para revisão 9 artigos da base *Scopus* e 6 da base IEEE.

## 4 RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentados os resultados da Bibliometria e da RSL.

### 4.1 Resultados e Discussão da Bibliometria

Por meio da análise dos dados da Bibliometria, identificou-se um cenário ativo, considerando o crescente número de publicações, colaborações e a participação de vários países no desenvolvimento desse tema.

Com base nos dados analisados na bibliometria, observa-se o seguinte:

- Há um aumento significativo no número de trabalhos científicos relacionados a estudos de caso, *machine learning* e a indústria no período de 2018 a 2022.
- É notável uma diminuição no número de citações no mesmo período, e essa tendência pode ser explorada em futuros estudos.
- O período de 2018 a 2022 apresenta fontes, autores e instituições bastante ativos em termos de publicações.
- Diversas palavras-chave, como *Machine learning*, *Deep learning*, *Case Studies*, *Learning Systems*, *Decision Trees* e *Manufacture*, foram amplamente utilizadas durante esse período.

Além disso, é evidente a diversidade de termos relevantes quando se busca o tema que abrange "estudo de caso, *machine learning* e indústria" em bases como a *Scopus* e a *IEEE*. Estes termos incluem *deep learning*, *Decision Trees*, *neural networks*, *decision making*, *Case Studies*, *study of case*, *manufacture*, entre outros.

#### 4.2 Resultados e Discussão da RSL

Várias foram as contribuições para a indústria, onde estudos de caso sobre *machine learning* foram conduzidos. Fernandes *et al.* (2018) apresentaram a redução do conjunto de recursos durante a análise exploratória do conjunto de dados de uma empresa metalúrgica. Eles empregaram diversos métodos de análise de dados e seleção de recursos, ajudando a identificar relacionamentos e padrões nos dados coletados. Como resultado, conseguiram reduzir o conjunto inicial de 47 recursos para um conjunto de 32 recursos mais relevantes. Essa redução permitiu simplificar a análise e concentrar-se nos recursos mais significativos para o estudo.

Fernandes *et al.* (2018) também apresentaram a criação de um modelo baseado em regras. Com base nas informações e associações identificadas durante a análise exploratória, desenvolveram um modelo baseado em regras que define as ações que o sistema *InValue* deve tomar quando certas condições são atendidas. Esse modelo é composto por um conjunto de máquinas, um conjunto de peças de trabalho, um conjunto de mensagens e um conjunto de regras. Ele permite fornecer informações sobre a condição de parâmetros operacionais específicos, com o objetivo de prevenir a deterioração dos componentes da máquina e maximizar a produtividade.

Lingitz *et al.* (2018) concluíram que o uso do método de estudo de caso contribuiu para uma compreensão aprofundada do problema de previsão do *lead time* da empresa de

semicondutores onde o estudo foi conduzido. Isso permitiu que os autores aplicassem algoritmos de aprendizado de máquina de forma relevante e obtivessem *insights* práticos e aplicáveis para melhorar a eficiência da cadeia de suprimentos nesse contexto específico.

Amihai *et al.* (2018) apresentaram que as principais contribuições do método científico "estudo de caso" no contexto em que o desenvolveram são a validação da abordagem de aprendizado de máquina para prever a condição de ativos industriais usando dados de campo reais e a identificação de desafios específicos e *insights* práticos relacionados à aplicação do aprendizado de máquina em ambientes industriais.

No estudo de Canizo (2019), evidenciou-se que as principais contribuições do método científico "estudo de caso" mencionadas incluem uma nova arquitetura DL para detecção supervisionada de anomalias em sistemas multi-sensores e aprendizagem de transferência para adaptação de modelos a configurações variáveis de sensores. Essas contribuições abordam os desafios da detecção de anomalias em sistemas multi-sensores, como lidar com dados de sensores heterogêneos, reduzir os esforços de pré-processamento de dados, adaptar-se a mudanças nas configurações do sensor e melhorar a precisão e a eficiência da detecção de anomalias.

Márquez, Ramírez e Marugán (2019) deixam claro que o método de estudo de caso contribuiu fornecendo uma estrutura e uma abordagem sistemática para coletar, analisar e interpretar os dados relacionados à fabricação de turbinas eólicas. Isso ajudou os pesquisadores a obter *insights* específicos e a extrair conclusões que podem ser aplicáveis a outros contextos e setores.

