

UNIVERSIDADE DO PLANALTO CATARINENSE – UNIPLAC

UNIVERSIDADE DO CONTESTADO – UNC

UNIVERSIDADE DO EXTREMO SUL CATARINENSE – UNESC

UNIVERSIDADE DA REGIÃO DE JOINVILLE – UNIVILLE

PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS PRODUTIVOS - PPGSP

MARIA JOANA MARTINS

CIÊNCIA DE DADOS APLICADA À TRANSFORMAÇÃO DIGITAL DA GESTÃO DE INVENTÁRIO EM UMA
INDÚSTRIA CARBONÍFERA

LAGES/SC

2024

MARIA JOANA MARTINS

**CIÊNCIA DE DADOS APLICADA À TRANSFORMAÇÃO DIGITAL DA GESTÃO DE INVENTÁRIO EM UMA
INDÚSTRIA CARBONÍFERA**

Dissertação de Mestrado, vinculada ao Programa de Pós-graduação em Sistemas Produtivos – PPGSP em forma associativa entre UNIPLAC, UNC, UNESC e UNIVILLE, como requisito parcial para obtenção do título de Mestre(a) em Sistemas Produtivos.

Orientador(a): Profa. Dra. Merisandra Côrtes de Mattos

Coorientador(a): Prof. Dr. Marcelo Leandro de Borba

LAGES/SC

2024

Ficha Catalográfica

M379c

Martins, Maria Joana

Ciência de dados aplicada à transformação digital da gestão de inventário em uma indústria carbonífera / Maria Joana Martins ; orientadora Dra. Merisandra Côrtes de Mattos ; coorientador Dr. Marcelo Leandro de Borba. – 2024.

119 f. ; 30 cm.

Dissertação (Mestrado em Sistemas Produtivos) – Programa de Pós-Graduação em Sistemas Produtivos em forma associativa entre a Universidade do Planalto Catarinense ; Universidade do Contestado ; Universidade do Extremo Sul Catarinense ; Universidade da Região de Joinville. Lages, SC, 2024.

1. Inteligência artificial. 2. Mineração de dados (Computação). 3. Associação. 4. Manutenção. 5. Indústria carbonífera. I. Mattos, Merisandra Côrtes de (orientadora). II. Borba, Marcelo Leandro de (coorientador). III. Universidade do Planalto Catarinense. IV. Universidade do Contestado. V. Universidade do Extremo Sul Catarinense VI. Universidade da Região de Joinville. VII. Programa de Pós-Graduação em Sistemas Produtivos. VIII. Título.

CDD 658.5

FOLHA DE APROVAÇÃO

MARIA JOANA MARTINS

**CIÊNCIA DE DADOS APLICADA À TRANSFORMAÇÃO DIGITAL DA GESTÃO DE INVENTÁRIO EM UMA
INDÚSTRIA CARBONÍFERA**

A Comissão Examinadora, abaixo assinada, aprova a Dissertação apresentada no Programa de Pós-Graduação em Sistemas Produtivos – PPGSP, Linha de Pesquisa Gestão e Conhecimento em Sistemas Produtivos, em forma associativa entre a Universidade do Planalto Catarinense – UNIPLAC, a Universidade do Contestado – UNC, a Universidade do Extremo Sul Catarinense – UNESC e a Universidade da Região de Joinville - UNIVILLE, como requisito parcial para obtenção do título de **Mestre em Sistemas Produtivos**.

Banca Examinadora

Documento assinado digitalmente



MERISANDRA CORTES DE MATTOS

Data: 27/02/2025 13:23:25-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dra. Merisandra Côrtes de Mattos / UNESC
Presidente da Banca / Orientador(a)

Documento assinado digitalmente



MARCELO LEANDRO DE BORBA

Data: 27/02/2025 07:52:43-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Marcelo Leandro de Borba / UNIVILLE
Coorientador(a)



Documento assinado digitalmente

Cristian Cechinel

Data: 27/02/2025 08:34:35-0300

CPF: ***.628.449-**

Verifique as assinaturas em <https://v.ufsc.br>

Prof. Dr. Cristian Cechinel
Membro externo da banca

Documento assinado digitalmente



JACIR FAVRETTO

Data: 27/02/2025 13:08:31-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Jacir Favretto
Membro interno da banca

Lages, SC, 25, junho e 2024

DEDICATÓRIA

Dedico essa dissertação a minha família, minha esposa Suelen, minha mãe Edilene, ao meu pai Adailton, por me apoiarem incansavelmente nessa trajetória. E aos meus professores que mesmo não me orientando diretamente me deram o suporte necessário para conseguir alcançar meus objetivos acadêmicos, Leopoldo Guimarães Filho, Vilson Bristot, Kristian Madeira. Por conta de vocês me torno cada dia uma profissional melhor.

AGRADECIMENTOS

Primeiramente agradeço a Deus, por me permitir ter saúde para elaborar uma dissertação e me abençoar com pessoas incríveis o tempo todo. Gostaria de agradecer também a minha esposa por ter me amparado em todo o meu período de estudo e dedicação ao mestrado, me fazendo acreditar que sou capaz. Agradeço aos meus pais, por não pararem de acreditar em mim e no meu potencial, sempre me incentivando a ser e me tornar alguém melhor.

Sigo com os meus pais profundos agradecimento a minha orientadora, que ressignificou o prazer de lecionar para mim, e sempre me deu suporte independente do horário e do dia ao longo desses 2 anos, suporte esse, como orientadora e amiga, me acalmando em dias difíceis e me guiando em dias produtivos.

- À Universidade do Planalto Catarinense – UNIPLAC, a Universidade do Contestado – UNC, a Universidade do Extremo Sul Catarinense – UNESC e a Universidade da Região de Joinville – UNIVILLE, que me oportunizaram a participação e aquisição de conhecimento no Programa de Pós-Graduação em Sistemas Produtivos;

- À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pela contribuição e investimentos na formação de recursos de alto nível por meio da pesquisa científica e estímulo na consolidação da pós-graduação no País;

- À Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina – FAPESC.

RESUMO

MARTINS, Maria Joana. **Ciência de dados aplicada à transformação digital da gestão de inventário em uma indústria carbonífera.** (2024), 119 f. (*número de folhas*). Dissertação (Mestrado em Sistemas Produtivos). Programa de Pós-graduação em Sistemas Produtivos – PPGSP em forma associativa entre a Universidade do Planalto Catarinense – UNIPLAC, da Universidade do Contestado – UNC, da Universidade do Extremo Sul Catarinense – UNESC e da Universidade da Região de Joinville – UNIVILLE, Lages, (2024).

A cadeia de suprimentos é responsável pela gestão de matéria-prima, componentes do processo e inventário de suprimentos de manutenção, garantindo a continuidade da produção e a conexão direta com os resultados empresariais, evitando interrupções no processo produtivo devido a falhas nos equipamentos. No contexto da manutenção, a gestão de suprimentos deve assegurar a disponibilidade dos itens necessários para a execução eficiente das atividades de reparo. Na gestão de inventário, ferramentas de apoio podem ser utilizadas para prever demandas futuras e armazenar a quantidade adequada de materiais. Assim, esta pesquisa visa analisar como a ciência de dados pode impulsionar a transformação digital na gestão de inventário de manutenção em uma indústria carbonífera. Para a concepção do estudo, os métodos de tratamento de dados foram divididos em três partes: captura dos dados do inventário de manutenção de uma empresa de extração mineral e pré-processamento do conjunto de dados, execução da mineração de dados utilizando o algoritmo Apriori para realizar as relações no conjunto de dados, e pós-processamento dos dados e análise de valor com base nas regras de associação identificadas. A etapa final apresenta os resultados obtidos, incluindo métricas estatísticas de IA que identificam componentes e quantidades ideais para a gestão de estoque. A implementação dessas etapas permitiu identificar os conhecimentos necessários para a gestão do inventário no setor de manutenção hidráulica em uma indústria carbonífera.

Palavras-chave: Inteligência artificial; Mineração de dados; Associação; Análise de valor; Manutenção; Indústria carbonífera.

ABSTRACT

MARTINS, Maria Joana. **Data science applied to the digital transformation of inventory management in a coal industry.** (2024), 119 f. (*número de folhas*). Dissertação (Mestrado em Sistemas Produtivos). Programa de Pós-graduação em Sistemas Produtivos – PPGSP em forma associativa entre a Universidade do Planalto Catarinense – UNIPLAC, da Universidade do Contestado – UNC, da Universidade do Extremo Sul Catarinense – UNESC e da Universidade da Região de Joinville – UNIVILLE, Lages, (2024).

The supply chain is responsible for managing raw materials, process components, and maintenance supply inventory, ensuring production continuity and a direct connection to business outcomes by preventing interruptions in the production process due to equipment failures. In the context of maintenance, supply management must ensure the availability of necessary items for the efficient execution of repair activities. In inventory management, support tools can be used to predict future demands and store the appropriate quantity of materials. Thus, this research aims to analyze how data science can drive digital transformation in maintenance inventory management within a coal industry. For the study's design, data processing methods were divided into three parts: capturing maintenance inventory data from a mineral extraction company and preprocessing the dataset, executing data mining using the Apriori algorithm to identify relationships within the dataset, and post-processing the data and value analysis based on the identified association rules. The final stage presents the obtained results, including AI statistical metrics that identify ideal components and quantities for inventory management. The implementation of these steps allowed for the identification of the necessary knowledge for inventory management in the hydraulic maintenance sector of a coal industry.

Palavras-chave: Artificial intelligence; Data mining; Association; Value analysis; Maintenance; Coal industry.

LISTA DE FIGURAS

Figura 1	<i>Representação do conceito interdisciplinar do PPGSP</i>	22
Figura 2.	<i>Figura explicativa sobre os campos que compõe a ciência de dados e como o campo trabalho no meio interdisciplinar para o processamento do conhecimento</i>	23
Figura 3.	<i>Publicações por períodos</i>	29
Figura 4.	<i>Concentração de publicações nos países</i>	33
Figura 5.	<i>País de origem das publicações dos autores e coautores</i>	33
Figura 6.	<i>Apresentação das ferramentas que integram a ciência de dados</i>	41
Figura 7.	<i>Perspectiva da ciência de dado para a extração do conhecimento de bases de dados</i>	43
Figura 8.	<i>Capacidades de armazenamento e computacional entre 1986 e 2014</i>	44
Figura 9.	<i>Plataforma da ciência de dados</i>	46
Figura 10.	<i>Tarefas executadas pela mineração de dados</i>	47
Figura 11.	<i>Fluxo de informação no modelo de classificação</i>	49
Figura 12.	<i>As imagens representam os vizinhos mais próximos de um ponto no centro, o raio do círculo que abrange os vizinhos é definido por k</i>	50
Figura 13.	<i>Algoritmo Apriori</i>	51
Figura 14.	<i>Demonstração da relação entre os conjuntos de associação</i>	52
Figura 15.	<i>Representação da utilização da técnica K-means para localizar informações similares no espaço</i>	53
Figura 16.	<i>Demonstração do agrupamento hierárquico via dendrograma</i>	53
Figura 17.	<i>Comportamento do fluxo de suprimentos de manutenção</i>	68
Figura 18.	<i>Tratamento da base de dados de requisições de materiais</i>	71
Figura 19.	<i>Tratamento da base de dados de Ordens de serviço</i>	72
Figura 20.	<i>Estrutura da base de dados</i>	73
Figura 21.	<i>Gráfico de dispersão do conjunto de dados</i>	76
Figura 22.	<i>Gráfico de histograma de cadastros que requisitavam componentes no almoxarifado</i>	77
Figura 23.	<i>Exemplos de imputação de dados</i>	79
Figura 24.	<i>Cilindro hidráulico de elevação da cabeça de corte do minerador JOY</i>	87

Figura 25. <i>Algoritmo Apriori aplicado para associação de peças no almoxarifado e componentes.</i>	88
Figura 26. <i>Algoritmo Apriori aplicado para associação de peças no almoxarifado e equipamentos.</i>	91
Figura 27. <i>Tempo entre início da manutenção x requisição da peça 135072 no almoxarifado.</i>	98
Figura 28. <i>Tempo entre início da manutenção x requisição da peça 48735 no almoxarifado.</i>	99
Figura 29. <i>Número de estudos por domínio de aplicação</i>	101

LISTA DE TABELAS

Tabela 1. <i>Apresentação do número de artigos encontrados nas bases de dados, sobre diferentes critérios</i>	28
Tabela 2. <i>Tipo de publicações dos artigos</i>	29
Tabela 3. <i>Revistas nas quais os estudos foram publicados</i>	30
Tabela 4. <i>Estudos que demonstram o maior número de citações</i>	32
Tabela 5. <i>Instituições presentes nas publicações</i>	34
Tabela 6. <i>Frequência das palavras chaves</i>	35
Tabela 7. <i>Estrutura dos relatórios extraído do sistema de gestão da empresa</i>	69
Tabela 8. <i>Apresentação das seis associações com os maiores lift</i>	96
Tabela 9. <i>Apresentação das seis associações com os maiores lift</i>	97

LISTA DE ABREVIATURAS E SIGLAS

ANN	Rede Neural Artificial
API	Interface de programação de aplicações
CART	Classification and Regression Tree
FL	Lógica Fuzzy
FP-Growth	Frequent Pattern Growth
HOG	Histogram of Oriented Gradients
IA	Inteligência artificial
k-NN	k-Nearest Neighbor
LHS	Lado esquerdo ou left-hand side.
MAD	Mean Absolute Deviation
MAE	Erro Absoluto Médio
MSE	Mean Squared Error
ODS	Objetivo de Desenvolvimento Sustentável
PPGSP	Programa de pós-graduação em sistemas produtivos
RF	Random Forest
RFID	Identificação por radiofrequência
RHS	Lado direito ou right-hand side.
RMSE	Erro Quadrado Médio Médio
SVM	Vetor de suporte
WSN	Wireless sensor network

SUMÁRIO

1 INTRODUÇÃO	14
1.2 OBJETIVOS.....	18
1.2.1 Objetivo geral	18
1.2.2 Objetivos específicos	18
1.3 JUSTIFICATIVA	18
1.3.1 Inserção social	20
1.4 CARACTERIZAÇÃO INTERDISCIPLINAR E ADERÊNCIA AO PROGRAMA	21
1.5 ESTRUTURA GERAL DO DOCUMENTO.....	24
2 PANORAMA TEÓRICO	26
2.1 REVISÃO BIBLIOGRÁFICA.....	26
2.1.1 Uma análise bibliométrica de decisão orientada por mineração de dados na gestão de inventário.....	26
2.2 GESTÃO DE INVENTÁRIO	36
2.2.1 Gestão de inventário e inteligência artificial	38
2.2.2 Gestão de inventário e ciência de dados.....	39
2.3 TRANSFORMAÇÃO DIGITAL.....	41
2.4 CIÊNCIA DE DADOS.....	43
2.5 MINERAÇÃO DE DADOS	46
2.5.1 Classificação	48
2.5.2 Associação	50
2.5.3 Agrupamento.....	52
2.6 ESTADO DA ARTE.....	55
3 METODOLOGIA DA PESQUISA	61
3.1 ABORDAGEM, OBJETIVOS, PROCEDIMENTOS E TÉCNICAS DE PESQUISA	61
3.2 DELIMITAÇÃO DO ESTUDO	63
3.2.1 Descrição do local e população em estudo	63
3.2.2 Etapas da pesquisa	64
3.3 PROCESSO DE COLETA DE DADOS.....	67
3.4 PROCESSO DE ANÁLISE DE DADOS.....	69
3.4.1 Pré-processamento dos dados	70
3.4.2 Mineração de dados	83
3.4.3 Pós-processamento dos dados.....	92
3.5 PROCEDIMENTOS ÉTICOS DA PESQUISA.....	93
4 RESULTADOS E DISCUSSÃO	95
4.1. ACHADOS CIENTÍFICOS.....	95
5 CONSIDERAÇÕES FINAIS.....	103

1 INTRODUÇÃO

As mudanças nos sistemas produtivos são derivadas do processo envoltório da indústria e sociedade, agregada a transformação digital, que está sendo introduzida nas tarefas rotineiras e as tem aperfeiçoado (Tran-Dang & Kim, 2021). No universo corporativo possibilita o reposicionamento da empresa em nível estratégico, oportunizando o trabalho com ferramentas tecnológicas para desencadear o fluxo de conhecimento entre as operações, como a comunicação e a conectividade (C. H. Lee et al., 2021).

A transformação digital é uma nova forma de executar as atividades e permitir a tomada de decisão sobre os dados gerados no processo, o que anteriormente era considerada uma prática com alto valor agregado. Está mudando o formato de trabalho, principalmente no que se refere a gestão das empresas.

A forma como a organização constrói sua estratégia tem a ver com a transformação digital, pois fomenta novas maneiras de pensar e exige uma mentalidade estratégica como evidência primordial no papel de liderança (Rogers & Serra, 2017). As entidades industriais que realizam a transformação digital dos seus processos, identificam ganhos significativos em áreas estratégicas, tornando-as modernas, simplificadas e com a tomada de decisão ágil (Tran-Dang & Kim, 2021).

A transformação digital nas empresas, segundo Muhuri et al. (2019a), foi fortalecida pela indústria 4.0 que permite a modernização das plantas fabris para a execução das atividades produtivas e monitoramento delas, porém o processo de aplicação de uma das vertentes de transformação digital se intensificou em função da pandemia da Covid-19 Dwivedi et al. (2020), que ocasionou uma mudança de cultura, alterando o modo de trabalhar e de gestão das empresas (Labucay, 2022).

O conceito da indústria 4.0 é fundamentada na utilização de ferramentas digitais para a execução de atividades, via sensores e outros mecanismos que possibilitam a coleta de dados durante o processo e a sua análise para a descoberta de conhecimento para o próprio meio, auxiliando no processo de tomada de decisão (Pejic-Bach et al., 2020), A indústria 4.0 é composta por pilares, sendo eles a integração de sistemas, internet industrial, segurança cibernética, armazenamento em nuvem, manufatura aditiva, realidade aumentada, ciência de dados, automação e simulação (Souza et al., 2021).

A tendência exponencial da evolução da tecnologia, fortalece a digitalização das cadeias produtivas, derivada da transformação digital agregada aos pilares da indústria 4.0. A digitalização acentua a importância das atividades, como na cadeia de suprimentos Tran-Dang & Kim (2021), que é um dos principais setores estratégicos da indústria.

A cadeia de suprimentos é responsável pelo conjunto de operações que compreendem desde o processamento do material na linha de produção até a entrega do produto ao cliente, impactando

por meio do seu desempenho o crescimento empresarial Zhang et al., (2021). Assim, pode-se garantir as operações de modo confiável, eficaz e flexível, como também o melhor gerenciamento da mão de obra, dos materiais e da logística, a fim de garantir a confiabilidade do cliente (Ng et al., 2022). Este setor passou a se desenvolver pelo incremento das peças personalizadas e decréscimo dos lotes padrão, resultando em uma diversidade dos suprimentos que devem ser disponibilizados para a produção.

Áreas como marketing, manutenção, logística e inventário são alicerces da cadeia de suprimentos e devem trabalhar sincronizadas com a demanda produtiva. A gestão de inventário tem papel essencial, pois está ligada diretamente com o custo de produção e a receita da empresa (Namir et al., 2021). Desta forma, deve-se guiar o processo para que se tenha a quantidade de materiais necessários no momento coerente, evitando-se o acúmulo de itens desnecessários em estoque e proporcionando a circulação dos suprimentos fundamentais na linha produtiva (Demeter et al., 2019).

A inadequada gestão de inventário faz com que o ciclo de suprimentos aconteça de forma mais lenta e com o custo mais elevado, pois ocasiona o acúmulo de materiais obsoletos, a carência de itens necessários para a evolução do processo e as compras de suprimentos realizadas em caráter de urgência (Guo et al., 2014).

A cadeia de suprimento é responsável pela gestão de matéria prima, componentes de execução do processo e inventário de suprimentos de manutenção para manter o processo na linha produtiva, com isso, cria-se uma conexão direta com o resultado, pois o processo produtivo não deve ser interrompido em função do mau funcionamento dos equipamentos (Agostinelli & Cumo, 2017).

A automatização da linha produtiva tem dado ênfase à área de manutenção, que assume a responsabilidade de manter a eficiência da cadeia produtiva, com vistas ao aumento da eficácia e da qualidade do processo de conserto dos equipamentos (Sajid et al., 2021). Os equipamentos, presentes na linha produtiva, são montados com componentes comerciais que se constituem em peças denominadas como suprimentos de manutenção. O conjunto dessas peças resulta no inventário de suprimentos de manutenção, que deve ser gerenciado e monitorado para proporcionar uma maior margem de produtividade, assim não interrompendo o processo produtivo.

O processo produtivo se programado e controlado, podem gerar dados que são armazenados em banco de dados, as informações gravadas podem ser analisados e viabilizar a extração do conhecimento (Benabdellah et al., 2019). Este processo se refere a ciência de dados que faz uso do

data mining e data analytics, empregando-se para isso o aprendizado estatístico ¹e o aprendizado de máquina² (Waller & Fawcett, 2013).

A Universidade de Harvard determina que a ciência de dados é a atividade voltada aos sistemas produtivos mais exigida no século 21, destacando a importância da área (Sajid et al., 2021). A ferramenta é utilizada por órgãos reguladores, como o Fórum Econômico mundial, que identifica o modo como as empresas operam e quais os possíveis impactos ambientais de suas ações no período de dez anos, (Qazi & Al-Mhdawi, 2023), desta forma, trabalha com pautas de confiabilidade de dados na tomada de decisão e qual o nível de exigência da sociedade na busca do conhecimento digital (Cin et al., 2022).

A aplicação das ferramentas e tarefas da ciência de dados no processo produtivo, permitem a empresa uma posição competitiva, pois altera a forma como a cadeia de suprimento realiza a tratativa das informações e realiza a tomada de decisão, já que a ciência de dados possibilita visualizar as despesas, se integrar sobre a sazonalidade das tendências e ajustar as demandas para o atendimento de tal fator (Hazen et al., 2014).

O campo de inovação e tecnologia no Brasil está influenciando o aperfeiçoamento de diversas áreas com a introdução de ferramentas e tecnologias digitais para efetivar a transformação digital nos parques fabris, como a indústria carbonífera do sul de Santa Catarina Soares (2022), responsável pela extração mineral utilizando o método de lavra, via subsolo. De acordo com a Federação das Associações Empresariais de Santa Catarina, este setor da economia gera 20 mil empregos diretos e indiretos, sendo responsável por 4% do valor adicionado³ das cidades ligadas à Associação dos Municípios da Região Carbonífera (SIESESC, 2022).

Dentre as áreas de uma indústria carbonífera, tem-se a unidade de manutenção que possui como foco principal a reposição de componentes mecânicos no processo produtivo, os quais ao serem disponibilizados devem estar funcionando e em conformidade com as normas de segurança. Na gestão de inventário de manutenção pode-se aplicar a ciência de dados para a análise dos dados existentes referente ao uso das matérias primas e a demanda produtiva real, a fim de se proporcionar a criação de um controle de suprimentos eficaz e o aprimoramento das tecnologias empregadas durante o processo.

¹ O aprendizado estatístico é definido como a área que extrai padrões de dados e proporciona o entendimento sobre variáveis encontradas no processo que represente o ambiente atual e futuro (P. H. T. de O. Sousa, 2016).

² O aprendizado de máquina é uma subárea da inteligência artificial, utilizada para desenvolver algoritmos de geração de conhecimento e tratar informações geradas com rapidez por meio de sistemas digitais (FaceLi et al., 2011).

³ Valor adicionado é toda a riqueza no âmbito nacional, gerada em um período determinado (Santos & Hashimoto, 2003).

Mediante o exposto, a presente pesquisa analisa como a ciência de dados pode alavancar a transformação digital na gestão de inventário de manutenção, no contexto de uma indústria carbonífera no sul de Santa Catarina, a fim de proporcionar uma tomada de decisão eficaz sobre pontos estratégicos da organização.

1.1 PROBLEMA

A área de manutenção das indústrias encontra-se em constante aprimoramento, pois está diretamente ligada a evolução das máquinas. A falha de um equipamento mecânico no processo produtivo, pode ocasionar a interrupção da produção, resultando em perda de capital à empresa (Ridzuan & Wan Zainon, 2021).

A manutenção tem relação com o faturamento, visto que mantém as atividades no ritmo proposto e enfrenta diversos desafios, como a sincronização da demanda produtiva com a disponibilidade dos equipamentos, por consequência a capacidade de produção de uma indústria está diretamente associada ao funcionamento adequado das suas máquinas.

A execução do processo de conserto dos equipamentos na linha de produção, demanda de suprimentos, que compõe o inventário de manutenção, respeitando variáveis de riscos que pode influenciar no resultado da empresa. O fluxo de informações sobre a necessidade e a utilização do componente são armazenados em sistemas gerenciadores de bancos de dados, que se tratados de forma adequada, transformam os dados em importante ativo para a empresa, Han & Trimi (2022), auxiliam na tomada de decisão dirigida a dados, no bom aproveitamento dos recursos e na vantagem competitiva.