No trabalho conduzido por Leyva *et al.* (2020), evidenciou-se que o método do estudo de caso foi utilizado para comparar e analisar diferentes métodos de previsão de demanda para produtos têxteis. Essa comparação permitiu identificar os métodos mais adequados para prever a demanda por produtos têxteis, levando em consideração as mudanças recentes no contexto equatoriano. Uma segunda contribuição percebida foi que o estudo de caso também contribuiu para o desenvolvimento de um *framework* de previsão de demanda para a indústria têxtil estudada. Esse *framework* combina métodos clássicos e de aprendizado de máquina, projetado para melhorar a precisão das previsões de demanda e facilitar a tomada de decisões na indústria têxtil.

No trabalho desenvolvido por Nguyen-duc *et al.* (2020), evidenciou-se a ausência de métricas orientadas por negócios no desenvolvimento de sistemas de IA. O texto afirma que é comum que oportunidades de negócios envolvendo sistemas de IA não sejam validadas e que haja falta de métricas orientadas por negócios para guiar o desenvolvimento desses

sistemas de IA. Isso sugere que uma das principais contribuições do método de estudo de caso é destacar a ausência de métricas orientadas por negócios no desenvolvimento de sistemas de IA. Ao realizar entrevistas e analisar casos do mundo real, os pesquisadores identificaram essa lacuna e enfatizaram a necessidade de métricas alinhadas aos objetivos de negócios para orientar o desenvolvimento de sistemas de IA.

O trabalho desenvolvido por Nguyen-duc *et al.* (2020) também tornou clara a exploração de fatores contextuais e abordagens de desenvolvimento de IA. Isso indica que o método de estudo de caso contribuiu para compreender os fatores contextuais que influenciam o desenvolvimento de IA e identificar diversas abordagens usadas no desenvolvimento de sistemas de IA. Ao examinar vários casos e realizar entrevistas com profissionais, os pesquisadores obtiveram *insights* sobre a diversidade de sistemas de IA e métodos de desenvolvimento, permitindo uma compreensão mais ampla do campo.

Ma *et al.* (2020) apresentaram o desenvolvimento de um modelo de aprendizado de máquina (*machine learning*) para estimar a concentração diária de massa de PM<sub>2.5</sub> em Xangai, China. O estudo combinou previsões de concentração de massa de PM<sub>2.5</sub> de um sistema de previsão numérica da qualidade do ar operacional (WRF-Chem) com medições abrangentes de poluentes atmosféricos e condições meteorológicas para desenvolver um modelo de aprendizado de máquina. Esse modelo proporcionou uma previsão diária significativamente melhor da concentração de PM<sub>2.5</sub> do que o modelo WRF-Chem, com coeficientes de correlação mais altos e desvio padrão mais baixo.

Ma *et al.* (2020) apresentaram ainda o aprimoramento da previsão da qualidade do ar na China, especialmente em relação à concentração de PM<sub>2.5</sub>. Ao utilizar o modelo de aprendizado de máquina desenvolvido, é possível obter previsões mais precisas e confiáveis da concentração diária de PM<sub>2.5</sub>, contribuindo para a redução da poluição do ar por indústrias e seus efeitos adversos na saúde humana.

No trabalho desenvolvido por Papachristou, Chrysopoulos e Bilalis (2020), evidenciou-se a possibilidade de consideração da segurança em roupas infantis. O estudo menciona o mercado de roupas infantis como o de crescimento mais rápido e um dos mais lucrativos na indústria do vestuário. As crianças que consomem produtos de roupas têm consciência de moda e demandam roupas de designers mais *fashion*. No entanto, além da aparência, a segurança é um requisito fundamental que as roupas infantis devem atender. A aplicação de ferramentas de inteligência artificial no processo de geração de designs ou tomada de decisão pode ajudar a prever quais produtos serão mais atraentes e vendáveis.

No entanto, é necessário levar em consideração requisitos de conforto, ajuste para brincadeiras, descanso e segurança. O estudo buscou contribuir identificando os desafios que as aplicações de tecnologia de inteligência artificial enfrentam no que diz respeito às roupas infantis, considerando as restrições e padrões de segurança, utilizando um fabricante e varejista grego de roupas infantis como estudo de caso.