O estoque de manutenção analisado, é baseado sobre o valor limite que não está relacionado com a demanda real, e é composto por todos os tipos de suprimentos de manutenção, pois é responsável pelo abastecimento de todas as áreas da empresa, porém o presente estudo está desenhado com foco em entender o comportamento das peças na gestão de estoque de manutenção de uma indústria carbonífera e promover uma gestão efetiva, minimizando impactos ao processo produtivo. Desta forma, busca-se entender como a ciência de dados pode ser utilizada para a transformação digital do setor de gestão de inventário em uma indústria carbonífera?

1.2 OBJETIVOS

1.2.1 Objetivo geral

Analisar as regras geradas do algoritmo Apriori, para apoio à tomada de decisão na gestão de inventário em uma indústria carbonífera.

1.2.2 Objetivos específicos

Os objetivos específicos desta pesquisa consistem em:

- a) Demonstrar por meio de uma análise bibliométrica a aplicação da ciência de dados para apoio à tomada de decisão na gestão de inventário;
- b) Identificar os métodos de ciência de dados a serem empregados para a descoberta de conhecimento na base de dados de inventário de manutenção;
- c) Aplicar algoritmos de aprendizado de máquina e métricas de qualidade em ciência de dados;
- d) Descrever o conhecimento descoberto nos dados, combinados com a análise de valor, para a tomada de decisão eficaz na gestão de inventários de manutenção.

1.3 JUSTIFICATIVA

A indústria, assim como a sociedade, está sendo impactada pela transformação digital, que está conectando os meios de comunicação e tornando possível a sincronicidade dos dados de modo eficaz, facilitando a tomada de decisão sobre diversos aspectos, sejam profissionais ou pessoais (Tran-Dang & Kim, 2021).

As fortes tendências mundiais, como o remodelamento dos setores industriais para conseguir acompanhar a globalização fortaleceram o maior contato entre pessoas em longas distâncias e a aceleração de procedimentos utilizados em empresas no dia a dia. A modernização dos parques fabris e diretrizes de trabalho, alavancam o sucesso da organização, pois permite a singularidade sobre procedimentos anteriormente complexos de fazer e uma melhor relação entre funcionário, cliente e parceiros (Forum, 2022b).

A transformação digital também foi intensificada pela pandemia da Covid-19, pois os ambientes de trabalho, estudos ou lazer tiveram que se adaptar as lacunas construídas pela impossibilidade da circulação das pessoas nestes espaços (Dwivedi et al., 2020). A aceleração da

digitalização de atividades, permitiu o fortalecimento de diferentes sistemas de informações para a tomada de decisão nos diferentes meios, como por exemplo, os econômicos, sociais e ecológicos.

O aumento exponencial do uso de instrumentos geradores de informações em diferentes perspectivas da vida urbana e social, está desenvolvendo um novo conceito denominado revolução de dados. Este conceito é considerado um gerador de possibilidades já que permite a análise das informações, porém para a extração do conhecimento predominante nas bases de dados requer a capacitação em ciência de dados, que sincroniza estatística, ciência da computação e necessita conhecimento técnico sobre a área que está sendo analisada (Initiative, 2022).

A tecnologia aplicada nos parques fabris de países desenvolvidos estimula a necessidade da atualização da produção nacional e para incentivar o fortalecimento de ferramentas tecnológicas na extração mineral o governo federal brasileiro assinou o decreto N° 10.438 em abril de 2002, que sofreu alterações em Janeiro de 2022 N° 14.299⁴ referente ao “Programa de Transição Energética Justa”, construído sobre uma série de diretrizes, subsidiadas pelo governo, no intuito de reduzir impactos ambientais e aumentar a eficiência da produção, utilizando o conceito de produção mais limpa (Tomaz, 2022).

O processo de extração do carvão mineral é realizado por máquinas que devem estar em funcionamento desde a mineração, beneficiamento e expedição do produto. Os equipamentos em atuação na empresa, são responsabilidades do setor de manutenção, estruturada sobre o conceito de produção mais limpa, que busca constantemente estratégias econômicas, sustentáveis e tecnológicas para garantir o melhor resultado (Cosmo et al., 2020).

O setor carbonífero é afetado diretamente pela gestão incoerente de inventários de manutenção, resultando em um maior *lead time*⁵ no processo de reparo dos equipamentos, conseqüentemente em paradas de produção, e na deflagração da manutenção, por demonstrar inércia no momento de resposta (George et al., 2022). O alto *lead time* na resolução de tarefas de manutenção prejudica também outros setores da economia, como por exemplo, no campo da aviação, onde o período de parada do equipamento é contabilizado e deve ser realizadas as manutenções preventivas, preditivas e corretivas em um curto espaço de tempo, para então disponibilizar o equipamento para a operação com total segurança (Stadnicka & Ratnayake, 2017).

A falta de gerenciamento de suprimentos prejudica as atividades em suas mais diversas áreas e a presente pesquisa é elaborada no setor carbonífero, porém a forma como é modelada proporciona

⁴ [L14299 \(planalto.gov.br\)](https://planalto.gov.br)

⁵ Lead time é definido sobre o intervalo decorrido entre a solicitação de um produto realizado pelo cliente até a entrega do produto ao mesmo (Abreu & Gomes, 2013).

a aplicação da ferramenta de ciência de dados em problemas de gestão de inventário de setores hospitalares, farmacêuticas, industriais, hoteleiras, entre outras.

A utilização de dados gerados nas atividades de gestão de suprimentos, geralmente são desconsideradas pela idealização de que os dados são inconsistentes e não estão ligados entre si (C. K. M. Lee & Palaniappan, 2014), logo no decorrer da pesquisa, está demonstrado a estrutura ideal para o tratamento de dados e como cada modelo pode ser desenvolvido para um melhor resultado.

No que se refere ao processo eficiente de conserto dos componentes, este é construído sobre a disponibilidade de peças de manutenção para a troca em um cenário de falha. A gestão de suprimentos em inventário de manutenção garante a execução eficaz da atividade, pois disponibiliza os itens necessários no momento de execução da manutenção. Na gestão de inventário, pode ser empregadas ferramentas de apoio que permite a visualização das futuras demandas e o armazenamento da quantidade correta dos materiais (López-Soto et al., 2017). A programação incoerente da disponibilidade de suprimentos no inventário, pode ocasionar para a empresa perdas de produção, de recursos e de confiabilidade do cliente, influenciando diretamente na receita da empresa (Sangkhasuk et al., 2018).

No ramo empresarial a melhoria contínua, inovação de serviços e produtos representa a gestão do conhecimento, pois utiliza do conhecimento explícito e tácito para resolução de problemas diários e permite o ganho de resultados sobre as diferentes perspectivas de pensar (Gaviria-Marin et al., 2019).

A presente pesquisa, demonstra aderência aos Objetivos de Desenvolvimento Sustentável (ODS) da Organização das Nações Unidas, no que se refere a Indústria, Inovação e Infraestrutura (ODS 9), pois fomenta a infraestrutura resiliente, promove a industrialização inclusiva e sustentável, e fomenta a inovação dentro da indústria, permitindo a modernização de parques fabris, com acesso à tecnologia até 2030; e ao Consumo e Produção Responsáveis (ODS 12), pois promove a importância da aplicação de recursos para diminuir os impactos da exploração de energias não renováveis e o impacto ambiental (Forum, 2022a).

1.3.1 Inserção social

A inserção social está ligada diretamente aos impactos sociais, educacionais, tecnológicos, econômicos, sanitários, culturais, artístico, profissional, legal e ambiental que a pesquisa pode atingir com o tema em pauta, desta forma a pesquisa deve estar dentro dos parâmetros de inserção social para ser efetiva e construir pontes entre o social e o acadêmico para trabalhar as possibilidades de aperfeiçoamento de ambos (Ferraço & de Farias, 2021).

O impacto social gerado pela presente pesquisa possui correlação com a recolocação estratégica das empresas da região uma vez que fomenta a atualização das plantas fabris, principalmente no ramo carbonífero, responsável pela movimentação de parte significativa da economia regional. Assim, se entrelaça com os impactos educacionais, fomentando o preparo do profissional para o manuseio das ferramentas e novos desafios, logo a quadrupla hélice da inovação deve acontecer de forma síncrona entre as quatro áreas que a compõe, política, social, acadêmica e empresarial, proporcionando a atualização do parque fabril.

A transformação digital é uma forma de inserção social, pois se trata de uma mudança significativa na cultura organizacional. A população atual tem sido impactada pela tecnologia, buscando se adaptar, no entanto, a não acessibilidade a ferramentas tecnológicas é considerada uma dificuldade profissional e social. Dessa forma, o Estado deve assegurar treinamento e integração a todos os integrantes da sociedade a fim de proporcionar a devida inclusão digital (Ó. Sousa & Rodrigues, 2022).

A ciência da inovação tem permitido a inserção e integração de todos os públicos na sociedade em um cenário de rápidas mudanças, financiadas pela evolução tecnológica, inteligência artificial e realidade virtual presentes na área profissional, permitindo o gerenciamento de atividades e a participação dos mais diversos grupos que integram o meio social (Institution, 2022).

O impacto legal é fundamentado sobre o fortalecimento de leis implantadas e trabalhadas pelo próprio governo federal Nº 14.299, que obrigam as carboníferas a diminuir os recursos ambientais e valorizar os sociais e econômicos na extração e beneficiamento do carvão mineral, utilizando recursos atuais.

1.4 CARACTERIZAÇÃO INTERDISCIPLINAR E ADERÊNCIA AO PROGRAMA

A interdisciplinaridade é um paradigma metodológico que visa o encontro de diversas áreas do conhecimento, em uma perspectiva colaborativa em busca da resolução de problemas compartilhando informações, resultados e metodologias Carneiro (1994); Pátaro & Bovo, (2017).

O Programa de Pós-graduação em Sistemas Produtivos (PPGSP) é da área interdisciplinar, e a presente pesquisa é aderente a linha de Gestão e Conhecimento em Sistemas Produtivos. A referida linha de pesquisa objetiva, segundo (PPGSP, 2021) “investigar os sistemas produtivos por meio de conhecimentos científicos e/ou tecnológicos em distintos contextos. Propõe atuar de forma interdisciplinar as temáticas: estratégia e aprendizagem organizacional, gestão do conhecimento, inovação e sustentabilidade; inteligência empresarial; gestão de processos e produção.”. Na figura 1 tem-se a representação da interdisciplinaridade envolvida na linha de pesquisa de Gestão e Conhecimento em Sistemas Produtivos, a que a presente pesquisa se relaciona.

Figura 1

Representação do conceito interdisciplinar do PPGSP.



Nota. Autores (2022).

A presente pesquisa relaciona-se com a área estratégica, pois está associada diretamente a possibilidade de entender e analisar as demandas e preparar os possíveis cenários para os futuros pedidos, como em uma cadeia de suprimentos, que segundo (Tsou, 2013) pode se utilizar de ferramentas estatísticas para compreender o fluxo de determinada mercadoria em um período e consegue elaborar sua programação.

O aprendizado organizacional passou por diferentes perspectivas desde a revolução industrial, porém a partir dos anos 2000, as empresas têm se remodelado em um curto espaço de tempo, para proporcionar uma entrega ágil, com qualidade excelente, além de sincronizar os serviços com os ambientes virtuais para poder manter a organização em um patamar estratégico (Schwandt & Marquardt, 1999), trabalhando diretamente com o campo de inovação, para buscar as atualizações no mercado.

A capacidade de realizar o gerenciamento de processos e identificar o conhecimento presente nas operações, diferencia a predominância da empresa no mercado estratégico organizacional e fortalece as tratativas competitivas, logo em um campo empresarial é comum encontrar áreas correlacionadas, como a gestão do conhecimento e gestão de processos e produção para alavancar resultados (Strauhs et al., 2012).

A introdução da tecnologia no ambiente organizacional é uma característica interdisciplinar, por conta do controle e suporte de atividades, apresentadas e monitoradas em uma plataforma, concentrando o conhecimento gerado por diversos âmbitos e produzindo novas formas para o seu entendimento e compartilhamento (Philippi. & Fernandes, 2021).

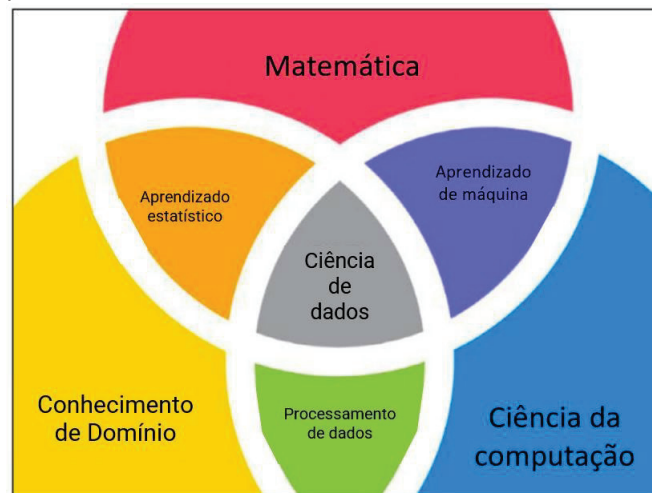
A quarta revolução industrial é resultante da inclusão da transformação digital na indústria 4.0, fortalecendo tecnologias inovadoras em diferentes áreas econômicas organizacionais (Diogo et al., 2019). A ciência de dados é considerada uma tecnologia inovadora, pois realiza a gestão do

conhecimento em conjunto com a tecnologia da informação, utilizando-se de algoritmos de inteligência artificial e de estatística, para auxiliar na tomada de decisão (Bellomarini et al., 2022).

A ciência de dados identifica as relações e padrões existentes em uma base de dados, trabalhando com ferramentas matemáticas, estatísticas, da ciência da computação e da ciência comportamental, reunindo-as de forma interdisciplinar, contextualizando o conhecimento extraído de todos os pontos que a compõe e construindo resultados sobre diversas perspectivas (Hazen et al., 2014).

Figura 2.

Figura explicativa sobre os campos que compõe a ciência de dados e como o campo trabalho no meio interdisciplinar para o processamento do conhecimento.



Nota. Figura adaptada de 'Data science applications for predictive maintenance and materials science in context to Industry 4.0' de (Sajid et al., 2021).

As fronteiras estabelecidas sobre as disciplinas são ultrapassadas quando tratadas de forma interdisciplinar, o conceito constrói pontes de ligação entre os diferentes pontos que a compõe (Tsui et al., 2019), sejam eles, matemáticos, estatísticos e tecnológicos para a dissipação do conhecimento (L. Huang et al., 2018).

O profissional da ciência de dados é treinado para pensar de forma interdisciplinar, para proporcionar o melhor tratamento de dados, realizando uma análise sobre quais ferramentas e frameworks devem ser utilizados para a extração dos dados, e em quais dados estão as informações geradoras de conhecimento (L. Huang et al., 2018).

A ciência de dados trabalha com a agregação de diversas áreas para permitir a tomada de decisão, sobre a informação presente nas bases de dados (Boehmke et al., 2020), porém a efetivação da ferramenta só é possível com a definição do tamanho da população e o esclarecimento do problema em questão, assim inicia-se o processo de avaliação de ferramentas apropriadas para a execução das ações, obedecendo os critérios de sequência lógica (Maass & Storey, 2021).

A efetivação da ciência de dados é realizada sobre a extração de informações coerentes de banco de dados para a tomada de decisão, dentro das organizações vem sendo aplicada frequentemente na cadeia de suprimentos, como também na gestão de inventário, desde modo, realiza o entendimento de demandas e o posicionamento estratégico da empresa, que podem ser construídas sobre diferentes ferramentas de aplicação como a data mining e a big data, que são aplicadas juntas para a descoberta de conhecimento (L. Huang et al., 2018).

A ciência de dados auxilia no entendimento do conhecimento por meio de métodos de extração de informação, sendo elas estatísticas e algoritmos de inteligência artificial, que trabalham com a modelagem das informações para um melhor resultado, e gestão de inventário de forma efetiva, fortalecendo o conceito interdisciplinar na ferramenta desenvolvida (Ijadi et al., 2018).

A gestão de inventário é construída sobre a coordenação eficiente dos suprimentos dentro da empresa, correlacionada com o posicionamento estratégico. O profissional deve ter uma visão ampla sobre as áreas que integram o inventário, como a logística, custo e fornecedores, para fortalecendo o conceito interdisciplinar na área estudada (Govindasamy & Antonidoss, 2021).

A dinâmica da gestão de inventário é definida por Huang (2021), como multidisciplinar, pois permite a utilização de ferramentas de gerenciamento, independente do ramo. O conhecimento e o domínio devem ser trabalhados juntos, e articulados de maneira interdisciplinar, onde é utilizado a ciência de dados em metodologias qualitativas e quantitativas nos âmbitos de trabalho (Waller & Fawcett, 2013).

Após a análise das áreas estudadas pelos demais autores que estão correlacionados diretamente com o tema, podem-se afirmar a dissertação possui aderência no programa interdisciplinar, por tratar de diversas áreas na construção de uma solução sobre diversas perspectivas do conhecimento.

O objetivo geral da presente pesquisa é elaborar métodos que integram a ciência de dados no processo de tomada de decisão, na gestão de inventário no setor de manutenção carbonífero. Visualiza-se então, a possível atuação de diferentes ferramentas para a extração do conhecimento como a utilização do data mining para a extração de dados e a utilização de ferramentas estatísticas para a demonstração dos resultados extraídos.

1.5 ESTRUTURA GERAL DO DOCUMENTO

A estrutura da presente pesquisa foi desenvolvida para proporcionar o entendimento e a fundamentação do tema estudado, além de encontrar a solução para o problema apresentado, constituindo-se por metodologias e métodos que se complementam para mensurar os dados sob e

proporcionando um entendimento interdisciplinar da aplicação da ciência de dados na gestão de inventário.

Na primeira etapa, foi elaborado o estudo bibliométrico, a fim de entender o comportamento dos estudos que envolvem os temas da presente pesquisa atualmente, seguindo-se pela estruturação do panorama teórico, onde estão definidos os principais assuntos abordados ao longo do trabalho. No decorrer do texto foram extraídos via critérios estabelecidos, seis trabalhos e sobre eles, foi redigido o estado da arte.

A etapa de metodologia, segunda da pesquisa, é estruturada com métodos que permitem a execução dos objetivos propostos, a etapa está descrita de forma clara, apresentando os processos e ferramentas utilizadas para o tratamento dos dados, dentre suas principais etapas, estão a descrição da população, o pré-processamento, a mineração de dados e os pós-processamento dos dados.

Durante a seção subsequente, de resultados, estão demonstradas as relações extraídas da mineração de dados e correlacionadas com a análise de valor, conforme alinhado nas linhas de pesquisa do presente mestrado, e sobre a visão de interdisciplinaridade. Nessa seção está alinhado o atingimento do objetivo geral, disposto na presente pesquisa.

Por fim, as conclusões estão descritas, abordando o impacto da pesquisa no meio acadêmico e empresarial. Na presente etapa também se apresentam as futuras oportunidades de estudo no tema abordado.

2 PANORAMA TEÓRICO

O panorama teórico está estruturado em um artigo bibliométrico e na revisão descritiva das quatro principais áreas envolvidas na pesquisa e como elas se relacionam, as quais são gestão de inventário, transformação digital, ciência de dados e mineração de dados.

2.1 REVISÃO BIBLIOGRÁFICA

2.1.1 Uma análise bibliométrica de decisão orientada por mineração de dados na gestão de inventário

A inclusão da tecnologia em empresas de manufatura, é efetiva junto aos meios estratégicos presente diariamente na esfera social e industrial, como os aplicativos de gerenciamento de tarefas, calendários online, redes sociais e facilitadores de atividades, que estão inseridos a estruturas organizacionais.

As empresas de manufatura geram valor a partir de processos físicos e químicos, para a produção de produtos, tendo-se as etapas arquitetadas sobre diversos setores (Groover, 2017). A área responsável pelo fluxo do produto na empresa, desde a entrada da matéria prima e entrega ao cliente do material acabado, é chamada de cadeia de suprimento, e integra setores como compras, produção, logística, inventário e armazenamento (Queiroz et al., 2019).

A cadeia de suprimento pode ser definida como um direcionador dos produtos e serviços não acabados na empresa (Hussain et al., 2021). É responsável pela transformação do material na linha de produção, que resulta na entrega do produto final ao cliente (Zhang et al., 2021).

As empresas manufatureiras se mostram eficientes em capturar, organizar e armazenar dados, obtidos de suas operações rotineiras, porém, ainda pode-se melhorar à gestão de dados e transformá-los em ganhos financeiros. Desta forma, a introdução da mineração de dados na estratégia das empresas é crucial, pois ela é conceituada como uma ferramenta de descoberta de conhecimento, em princípios de análise de dados exploratórios e de modelagem de informação (Côrtes et al., 2002).

O posicionamento das empresas manufatureiras no mercado é resultado de uma tomada de decisão eficaz, que está sofrendo interferência com a transformação digital (Rogers & Serra, 2017). A digitalização de atividades empresariais, resulta na criação de banco de dados, que tratados com métodos de análise eficientes, possibilita dissipar conhecimento estratégico. A utilização de dados no gerenciamento das empresas, surgiu no final dos anos noventa, mas só foi definido como viável em 2001 e desde então vem se aperfeiçoando exponencialmente (Mariani & Fosso Wamba, 2020). Os resultados deste estudo bibliométrico estão demonstrados por meio do mapeamento e da análise das publicações correlacionadas ao assunto, em diferentes bases de dados científicas.

A possibilidade de analisar temas em períodos restritos, definidos pelo viés de pesquisa que permitem interpretar os acontecimentos que influenciaram volumes de publicações, é um dos entendimentos que a bibliometria proporciona e o direcionamento para futuras pesquisas (Muhuri et al., 2019b).

O procedimento bibliométrico, descrito pela Lei de Bradford, Lokta e Zipf, é um método que busca trabalhos relacionados a um tema utilizando uma série de critérios e resulta no volume final. A este volume final são aplicadas a Lei de Bradford, método que permite julgar a produtividade das revistas em que os trabalhos foram encontrados, a Lei de Lokta que avalia o rendimento dos autores e a Lei de Zipf que a frequência das palavras definidas como palavras-chaves (Voese & Mello, 2013).

A elaboração da bibliometria foi na esfera descritiva e exploratória sobre os estudos encontrados nas bases de dados IEEE Xplore, Science Direct, Web of Science e Scopus. Foram considerados para a bibliometria, somente artigos encontrados pelo conjunto de palavras-chaves *"Inventory management"* AND *"Data mining"*, resultando 2.191 artigos. O mapeamento de dados ocorreu em agosto de 2022.

Para o alcance de artigos que estão diretamente correlacionados com o tema da pesquisa, foi efetuada a busca nas quatro bases de dados selecionados, subdivida em quatro etapas. A primeira, foi realizada para encontrar as palavras chaves *"inventory management"* AND *"data mining"* em qualquer parte dos textos, seguida pelo primeiro critério da pesquisa, que considera apenas estudos publicados no período de 2012 a 2022. Para filtrar os artigos que abordavam o tema proposto, foi realizada uma pesquisa avançada para encontrar periódicos que contém *"inventory management"* AND *"data mining"* no título, resumo e palavras-chaves, formalizando o segundo critério de pesquisa estabelecido, obtendo-se 127 artigos.

O tratamento de dados na gestão de estoque é utilizado em algumas situações, como para melhorar o layout do setor e mapear a rota de transporte de suprimentos, pois os dados permitem visualizar a frequência da utilização do suprimento e podem sugerir o local correto de armazenamento. Porém, esse não é o foco principal do estudo, desenvolvido para compreender o comportamento do campo de pesquisa que integra a mineração de dados na tomada de decisão da gestão de inventário. Para trabalhar somente com publicações relacionadas com a pesquisa, foi estabelecido o critério três, que pede para excluir artigos que contém *"logistic"* em seus resumos, somando 120 artigos.

Tabela 1.

Apresentação do número de artigos encontrados nas bases de dados, sobre diferentes critérios

Base de dados	Web of science	Scopus	Science Direct	IEEE Xplore
Artigos encontrados	70	1.341	699	81
1º Critérios = Artigos publicados entre 2012 e 2022	50	1.108	502	46
2º Critério = Contém "Inventory management" and "data mining" no título, resumo e palavras chaves.	29	80	7	11
3º Critério = Não contém "Logistic" no resumo	29	73	7	11

Nota. Autores (2022).

Conforme apresentado na tabela 1, o estudo é composto por 120 artigos, sendo 29 provenientes da Web of Science, 73 da Scopus, 7 da Science Direct e 11 da IEEE Xplore. Durante o processo de filtros e estabelecimento de critérios, foram desconsiderados 2.071 estudos.