No estudo desenvolvido por Kabugo *et al.* (2020), evidenciou-se a implementação de *soft* sensores baseados em dados, onde notou-se um estudo de caso realizado em uma usina de energia a partir de resíduos (WTE, na sigla em inglês). O estudo investigou *soft* sensores baseados em dados para prever o valor de aquecimento do gás de síntese (*syngas*) e a temperatura do gás de exaustão quente. Entre os métodos baseados em dados estudados, um modelo NARX baseado em redes neurais demonstrou melhor desempenho na previsão tanto do valor de aquecimento do *syngas* quanto da temperatura do gás de exaustão. O uso de *soft* sensores baseados em dados permite a análise preditiva de dados e prova ser uma ferramenta útil quando o conhecimento do processo é limitado.

No trabalho de Sacco *et al.* (2020), evidenciou-se que o método de estudo de caso contribuiu fornecendo uma estrutura e uma abordagem sistemática para a coleta de dados, aplicação de técnicas de *machine learning*, avaliação dos resultados e discussão das implicações práticas. Isso ajudou os pesquisadores a investigar e apresentar a viabilidade e a eficácia do uso de *machine learning* na inspeção automatizada de colocação de fibras em manufatura de compósitos.

No trabalho desenvolvido por Wang *et al.* (2021), verificou-se que o método de estudo de caso propiciou a aplicação prática do *deep learning* em um problema real de fabricação. O estudo utilizou o aprendizado profundo (*deep learning*) para resolver um problema específico relacionado à soldagem na indústria. Ele fornece um exemplo concreto de como o *deep learning* pode ser aplicado no monitoramento de processos e na previsão da qualidade do produto em um contexto de fabricação.

O trabalho de Behdinian *et al.* (2022) permitiu a identificação dos fatores que impactam o sucesso do projeto. O artigo demonstrou que o objetivo era identificar os fatores que impactam o sucesso do projeto e apontar os fatores mais relevantes para facilitar a implementação do projeto. Ele destacou a utilização de algoritmos de *Machine learning* e modelos de simulação para detectar a eficácia dos principais fatores organizacionais no sucesso do projeto. Por meio do uso de uma regressão logística, um modelo preditivo foi criado com variáveis independentes para prever se o projeto de *software* seria bem-sucedido ou fracassaria. Além disso, a importância do recurso de regressão logística identificou a

gamificação como o fator mais influente no objetivo. Em seguida, os modelos gamificados e não gamificados foram comparados pelo método de simulação, mostrando que a gamificação proporcionou uma melhoria de 36,26% no tempo do ciclo de projetos e um aumento de 15% na qualidade do desempenho dos funcionários, reduzindo os erros dos projetos. Isso demonstra como o método de estudo de caso pode ajudar a identificar e avaliar os fatores que contribuem para o sucesso do projeto.

No estudo conduzido por Gaugel e Reichert (2023), evidenciou-se a análise prática dos benefícios do uso de transferência de aprendizado (TL) para segmentação de séries temporais (TSS) em um ambiente industrial. O artigo fornece um dos primeiros estudos a abordar a TL para TSS. Ele analisa sistematicamente como o pré-treinamento da rede com diferentes conjuntos de dados de origem, variando em similaridade com o conjunto de dados de destino, afeta o desempenho do modelo de destino após o ajuste fino.

Após a análise dos artigos selecionados, os ganhos identificados com o uso de estudos de caso sobre *machine learning* na indústria foram expressos na Matriz 1:

**Matriz 1** Artigos por grupo de ganho

continua

Artigo	Ganhos em segurança e confiabilidade	Ganhos em eficiência e produtividade	Ganhos em tomada de decisão e gestão de projetos	Ganhos em detecção e prevenção de anomalias e falhas	Ganhos em análise de dados e segmentação de séries temporais	Ganhos em previsão e análise de dados	Ganhos em segurança e combate à falsificação
<i>Machine learning in composites manufacturing: A case study of Automated Fiber Placement inspection</i>		X					
<i>Multi-head CNN-RNN for multi-time series anomaly detection: An industrial case study</i>		X		X			
<i>A tutorial on deep learning-based data analytics in manufacturing through a welding case study</i>		X	X				
<i>Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer</i>	X					X	
<i>A Multiple Case study of Artificial Intelligent System Development in Industry</i>			X				
<i>An Integrating Machine learning Algorithm and Simulation Method for Improving Software Project Management: A Case study</i>	X		X				
<i>An Industrial Case study Using Vibration Data and Machine learning to Predict Asset Health</i>	X			X			