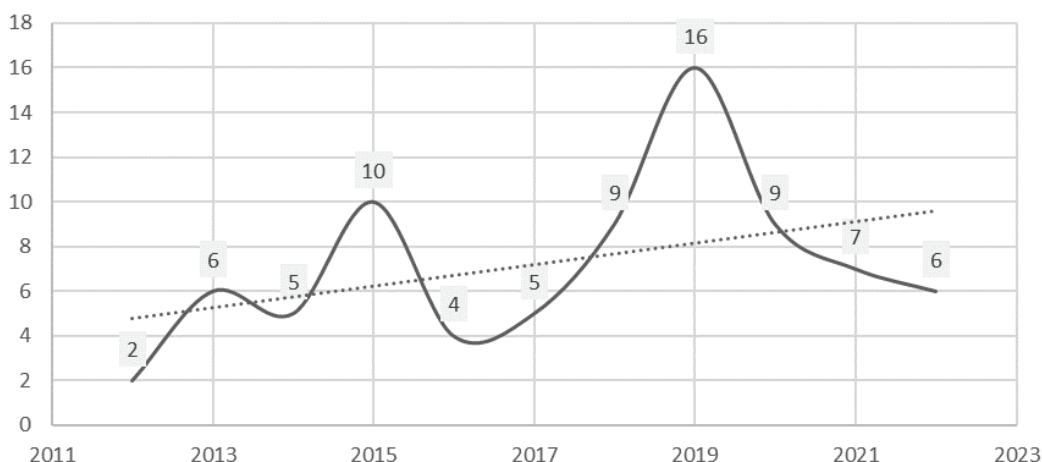
Foram identificados 2.191 artigos relacionados ao tema. O mapeamento desses artigos foi elaborado em uma planilha no Excel, na qual foram subdivididos por: título, autor, instituição de afiliação do autor, país da instituição de afiliação do autor, coautor, instituição de afiliação do coautor, país da instituição de afiliação do coautor, revista onde o artigo foi publicado, ano de publicação, fator de impacto da revista, Qualis CAPES do periódico, palavras-chaves utilizadas no estudo.

O processo de qualificação de um periódico é via o fator de impacto, que é mensurado considerando o número de citações da revista, o índice é classificado bom quando seus valores são altos, resultado de um elevado número de interações com o periódico (Pinto & Andrade, 1999), esse mecanismo de avaliação é utilizado mundialmente, já o Qualis CAPES é uma ferramenta estritamente brasileira, desenvolvida para classificar as revistas entre critérios de A1, mais elevado; A2; A3; A4; B1; B2; B3; B4; C e entre as áreas de pesquisa.

Durante o processo de mapeamento de artigos no Excel foram observados 40 estudos duplicados, para evitar o processamento de informações incoerentes durante o levantamento de dados, os estudos duplicados foram removidos, resultando em 80 pesquisas.

A interpretação dos parâmetros extraídos dos estudos encontrados, se desvincula na distribuição de publicação ao longo dos anos, quais foram as instituições que mais atuaram nas publicações, a influência dos periódicos e o entendimento do comportamento das publicações em diferentes regiões do mundo.

Figura 3.
Publicações por períodos



Nota. Autores (2022).

A visualização de publicações por meio de uma linha do tempo permite o questionamento sobre os acontecimentos no decorrer do período, apresentando o comportamento das publicações de 2012 a 2022 (figura 3). O Cume é alcançado em 2019, contabilizando 16 artigos publicados.

O ano de 2019 foi marcado com o surgimento da tecnologia 5G, que possibilita a navegação e o acesso a informações com mais eficácia, possibilitando um melhor desempenho de software de tratamento de dados dentro das organizações. O surgimento da internet 5G resulta em um grande impacto na cadeia de suprimento, pois permite uma maior agilidade no processamento de dados entre os setor (Hassija et al., 2021).

Tabela 2.
Tipo de publicações dos artigos .

	Conferência	Revista	Livro
Porcentagem de trabalhos	43.75%	50%	6.25%

Nota. Autores (2022).

Os artigos encontrados nas bases de dados foram publicados em um dos três tipos de publicações, apresentados na tabela 2, 50% dos estudos são derivados de conferências contabilizando 35 publicações, seguido por 43,75% relacionados a revistas, o que representa 40 trabalhos e 6,25% que são 5 livros.

Tabela 3.

Revistas nas quais os estudos foram publicados.

REVISTA	Nº DE PUBLICAÇÃO	FATOR DE IMPACTO	QUALIS CAPES
Computers & operations research	2	5.159	A2
Applied intelligence	1	5.019	B1
Applied mathematical modelling	1	5.336	A2
Computer Science and Information Systems	1	1.167	B2
Decision Support Systems	1	6.969	
Electronic Library	1	1.453	
European Journal of Industrial Engineering	1	1.783	
European Journal of Operational Research	1	6.390	A1
Expert Systems With Applications	1	8.665	
IEEE Journal of Biomedical and Health Informatics	1	5.223	
International Journal of Advanced Manufacturing Technology	1	3.226	B1
International Journal of Business Analytics	1	0.513	
International Journal of Data Analysis Techniques and Strategies	1	0.500	
International Journal of Decision Support System Technology	1	0.139	
International Journal of Engineering Business Management	1	2.243	
International Journal of Pharmacy and Technology	1	0.102	
International Journal of Production Management and Engineering	1		B2
International Journal of Production Research	1	9.018	B1
International Journal of Service Science, Management, Engineering, and Technology	1	3.420	
International Review of Retail Distribution and Consumer Research	1	2.760	
Journal of Advanced Research in Dynamical and Control Systems	1	0.308	
Journal of Information & Optimization Sciences	1		
Journal of Intelligent Information Systems	1	2.940	
Journal of Modelling in Management	1	2.770	B2
Knowledge and Information Systems	1	2.531	A2
Lecture Notes in Business Information Processing	1	0.899	B3

Lecture Notes in Computer Science	1	1.363	B1
Lecture Notes in Mechanical Engineering	1	0.554	
Management and Production Engineering Review	1	1.452	
Management Science Letters	1	2.064	
Recent Advances in Computer Science and Communications	1	1.060	
Remote Sensing	1	5.333	A1
Scientific Journal of Logistics	1		
Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University	1	0.870	
Slas Technology	1	2.813	
Smart Materials and Intelligent Systems	1	2.970	
Visualization in Engineering	1	5.000	
Wireless Communications & Mobile Computing	1	2.336	

Nota. Autores (2022).

A tabela 3 apresenta os trabalhos publicados em periódicos e demonstra informações como os números de citações e o fator de impacto, um índice criado que determina a importância da revista, segundo Portugal et al. (2011) no território brasileiro apresenta-se o Qualis CAPES que classifica como A1 os estudos com maior relevância; A2; A3; A4; B1; B2; B3; B4; C.

Tabela 4.

Estudos que demonstram o maior número de citações

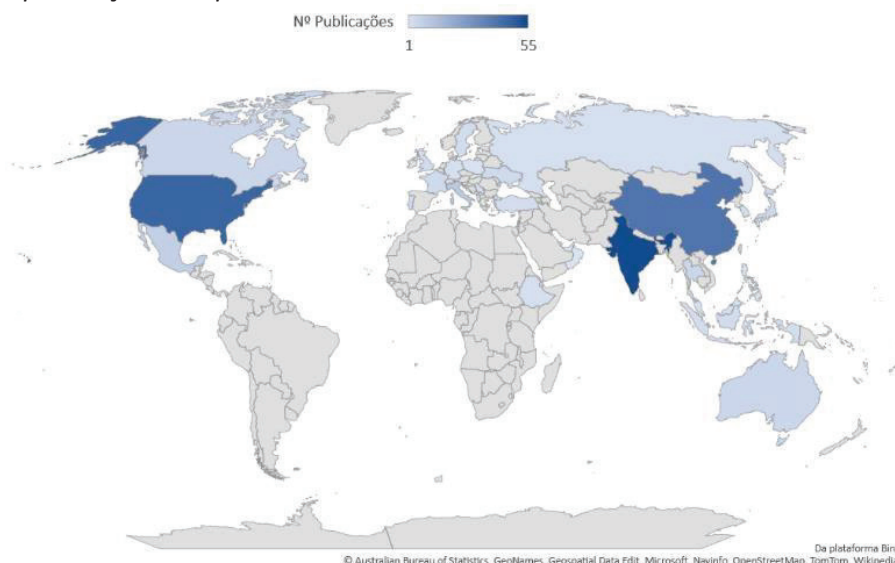
Nº	Base de dados	Título	Referência	Nº Citações	Qualis Capes
67	Scopus	Detection, classification, and mapping of U.S. traffic signs using google street view images for roadway inventory management	(Balali et al., 2015)	42	
68	Scopus	IoT-based asset management system for healthcare-related industries	(C. K. M. Lee et al., 2015)	35	
79	Scopus	Improving inventory performance with clustering based demand forecasts	(Bala, 2012)	30	A3
23	Web of Science / Scopus / Science Direct	Predicting solutions of large-scale optimization problems via machine learning: A case study in blood supply chain management	(Abbasi et al., 2020)	27	A1
26	Web of Science / Scopus	Forecasting flow time in semiconductor manufacturing using knowledge discovery in databases	(Tirkel, 2013)	24	A1
21	Web of Science / Scopus / Science Direct	On the strategy of supply chain collaboration based on dynamic inventory target level management: A theory of constraint perspective	(Tsou, 2013)	22	A1
54	Scopus / Science Direct	Coping with demand volatility in retail pharmacies with the aid of big data exploration	(Papanagnou & Matthews-Amune, 2018)	21	A1
70	Scopus / Science Direct	Estimating available supermarket commodities for food bank collection in the absence of information	(Brock & Davis, 2015)	21	A1
3	Web of Science / Scopus	Collaborative Predictive Business Intelligence Model for Spare Parts Inventory Replenishment	(Stefanovic, 2015)	17	B2
25	Web of Science / Scopus	A hybrid framework for mining high-utility itemsets in a sparse transaction database	(Dawar et al., 2017)	16	A3
52	Scopus	Lean principles for organizing items in an automated storage and retrieval system: An association rule mining – Based approach	(Bevilacqua et al., 2019)	14	

Autor: Autores (2022)

A importância das publicações na área estudada, tabela 4, foi mensurada pela quantidade de vezes que cada artigo foi citado em outro trabalho, demonstrando a importância com relação a área de pesquisa. Na avaliação dos artigos foi adicionado o Qualis CAPES na tabela, evidenciando a qualidade do estudo.

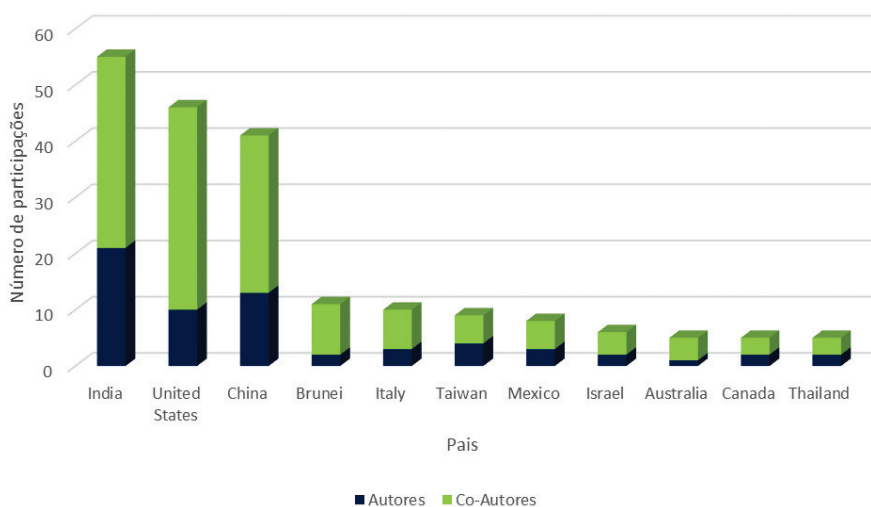
Realizar a leitura sobre a perspectiva de quais estão sendo as regiões ao redor do mundo com maior embasamento sobre o tema, fortalece a avaliação sobre a metodologia bibliométrica, onde se torna mais fácil o entendimento de parâmetros e a influência externa sobre a área de pesquisa.

Figura 4.
Concentração de publicações nos países.



Nota. Autores (2022)

Figura 5.
País de origem das publicações dos autores e coautores.



Nota. Autores (2022).

Nas figuras 4 e 5, é possível visualizar a localização das instituições de afiliação dos autores e coautores, na figura 4, permite o entendimento por quantidade total de publicações derivadas dos países, seja esse dado um autor ou coautor. Na figura 5, tem-se representada 80% das publicações, entre autores e coautores que demonstram a origem das instituições.

A Índia tem o maior número de publicações, somando 21 derivadas de autores indianos e 34 de coautores. Os Estados Unidos, possuem o maior volume de coautores somando 36, tendo-se 10 publicações de autores americanos, o que totaliza 46 trabalhos dos identificados no período e considerados no presente estudo.

A frequência da presença de uma instituição na área de pesquisa, representa o engajamento dela com o tema, a tabela 5 apresenta as instituições que tiveram mais de três publicações, considerando-se os autores e co-autores, e o país onde estão localizadas.

Tabela 5.
Instituições presentes nas publicações.

INSTITUIÇÕES	PAIS	ARTIGOS
Florida international university	ESTADOS UNIDOS	15
Universiti teknologi	BRUNEI	11
Xiamen university	MALÁSIA	9
Shanghai jiaotong university	CHINA	6
Marche polytechnic university	ITALIA	6
A p shah institute of technology	ÍNDIA	5
Renmin university of china	CHINA	5
Erode arts and science college	ÍNDIA	4
Towson university	ESTADOS UNIDOS	4
Wayne state university	ESTADOS UNIDOS	4
Amity school of science engineering and technology	ÍNDIA	3
Indraprastha inst informat technol	ÍNDIA	3
Manipal academy of higher education	ÍNDIA	3
Phranakhon rajabhat university	TAILANDIA	3
Sungkyunkwan university	COREIA DO SUL	3
Telkom university	INDONÉSIA	3
The hong kong polytechnic university	HONG KONG	3
Universidade do minho	PORTUGAL	3
University of skövde	SUÉCIA	3
Vellore institute of technology	ÍNDIA	3

Nota. Autores (2022).

Tabela 6.

Frequência das palavras chaves

PALAVRAS-CHAVE	FREQUÊNCIA	PERCENTUAL
Data Mining	36	10,29
Inventory Management	25	7,14
Association Rule	7	2,00
Forecasting	5	1,43
ABC Classification	4	1,14
Internet of Things	4	1,14
Cross-selling	3	0,86
FP-Growth	3	0,86
Healthcare	3	0,86
Industry 4.0	3	0,86
Knowledge Discovery	3	0,86
Machine Learning	3	0,86
Management	3	0,86
Medical Stores	3	0,86

Autor: Autores (2022).

As palavras-chave são ferramentas de endereçamento do artigo, devem ser escolhidas de acordo com o tema dissertado. Na tabela 6 estão demonstradas 30% das palavras-chaves mais presentes nos estudos, resultado da soma de 105 palavras encontradas. *Data mining*, representa 10,29% das palavras e estava presente em 36 artigos. Para mensuração, foram contabilizados nomes compostos que integravam as palavras chaves, como “*inventory management*” e “*Supply chain inventory management*”.

A similaridade entre o tema proposto na presente pesquisa e os estudos encontrados pelos critérios estabelecidos, foi demonstrada via mensuração das palavras-chaves utilizadas nas publicações, que apresentaram termos que coadunam com a pesquisa.

A atualidade do tema é destacada novamente na figura 4 que apresenta os países das instituições que originaram os artigos, trazendo a Índia em primeiro lugar no quesito autores e coautores, somando 55 publicações, e quando mensurado somente por autores contabilizou, 21 publicações.

O presente estudo conclui que a mineração de dados pode ser relacionada com a gestão de inventário em diferentes perspectivas. Os trabalhos encontrados foram desenvolvidos com o intuito de permitir a digitalização de atividades e o levantamento de dados para a execução de tarefas rotineiras do setor, além de guiar a organização para a tomada de decisão orientada a dados na gestão de inventário.

2.2 GESTÃO DE INVENTÁRIO

As empresas e organizações dependem de outros estabelecimentos como clientes e fornecedores para garantir a continuidade de suas operações. Essa interdependência é essencial, assim como a rede de suprimentos, que estrutura a empresa por meio da cadeia de suprimentos (Slack et al., 2013).

A cadeia de suprimentos de uma organização é arquitetada sobre o fluxo de materiais, de processos e de produtos, que circulam e interagem entre si, tendo como resultado o produto (Stefanovic, 2015).

O objetivo da cadeia de suprimentos é permitir, com o melhor resultado, a produção e a conclusão do produto no processo (Fahimnia et al., 2019). O setor é composto por várias áreas que se integram, como por exemplo, o gerenciamento de inventário de matéria-prima ou de produto acabado (Omar et al., 2020). O setor que mais influência na cadeia de suprimentos é o inventário, pois ele permite o fortalecimento da relação entre a empresa e o cliente, porém quando não está sincronizado com a demanda, resulta na perda de tempo produtivo, de recursos e no aumento dos custos produtivos (Corrêa, 2019).

O posicionamento da empresa no campo estratégico, está ligado diretamente com o inventário, pois os recursos gerenciados pelo setor geram ou têm a capacidade de gerar riqueza a organização, de maneira econômica (Alt & Martins, 2009). Os suprimentos necessários para a produção, que deve estar vinculada diretamente com a quantidade e os produtos em estoque, assim permitem a circulação dos produtos, inibindo a existência de materiais ociosos e obsoletos (López-Soto et al., 2017).

Em uma organização existem dois tipos distintos de gerenciamento de estoque, o inventário de suprimentos necessário para a fabricação do produto, que resulta no produto e o estoque de produto final acabado, que fica armazenado aguardando a solicitação do cliente (Bandaru et al., 2015).

O estoque de produto acabado é organizado para disponibilização imediata e fortalece a relação entre cliente e empresa, porém atinge diretamente o custo de produção, influenciado pela alta demanda de materiais em estoque, somando um maior custo de inventário de produto acabado.

A complexibilidade do gerenciamento de suprimentos é definida em garantir ao cliente o produto com um custo menor para a organização (Guo et al., 2014). O gerenciamento de estoque eficaz, acontece quando o fabricante consegue visualizar e entender as demandas futuras (Corrêa, 2019).

Os suprimentos que são utilizados na produção, podem ser organizados nos estoques por grupos, sobre o princípio de existir uma família de itens utilizados que compõe o produto e itens que são utilizados para a manutenção, como reparos de máquinas. Todos os suprimentos mapeados que

proporcionam o funcionamento da máquina devem estar disponíveis, para evitar a parada da produção (Stefanovic, 2015).

A gestão de inventário para ser eficaz e eficiente, deve estar programada para disponibilizar os itens necessários no período correto, evitar estoques obsoletos, gerar custos desnecessários para a organização (Namir et al., 2021) e trabalhar com o objetivo de disponibilizar o produto ao cliente.

A empresa para atingir o resultado deve alinhar diversas áreas e atividades que à integram, com o melhor *lead time*, cumprimento do cronograma de produção e não sofrer influência do efeito chicote⁶. Para isso, foi desenvolvido o gerenciamento de estoque eficiente, que deve ser definido e montado para atender ao objetivo principal do inventário, que consiste em disponibilizar os suprimentos para efetuar a entrega do produto final ao cliente (Corrêa, 2019).

Dessa forma, o gerenciamento de inventário deve ser construído para disponibilizar o produto no momento necessário, conforme a demanda e o *lead time* do processo, assim evita o acúmulo de componentes em períodos que não estão sendo processados e a falta do mesmo em períodos de necessidade. Com isso, a empresa apresentará como resultado um giro de estoque funcional dentro do período estimado (Tsou, 2013).

A gestão de estoque deve considerar a frequência com que ele é abastecido e qual a quantidade necessária de inventário (Guo et al., 2014), analisando-se os materiais disponíveis no almoxarifado e a compra de produtos faltantes, os quais são gerenciados pelo fator custo. O armazenamento de estoque insuficiente pode resultar na ineficiência produtiva da empresa, o que afeta no custo e na diminuição do lucro (Wang et al., 2018).

O gerenciamento eficiente do estoque é dificultado pela sincronização do real existente e da quantidade virtual, logo as empresas devem monitorar constantemente os almoxarifados para evitar a constatação de dados irreais (Wijffels et al., 2016). De acordo com Chen et al. (2019), um fator influenciador é a geração de cadastros duplicados, o que resulta em um material sendo gerenciado por dois endereços de almoxarifado.

Divergências no estoque físico e no sistema influenciam diretamente no resultado da empresa, devendo-se realizar a conferência no período padrão. A conferência do estoque, pode ser realizada de duas formas, sendo elas a cíclica, que consiste na análise e apuração de uma quantidade de materiais todos os dias, ou a inteira que compreende a conferência do estoque em um período curto (Wijffels et al., 2016).

O gerenciamento de inventário é estruturado sobre fundamentos de análise, como a demanda futura, que norteia os demais níveis de gerenciamento presentes na função, sendo eles a tomada de

⁶ O efeito chicote pode ser definido como o não alinhamento entre demanda e oferta que resulta em grandes hiatos entre demanda x produção (Coelho et al., 2009).

decisão do que comprar, a quantidade e o fornecedor correto (Bala, 2012). O sistema de gerenciamento de estoque efetivo, envolve quatro regras, descritas pelo armazenamento em local apropriado, monitoramento dos componentes, relatórios eficazes de estoque e controle efetivo dos processos e operações realizados no inventário (Govindasamy & Antonidoss, 2021).

O campo de pesquisa na área de gestão de inventário da cadeia de suprimentos está crescendo, as empresas estão buscando aprimorar a comunicação com os fornecedores, a fim de garantir valores e prazos de entrega com confiabilidade aos clientes (Fahimnia et al., 2019). Para isso, tem-se a investigação das atividades dessa área por meio da aplicabilidade de ferramentas de análise de dados que fazem uso da inteligência artificial e do aprendizado de máquina (Seeger et al., 2022).

2.2.1 Gestão de inventário e inteligência artificial

O fluxo de informações presente na gestão de inventário é complexo, necessitando-se de ferramentas de compilação e análise de dados a fim de oportunizar o alcance dos objetivos da área, tornando-a propícia para aplicação da inteligência artificial.

A inteligência artificial pode ser aplicada para auxiliar no desenvolvimento das atividades, alavancar os ganhos da organização (Roßmann et al., 2018), permitir o entendimento do estado do produto, proporcionar uma relação direta entre planejamento e programação de produção, e aumentar a qualidade dos produtos e serviços prestados (Y. Li et al., 2018).

A inteligência artificial é considerada um dos avanços dos últimos tempos, e vem sendo cada vez mais aplicada em modelos de negócios em corporações, pois pode melhorar a produtividade e possibilitar vantagem competitiva por meio do reposicionamento da organização no fornecimento de novos produtos e serviços (Y. Li et al., 2018).

A aplicação da inteligência artificial nas organizações é resultante da visão orientada sobre diversas áreas, do entendimento computacional, do reconhecimento da fala, além da tomada de decisão dirigida por dados, o que permite o emprego de inúmeras tecnologias como o aprendizado de máquina e o aprendizado profundo⁷ (Y. Li et al., 2018). A inteligência artificial tem sido aplicada na gestão de inventário para prever as demandas futuras, permitindo a preparação do estoque com suprimentos necessários (Giri et al., 2019).

O estoque de uma organização é baseado no planejamento estratégico de Roßmann et al. (2018), que deve ser estudado para se determinar a demanda de inventário, realizar a previsão de demanda e garantir a gestão integrada, a fim de se disponibilizar o material no prazo necessário ao

⁷ O aprendizado profundo, conhecido pela expressão em inglês deep learning, é um ramo do aprendizado de máquina que é capaz de entender e realizar classificações hierárquicas sobre informações complexas (Schmitt, 2023).

cliente e com o menor custo à empresa (H. Kumar et al., 2020). As organizações têm aplicado a inteligência artificial em seus estoques por meio do aprendizado de máquina para a previsão da demanda, redução dos estoques a níveis baixos e funcionais (Roßmann et al., 2018). Bala (2012) afirma que a inteligência artificial pode proporcionar à organização o posicionamento vantajoso com relação à gestão de estoque e ao aumento da lucratividade.

Os modelos gerados pela inteligência artificial podem auxiliar na identificação de padrões, na análise e previsão de demandas futuras, no entendimento de quais itens e a quantidade destes que devem ser adquiridos e armazenados, bem como no gerenciamento de incertezas para evitar a ocorrência de erros no processo decisório (Núñez-Merino et al., 2022).

A inovação tecnológica no gerenciamento do estoque pode permitir o desenvolvimento de mecanismos inteligentes que auxiliem na reposição de suprimentos de produção (Guo et al., 2014). A previsão de estoque deve ser baseada na compreensão de alguns fatores como, por exemplo, a produção da própria empresa, o volume e o custo do estoque, a distribuição do pedido, o tempo de recebimento e a forma de realização do pedido em lote ou unitário (H. Kumar et al., 2020). A tomada de decisão deve estar de acordo com os critérios estabelecidos pela própria empresa (Sundarakani et al., 2021), proporcionando o seu posicionamento no mercado e uma visão integrada por meio de métodos estatísticos e de inteligência artificial (Jeble et al., 2018). Assim, a inteligência artificial pode auxiliar no gerenciamento da cadeia de suprimentos (Roßmann et al., 2018) e proporcionar uma maximização dos ganhos, pois direciona a uma produção com menor valor agregado, o que pode aumentar os lucros (Bala, 2012).

Deng & Liu (2021) afirmam que a eficiência do ciclo do produto pode ser influenciada diretamente pela ciência de dados, que faz uso da inteligência artificial por meio de métodos de aprendizado de máquina, sendo responsável pela descoberta de conhecimento e padrões entre os dados, o que pode fortalecer as atividades de compra, estoque, produção, manutenção e distribuição.