**Matriz 1** Artigos por grupo de ganho

continuação

Artigo	Ganhos em segurança e confiabilidade	Ganhos em eficiência e produtividade	Ganhos em tomada de decisão e gestão de projetos	Ganhos em detecção e prevenção de anomalias e falhas	Ganhos em análise de dados e segmentação de séries temporais	Ganhos em previsão e análise de dados	Ganhos em segurança e combate à falsificação
<i>Application case study of Machine learning techniques towards a Fault Diagnosis System for a Manufacturing Plant Environment</i>			X	X		X	
<i>Application of the XGBoost Machine learning Method in PM2.5 Prediction: A Case study of Shanghai</i>			X			X	
<i>Data analysis and feature selection for predictive maintenance: A case-study in the metallurgic industry</i>	X			X			
<i>Decision making using Logical Decision Tree and Binary Decision Diagrams: A Real Case study of Wind Turbine Manufacturing</i>			X				
<i>Industrial Transfer Learning for Multivariate Time Series Segmentation: A Case study on Hydraulic Pump Testing Cycles</i>					X		
<i>Industry 4.0 based process data analytics platform: A waste-to-energy plant T case study</i>		X					
<i>Machine learning for clothing manufacture as a mean to respond quicker and better to the demands of clothing brands: a Greek case study</i>		X					
<i>Novel Counterfeit Detection of Integrated Circuits via Infrared Analysis: A Case study based on the Intel Cyclone II FPGAs</i>	X						X

**Fonte:** os autores (2023)

Ganhos em segurança e confiabilidade: Referem-se aos benefícios obtidos ao aprimorar a segurança e a confiabilidade de um sistema, processo ou produto. Isso pode incluir a detecção e mitigação de riscos, a identificação de vulnerabilidades e a implementação de medidas de segurança sólidas e confiáveis, entre outros. Foram identificados 5 relatos.

Ganhos em eficiência e produtividade: Refletem os benefícios derivados do aumento da eficiência e produtividade de um sistema, processo ou operação. Isso pode envolver a automatização de tarefas, a otimização de fluxos de trabalho, a redução de



desperdícios e o aprimoramento na utilização de recursos, entre outros. Foram identificados 5 relatos.

Ganhos em tomada de decisão e gestão de projetos: Referem-se aos benefícios alcançados ao utilizar análises de dados e informações para embasar decisões estratégicas e para gerenciar projetos de forma eficaz. Isso envolve o uso de métodos e ferramentas para analisar dados, visualizar informações relevantes, identificar tendências, avaliar riscos, entre outros. Foram identificados 6 relatos.

Ganhos em detecção e prevenção de anomalias e falhas: Refletem os benefícios obtidos ao aplicar técnicas de análise de dados para identificar e prevenir anomalias e falhas em sistemas, processos ou produtos. Isso pode envolver o uso de algoritmos de detecção de anomalias, monitoramento contínuo, alertas automáticos, manutenção preditiva, entre outros. Foram identificados 4 relatos.

Ganhos em análise de dados e segmentação de séries temporais: Referem-se aos benefícios obtidos ao realizar análises de dados e segmentar séries temporais para compreender padrões, tendências e comportamentos ao longo do tempo. Isso pode incluir a aplicação de algoritmos de análise de séries temporais, identificação de pontos de mudança, segmentação de dados, visualização de padrões, entre outros. Foi identificado 1 relato.

Ganhos em previsão e análise de dados: Refletem os benefícios alcançados ao utilizar técnicas de previsão e análise de dados para obter *insights* e embasar decisões. Isso pode envolver a aplicação de algoritmos de aprendizado de máquina, modelos estatísticos, análise exploratória de dados, visualização de informações relevantes, entre outros. Foram identificados 3 relatos.

Ganhos em segurança e combate à falsificação: Referem-se aos benefícios obtidos ao fortalecer a segurança e combater a falsificação de produtos, documentos ou sistemas. Isso pode envolver o uso de tecnologias de autenticação, criptografia, rastreabilidade, métodos de detecção de falsificações, entre outros. Foi identificado 1 relato.

## 5 CONSIDERAÇÕES FINAIS

O método científico de estudo de caso tem se mostrado relevante e aplicável na pesquisa que envolve *machine learning* no segmento industrial. O estudo de caso permite uma análise aprofundada, contextualizada e específica dos desafios enfrentados pelas indústrias, oferecendo benefícios, tais como a melhoria da segurança, o aumento da eficiência e produtividade, o apoio à tomada de decisões e à gestão de projetos, a detecção e

prevenção de anomalias e falhas, a análise de dados e segmentação de séries temporais, previsão e análise de dados, além do fortalecimento da segurança e do combate à falsificação.