2.2.2 Gestão de inventário e ciência de dados

A ciência de dados pode ser definida como a responsável pela identificação de conhecimento, utilizando-se de tecnologias para a modelagem da informação (Saritha et al., 2021), como data mining, big data e data analytics, que são aplicadas para entender e extrair informações concentradas em bases de dados (Sajid et al., 2021).

A ciência de dados quando aplicada na gestão de inventário da cadeia de suprimentos, transforma o modo de gerenciamento da área, possibilitando novos entendimentos de logística e ferramentas de gerenciamento para as organizações do século 21 (Stefanovic, 2015). Estas ferramentas permitem ao gerente um controle aprimorado de gastos e a rápida tomada de decisão

sobre informações precisas, que demonstram custos futuros e tendências empresariais (Hazen et al., 2014).

A integração entre a ciência de dados e a gestão de inventário possibilita ganhos produtivos ao setor, como proporcionar o giro do estoque de forma a disponibilizar nos períodos corretos a quantidade de materiais necessários, respeitando o orçamento da empresa e suprindo a necessidade de demanda da fábrica (Hamister et al., 2018). As análises de estoque realizadas por meio da ciência de dados são estruturadas com dados qualitativos e quantitativos, que descrevem e entendem o comportamento passado para prever o futuro (Waller & Fawcett, 2013) permitindo à empresa conhecer as dinâmicas efetivas dos processos que a integram e compreender o potencial para melhorias (Hamister et al., 2018).

Na gestão de inventário a atividade de logística de materiais internos é uma das mais estudadas pela ciência de dados, a fim de se determinar o local apropriado de armazenamento de produtos (Waller & Fawcett, 2013), o que pode ser influenciado por falhas relacionadas a qualidade do produto ou a mecanismos do processo na linha de produção. A ciência de dados aplicada à gestão de inventário, permite entender as falhas ocorridas durante o processo, e adequar a variabilidade da demanda e o balanceamento de carga (Diez-Olivan et al., 2019).

Nas empresas a circulação de suprimentos na linha de produção tem sido monitorada por meio da instalação de sensores que geram informações e as armazenam nos bancos de dados, sendo a ciência de dados aplicada no processo de transformação desses dados em conhecimento para auxiliar na tomada de decisão (Diez-Olivan et al., 2019).

O crescimento exponencial de equipamentos e operações, geradoras de dados instauradas na gestão de inventário, é responsável pela alta demanda de informações armazenadas em banco de dados, logo incentiva a expansão da ciência de dados, pois esta quando aplicada aprimora o tratamento da informação e proporciona a construção do inventário correto, impossibilitando estoques impróprios nos períodos estudados (Philip Chen & Zhang, 2014).

No meio empresarial o conhecimento descoberto utilizando-se a ciência de dados, são analisados para a tomada de decisão consciente de gestores das áreas, desta forma frequentemente são apresentados por meio de relatórios de tomada de decisão e de Key Performance Indicator (KPI), conhecidos como indicadores de produção, que demonstram os resultados produtivo, energético, estoque e econômico (Diez-Olivan et al., 2019).

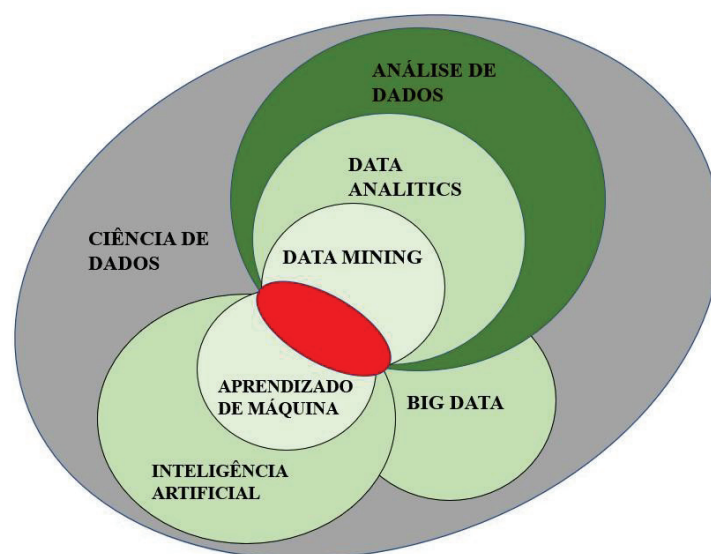
Nas empresas, no século 21, os dados são considerados o petróleo, pois possibilitam o entendimento das questões das organizações, auxiliam na tomada de decisão e na análise de risco (Han & Trimi, 2022). As abordagens existentes para o tratamento dos dados consistem na análise estatística, data mining, aprendizado de máquina, entre outras, as quais constituem a ciência de dados

(Philip Chen & Zhang, 2014). A ciência de dados vem crescendo nas organizações e os estudos apontam que o desenvolvimento tecnológico tem impactado nesta aplicação (Ijadi et al., 2018).

A ciência de dados é composta por diversas ferramentas e a junção delas permite a identificação de padrões úteis nas bases de dados (Figura 6), utilizando-se para isso de métodos que trabalham para mapear as oportunidades de transformar os dados em conhecimento (Seeger et al., 2022). Grandes empresas vêm trabalhando com a ciência de dados, para extrair conhecimento utilizando a mineração de dados (data mining), mas complementando com as demais ferramentas que integram a área.

Figura 6.

Apresentação das ferramentas que integram a ciência de dados.



Nota. Figura adaptada de 'Literature review on using data mining in production planning and scheduling within the context of cyber physical systems' de Seeger et al. (2022).

O gerenciamento de inventário pode ser explorado por ferramentas que integram a ciência de dados, como a mineração de dados Seeger et al. (2022), sendo este um campo de pesquisa e de aplicação emergente (Queiroz & Pereira, 2019). A mineração de dados permite a compreensão de padrões de dados armazenados e proporciona a extração deles, transformando os dados em conhecimento (Dawar et al., 2017).

2.3 TRANSFORMAÇÃO DIGITAL

O desenvolvimento de novas maneiras de pensar e suas estratégias são pilares da transformação digital, que está sendo aplicada nos negócios no intuito de aprimorar a comunicação com clientes, fomentar a competição entre as empresas do ramo, fortalecer a utilização de dados, incentivar a inovação e agregar valor ao produto oferecido (Rogers & Serra, 2017).

A transformação digital tem ocorrido na indústria, pois impulsiona o aumento da economia global pelas melhorias proporcionadas advindas da combinação tecnológica entre computadores,

comunicação e conectividade (C. H. Lee et al., 2021). O tema está sendo tratado como fator importante para o tratamento de informações e o aumento da eficiência em empresas manufatureiras (Felsberger & Reiner, 2020), pois garante maior visibilidade de parâmetros como a capacidade de processo, a antecipação de demanda e a colaboração entre as equipes (Núñez-Merino et al., 2022).

A indústria 4.0 é descrita como fortalecedora da implantação da transformação digital nas empresas (Muhuri et al., 2019a), que também se intensificou em função da pandemia da Covid-19, pois modelou com ferramentas tecnológicas as formas de ensino e de trabalho para a sociedade (Dwivedi et al., 2020). Nas organizações, ocasionou a mudança de cultura, alterando o modo de trabalhar e de gestão das empresas (Labucay, 2022).

O impacto causado pela Covid-19 foi fundamental para as organizações entenderem a necessidade de estarem agregadas aos dispositivos, técnicas e ferramentas tecnológicas que permitiram o funcionamento das mais diversas áreas (Dwivedi et al., 2020). Também, teve-se o fortalecimento do conceito de sustentabilidade nas organizações (Labucay, 2022), como por exemplo, pela adoção do home office em algumas funções o que proporcionou a diminuição da emissão de carbono, entre outros ganhos para a empresa e a sociedade (Dwivedi et al., 2020).

A transformação digital nas organizações modificou a execução das tarefas e operações, de modo a serem inteligentes e gerarem informações (Gunduz et al., 2021). Esta evolução permite a valorização do negócio e do atendimento ao cliente, pois faz uso da inteligência artificial, do aprendizado de máquina, da ciência de dados, da automação de processos, da internet das coisas, entre outras tecnologias, que quando alinhadas aos objetivos organizacionais, possibilitam a sua recolocação estratégica (ElMaraghy et al., 2021) e garantem a tomada de decisão com mais agilidade e assertividade (Felsberger & Reiner, 2020).

Nos países desenvolvidos, as empresas estão trabalhando na atualização dos parques para agregar a transformação digital, pois entendem que assim podem continuar competitivas nos negócios (Felsberger & Reiner, 2020) em uma perspectiva multifuncional, pois aplica-se ao campo estratégico, de avaliação estrutural e de capacidade digital (Ciampi et al., 2022).

A transformação digital nas cadeias produtivas pode ser facilitada, iniciando-se pela cadeia de suprimentos (Ciampi et al., 2022), resultando em ganhos à organização como a sincronização das operações e das funções de gerenciamento de suprimentos alinhadas com a sustentabilidade (Gunduz et al., 2021). Além disso, segundo Ciampi et al. (2022), este é um dos setores que se encontra preparado para a transformação digital e necessita dela para a sua sustentação no mercado, em especial, para a gestão de inventário, visto que esta influencia diretamente nos resultados da empresa.

2.4 CIÊNCIA DE DADOS

A ciência de dados, é comparada com as mudanças causadas pela eletricidade no século 19, descritas como disruptivas por alguns autores, pois permitiu a inovação de serviços e na vida social. A ciência de dados oportuniza a gestão, nas empresas viabiliza o processamento e a modelagem dos dados gerados pelas vias tecnológicas e permite a extração de conhecimento. O crescimento da área está diretamente interligado com a expansão tecnológica exponencial (Sajid et al., 2021).

A ciência de dados analisa as informações nas bases de dados, independente dos padrões e origina o conhecimento (Sajid et al., 2021), devendo ser realizado de forma ética, cumprindo as normas (Tilimbe, 2019). No caso do Brasil, deve-se atentar para a Lei Geral de Proteção de Dados.

Figura 7.

Perspectiva da ciência de dado para a extração do conhecimento de bases de dados.



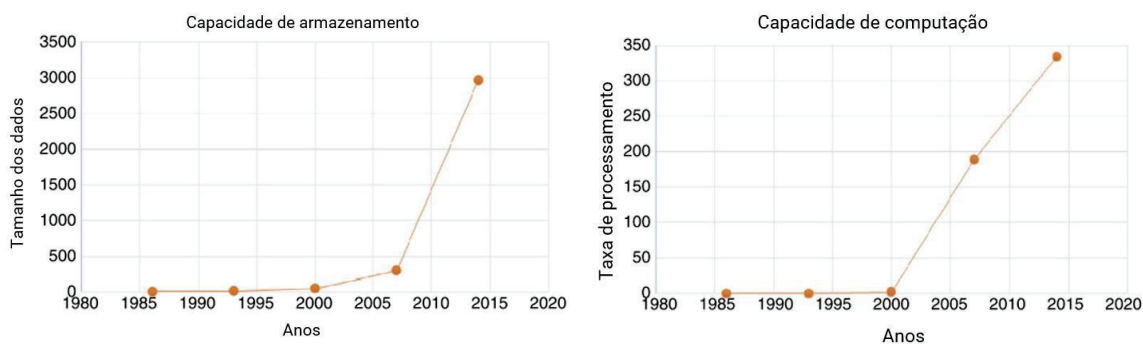
Nota. Figura adaptada 'Open source technologies in data science and big data analytics' de Saritha et al. (2021).

Os grupos de dados suportam um conjunto robusto de informações, derivadas de diferentes plataformas e formatos distintos (Waller & Fawcett, 2013), porém a ferramenta aplicada permite o entendimento dos dados e suas informações (Figura 7), transcrevendo-os em insights e conhecimento (Boehmke et al., 2020).

O aumento expressivo de dados gerados, alavancou a transformação digital, construída sobre o aprimoramento de ferramentas tecnológicas, capazes de realizar a leitura dos parâmetros das atividades executadas no mundo digital, auxiliar no entendimento e na descrição das informações (Maass & Storey, 2021).

Figura 8.

Capacidades de armazenamento e computacional entre 1986 e 2014.



Nota. Figura adaptada de 'Expanded cloud plumes hiding Big Data ecosystem' de Sharma (2016).

O crescimento da circulação de informações nos últimos anos, demonstra a necessidade de banco de dados mais robustos, conforme apresentando na figura 8, pois o processamento do conhecimento presente nesses bancos pode ser atrelado ao aumento de dados complexos e heterogêneos que necessitam de maior capacidade de computação para a modelagem do conhecimento existente (Sharma, 2016).