Por meio da revisão sistemática da literatura e dos resultados da entrevista, foi possível verificar a aderência das opiniões dos respondentes às ideias dos autores levantados, corroborando o cenário identificado pela bibliometria. Essas conclusões contribuem para o entendimento das contribuições do método de estudo de caso aplicado ao *machine learning* no segmento industrial no período de 2018 a 2022, que era o questionamento de pesquisa. No entanto, é importante ressaltar a necessidade de considerar as particularidades de cada estudo de caso, evitando generalizações além do contexto analisado.

Portanto, a pesquisa alcançou seus objetivos ao mapear as contribuições do método de estudo de caso aplicado ao *machine learning* no segmento industrial, apresentar o cenário de publicações sobre os temas agrupados, identificar os principais grupos de ganhos por meio da revisão sistemática da literatura e coletar declarações sobre as principais contribuições em *machine learning* no segmento industrial por meio da entrevista.

A pesquisa fica limitada ao recorte de 2018 a 2022 e à impossibilidade de generalização dos estudos de caso. Para futuros trabalhos, sugere-se focar em segmentos industriais específicos, como indústrias alimentícias, têxteis, metalúrgicas e outras.

## REFERÊNCIAS

- AMIHAI, I.; PARESCI, D.; GITZEL, R.; SUBBIAH, S.; KOTRIWALA, A.; SOSALE, G. *An Industrial Case study Using Vibration Data and Machine learning to Predict Asset Health*. In: *2018 IEEE 20th Conference on Business Informatics*. [S.l.]: IEEE, 2018. p. 28-35. DOI: 10.1109/CBI.2018.00028.
- BEHDINIAN, A.; AMANI, M. A.; AGHSAMI, A.; JOLAI, F. *An Integrating Machine learning Algorithm and Simulation Method for Improving Software Project Management: A Case study*. *Journal of Quality Engineering and Production Optimization*, v. 7, n. 1, p. 1214-1223, Winter & Spring 2022. Available at: <http://jqepo.shahed.ac.ir>. DOI: 10.22070/JQEPO.2022.15346.1214.
- CANIZO, M. *Multi-head CNN-RNN for multi-time series anomaly detection: An industrial case study*. *Neurocomputing*, v. 363, p. 20-31, 2019.
- CUNHA, L. J.; VALIM, P. R. O. **Rede de Sensores Sem Fio para Monitoramento de Variáveis de Ambiente**. In: *XI Computer on the Beach*, 2 a 4 de Setembro de 2020, Baln. Camboriú, SC, Brasil. Universidade do Vale do Itajaí - UNIVALI.
- FERNANDES, M.; CANITO, A.; CANEDO, V. B.; CONCEIÇÃO, L. *et al. Data analysis and feature selection for predictive maintenance: A case-study in the metallurgic industry*. *International Journal of Information Management*, v. 46, p. 252-262, 2019. doi:10.1016/j.ijinfomgt.2018.10.006.
- FINK, A. **Conducting research literature reviews: From the Internet to paper** (2nd ed.). Thousand Oaks: Sage, 2005.

- FRAZIER, P., D. *Novel counterfeit detection of integrated circuits via infrared analysis: A case study based on the Intel Cyclone II FPGAs*. In: *Proceedings of the 2016 International Conference on Machine learning and Cybernetics, Jeju, South Korea, 10-13 July, 2016*.
- GALLUP, G. *The Gallup Poll: Public Opinion 1935-1971*. New York: Random House, 1972.
- GARCÍA, M., FAUSTO, P. *Decision making using Logical Decision Tree and Binary Decision Diagrams: A Real Case study of Wind Turbine Manufacturing*. *Energies*, v. 12, n. 9, p. 1753, 2019. DOI: 10.3390/en12091753.
- GAUGEL, S., REICHERT, M. *Industrial Transfer learning for Multivariate Time Series Segmentation: A Case study on Hydraulic Pump Testing Cycles*. *Sensors*, 23, 3636. doi: 10.3390/s23073636.
- GLÄNZEL, W. *Bibliometrics as a Research Field: A Course on Theory and Application of Bibliometric Indicators*. Universiteit Antwerpen, 1996.
- KABUGO, J. C., JÄMSÄ-JOUNELA, S.-L., SCHIEMANN, R., BINDER, C. *Industry 4.0 based process data analytics platform: A waste-to-energy plant case study*. *Electrical Power and Energy Systems*, 115, 105508. doi: 10.1016/j.ijepes.2019.105508.
- LINGITZ, L.; GALINA, V.; ANSARI, F.; GYULAI, D. *et al. Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer*. *Procedia CIRP*, v. 72, p. 1006-1011, 2018.
- LEYVA, L. L. *et al. A Comparison of Machine Learning and Classical Demand Forecasting Methods: A Case Study of Ecuadorian Textile Industry*. In: *6th International Conference, LOD 2020, Siena, Italy, July 19–23, 2020, Revised Selected Papers, Part II*. Switzerland: Springer Nature, 2020. p. 131–142. (LNCS, 12566). Disponível em: [https://doi.org/10.1007/978-3-030-64580-9\\_11](https://doi.org/10.1007/978-3-030-64580-9_11).
- MA, J.; YU, Z.; QU, Y.; XU, J.; CAO, Y. *Application of the XGBoost Machine learning Method in PM2.5 Prediction: A Case study of Shanghai*. *Aerosol and Air Quality Research*, 20, 128-138. doi: 10.4209/aaqr.2019.08.0408.
- MAHESH, B. *Machine learning Algorithms - A Review*. *International Journal of Science and Research (IJSR)*, v. 9, n. 1, p. [381-386], jan. 2020. ISSN 2319-7064. Available at: <[www.ijsr.net](http://www.ijsr.net)>. Access on: [12/06/2023].
- MÁRQUEZ, F. P. G.; RAMÍREZ, I.; MARUGÁN, A.. *Decision making using logical decision tree and binary decision diagrams: a real case study of wind turbine manufacturing*. *Energies*, v. 12, n. 9, p. 1753, 2019.
- MERRIAM, S. B. *Qualitative Research and Case study Applications in Education*. 1998, San Francisco, CA: Jossey-Bass.
- MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN D., G. *et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. *PLoS Med.*, v. 6, n. 7, p. e1000097, 2009. doi: 10.1371/journal.pmed.1000097.
- NAVARRO, F., P. *Big Data no contexto de dados acadêmicos: o uso de machine learning na construção de sistema de organização do conhecimento*. *Informação & Tecnologia (ITEC)*, Marília/João Pessoa, v. 5, n. 2, p. 181-200, jul./dez. 2018. Universidade Estadual Paulista Júlio Mesquita Filho – UNESP.
- NGUYEN-DUC, A.; SUNDBO, I.; NASCIMENTO, E.; CONTE, T. *et al. A Multiple Case study of Artificial Intelligent System Development in Industry*. In: *Proceedings of the*

*Evaluation and Assessment on Software Engineering*, v. 1, p. 1-10, 2020. DOI: 10.1145/3383219.3383220.

OKOLI, C.. *A guide to conducting a standalone systematic literature review*. *Communications of the Association for Information Systems*, v. 37, n. 43, p. 879–910, nov. 2015. Disponível em: <http://aisel.aisnet.org/cais/vol37/iss1/43/>.

PAPACHRISTOU, E.; CHRYSOPOULOS, A.; BILALIS, N. **Machine learning for clothing manufacture as a mean to respond quicker and better to the demands of clothing brands**: a Greek case study. *The International Journal of Advanced Manufacturing Technology*. doi: 10.1007/s00170-020-06157-1.

SACCO, C.; RADWAN, A., B.; ANDERSON, A.; HARYK, R. *et al.* **Machine learning in composites manufacturing**: A case study of Automated Fiber Placement inspection. *Composite Structures*, v. 250, p. 112514, 2020.

SCHNEIDER, P. H. **Análise preditiva de Churn com ênfase em técnicas de Machine learning**: Uma Revisão. 2016. 82 - Fundação Getúlio Vargas, Escola de Matemática Aplicada. Rio de Janeiro, 2016.

STAKE, R. E. *The Art of Case study Research*. Thousand Oaks, CA: 1995, Sage Publications.

WANG, Q.; JIAO, W.; WANG, P.; ZHANG, Y. *A tutorial on deep learning- based data analytics in manufacturing through a welding case study*. 2021, *Journal of Manufacturing Processes*. DOI: 10.1016/j.jmapro.2020.04.044

YIN, R. K. *Case study Research and Applications: Design and Methods*. Thousand Oaks, CA: 2001, Sage Publications.