Os dados gerados, armazenados em sistemas gerenciadores de banco de dados e que não são explorados se tornam inúteis (Boehmke et al., 2020). Estes dados são volumosos, o que torna impossível o seu gerenciamento com as ferramentas antes utilizadas, desenvolvendo-se assim, a ciência de dados (Waller & Fawcett, 2013).

A área de ciência de dados é bastante pesquisada no google, com mais de 76 milhões de acesso (Sajid et al., 2021). Alguns autores defendem que o avanço da ciência de dados está correlacionado com o volume de dados gerados pelas redes sociais, que dissipam informações rapidamente (Sarmah & Moharana, 2017).

A ciência de dados foi desenvolvida para a construção de estratégias, sobre a avaliação dos dados e permitir um melhor gerenciamento das áreas que estão sendo avaliadas (Han & Trimi, 2022). A sua utilização nas organizações é fundamental para minimizar a dificuldade, ainda existente, em trabalhar com os dados gerados (Bellomarini et al., 2022; Boehmke et al., 2020), pois realiza a modelagem da informação extraída da empresa e de bancos de dados tangenciais para proporcionar o gerenciamento efetivo e a tomada de decisão ágil (Han & Trimi, 2022), disponibilizando relatórios via gráficos e modelos estatísticos que permitem o esclarecimento de parâmetros (Bellomarini et al., 2022). No entanto, a ciência de dados ainda não vem sendo utilizada com muita frequência em empresas de médio e pequeno porte, em função do seu custo de aplicação (Angeles, 2016), já em grandes empresas, como a Amazon, é utilizada em diferentes setores estratégicos a fim de proporcionar eficiência (Sharma, 2016).

O problema só pode ser estudado e avaliado pela ciência de dados, após a fundamentação real da questão a ser respondida. Portanto, antes de minerar e interpretar os dados, é necessário entender as variáveis e verificar a viabilidade de aplicação da ciência de dados (Boehmke et al., 2020). Dessa forma, a ferramenta pode proporcionar insights sobre os dados quantitativos e qualitativos, fornecendo análises de qualidade (Schoenherr & Speier-Peró, 2015).

A ciência de dados tem sido aplicada em diferentes áreas como a gestão pública, centros médicos e em empresas de pequeno e grande porte (Tsui et al., 2019). Nos Estados Unidos é muito utilizada no *National Institutes of Health* (NIH) e no *National Science Foundation* (NSF) para auxiliar na tomada de decisão (Philip Chen & Zhang, 2014).

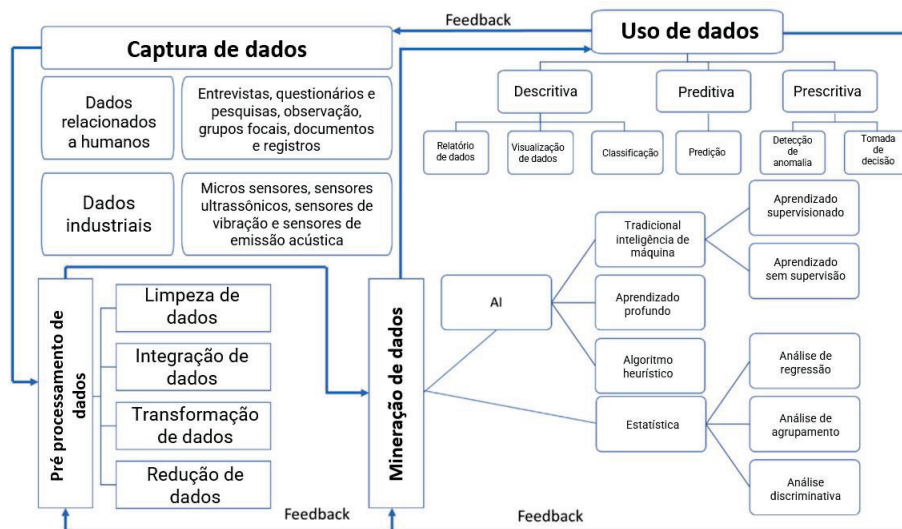
Nas organizações a ciência de dados vem sendo utilizada para aprimorar o processo decisório e a análise de risco, como por exemplo, na cadeia de suprimentos, onde é relacionada para compreender eventos internos e externos que influenciam diretamente na sua demanda produtiva (Salamai et al., 2019), utilizando-se da inteligência artificial por meio do aprendizado de máquina, para proporcionar análises em tempo real (Misra et al., 2022) e o acesso via mineração de dados as informações mapeadas como fundamentais (Rautenberg & Carmo, 2019).

A aplicação de métodos da ciência de dados para a tomada de decisão nos setores que compõe uma empresa, pode proporcionar melhorias como na eficiência, na entrega do produto final ao cliente, que pode ser facilitado com informações sendo atualizadas em tempo real da cadeia de suprimentos, fortalecendo a relação entre os envolvidos (Núñez-Merino et al., 2022).

Na cadeia de suprimentos, o processo de tomada de decisão objetiva melhorar a capacidade de armazenamento e produção, otimizar a utilização de recursos na linha produtiva, entender as demandas futuras para evitar o efeito chicote, viabilizar as lacunas dos setores e reduzir o custo empresarial (Schoenherr & Speier-Peró, 2015).

O suporte à tomada de decisão deve ser realizado sobre dados seguros, para direcionar os gestores com segurança, estabilidade e acessibilidade (Philip Chen & Zhang, 2014). No entanto, os dados presentes no conjunto não são todos viáveis, sendo denominados de “dados sujos”, os quais devem ser evitados durante uma análise de sistema (Ridzuan & Wan Zainon, 2021), pois impactam diretamente na gestão estratégica das empresas e inviabilizam a tomada de decisão (Hazen et al., 2014).

Figura 9.
Plataforma da ciência de dados.



Nota. Figura adaptada de 'Towards a data science platform for improving SME collaboration through Industry 4.0 technologies' de Han & Trimi (2022).

A figura 6 relata a necessidade da comunicação entre as demais atividades que permitem o entendimento do fluxo de informação presente em banco de dados. De acordo com Fayyad et al. (1996) existe uma necessidade urgente das atividades de descoberta de conhecimento em bases de dados estarem conectadas e se sustentando, pois somente assim será possível a extração de conhecimentos úteis e claros para interpretação.

2.5 MINERAÇÃO DE DADOS

A mineração de dados é definida como uma técnica para descoberta de conhecimentos ocultos e importantes em bases de dados (Liou et al., 2021), utilizando um conjunto de algoritmos e modelos estatísticos para entender os padrões de comportamento (Arivazhagan et al., 2022) e interpretar as previsões. Os dados são analisados por meio de uma série de tarefas, resultando no processamento e na geração do conhecimento a fim de auxiliar na tomada de decisão (Granillo-Macías, 2020).

O fortalecimento e a predominância da mineração de dados na indústria moderna, tem relação com os seus resultados e aplicabilidade, pois permite a identificação de padrões nos dados, dissipando a gestão do conhecimento nas organizações e auxiliando na tomada de decisão (Seeger et al., 2022) para o gerenciamento do processo que é assegurado pelos insights, disponibilizados (Rudnitckaia et al., 2022).

A mineração de dados permite visualizar as variáveis da organização (Maass & Storey, 2021) examinar demandas e identificar clientes (Deng & Liu, 2021), trabalhando com dados antigos e novos, o que resulta em uma linha do tempo sobre os fatores que estão sendo avaliados (Angeles, 2016).

Figura 10.

Tarefas executadas pela mineração de dados.



Nota. Figura adaptada de 'Literature review on using data mining in production planning and scheduling within the context of cyber physical systems' de Seeger et al. (2022).

A figura 10 apresenta as tarefas que constituem a mineração de dados, sendo elas classificação, regressão, agrupamento, associação e a previsão de séries temporais (V. Kumar et al., 2009). A classificação é fundamentada sobre a tarefa de modelar e permitir a relação entre o conjunto de dados, de acordo com a sua característica e objetivo (L. Huang et al., 2018). O agrupamento trabalha para identificar as similaridades entre os dados e excluir duplicatas na execução do processo. O entendimento de relações entre dados de um grupo é determinado pelas regras de associação (Mariano et al., 2020).

O conjunto de tarefas permite o cruzamento de informações de padrões diferentes, derivadas de uma ou mais bases de dados, no intuito de transformar os dados em conhecimento (Dawar et al., 2017). A mineração de dados é considerada como efetiva quando é possível a identificação das tarefas de classificação, agrupamento e associação entre os dados estudados (Arivazhagan et al., 2022). Entender o padrão dos dados, é importante para o resultado da mineração de dados, que deve trabalhar na modelagem do conhecimento (J. Chen et al., 2019) extraído da base de dados por meio de uma série de análises, utilizando algoritmos e softwares computacionais (Dawar et al., 2017).

As áreas de mineração de dados e de estatística se diferenciam pela forma como os dados são utilizados para a extração da informação. A mineração de dados realiza as descobertas sobre o assunto proposto, demonstrando ângulos e questões que não tinham sido identificadas anteriormente. A estatística é utilizada para fortalecer algo que já está sendo visto, ela trabalha como análise de números de situações (Senthilkumaran et al., 2016) e com dados específicos, porém a fidelidade do resultado depende da qualidade e do grau de confiabilidade dos dados, pois não realiza o cruzamento de informações para a validação de dados, como ocorre na mineração de dados (Deng & Liu, 2021). Destaca-se que a estatística e a mineração de dados se complementam na modelagem coerente dos resultados e na demonstração clara dos resultados encontrados nas bases de dados (Granillo-Macías, 2020).

O desenvolvimento da ciência de dados com a transformação digital, permitiu o avanço de ferramentas que compõe a ciência, como a mineração de dados, definida como recurso de avaliação de dados e no meio corporativo permite a recolocação estratégica de empresas no mercado (Balali et al., 2015) o recurso se associou bem com a gestão de inventários, pois viabiliza o entendimento da informação da circulação de suprimentos e analisa o banco de dados gerado durante as atividades processo, criando oportunidades constantes de análises de dados e cruzamento de informações (Liou et al., 2021).

A consolidação da tomada de decisão efetiva no processo de gestão de inventário, é justificada com a aplicação de mineração de dados, a ferramenta possibilita a visualização de itens caros e com pouca rotatividade em estoque e permite a conexão com a demanda (Bala, 2012). Nesse processo, a mineração de dados consegue identificar os principais clientes de um produto e os principais fornecedores de um suprimento, permitindo o fortalecimento do setor de inventário, que tem como foco trabalhar com menor custo, para alavancar o lucro da organização e desenvolver com eficiência as tarefas de compra, estoque, produção e distribuição (Deng & Liu, 2021).

A mineração de dados está sendo adotada para prevenir possíveis rupturas na estratégia adotada, direcionando a ferramenta para o fortalecimento das previsões e do volume ideal de suprimentos nos estoques, já que a ferramenta tem como princípio trabalham sobre situações reais e dados dinâmicos (V. Kumar et al., 2009).

De acordo com Seeger (2022), a mineração de dados tem sido empregada com frequência na indústria moderna para a gestão de processos, pois pode disponibilizar conhecimentos precisos e complexos acerca de fatores que influenciam diretamente no posicionamento estratégico da empresa (Rudnitchkaia et al., 2022).

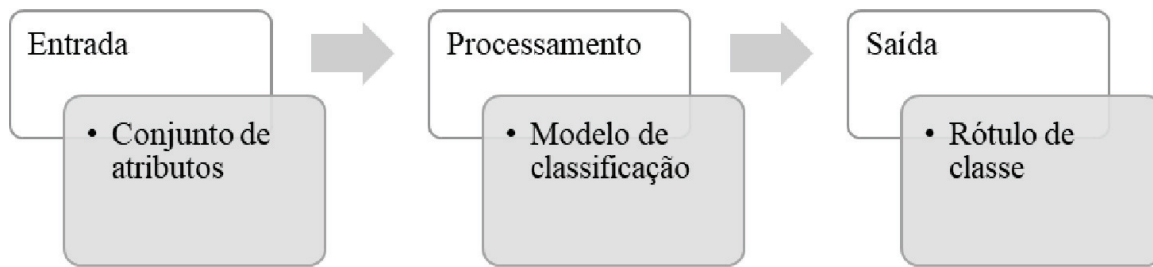
A mineração de dados por meio da aplicação das suas tarefas e algoritmos pode em um modelo de negócio demonstrar, por exemplo, a previsão de estoque do inventário de uma organização, identificando os itens parados, os fatores que dificultam o desenvolvimento do setor (L. Li et al., 2014), resultando no gerenciamento de inventário efetivo e enxuto (Bevilacqua et al., 2019).

2.5.1 Classificação

A classificação é construída sobre o conceito de fluxo dos dados entre o conjunto de atributos, que permite a concentração dos registros com características similares. O tratamento e o entendimento da informação são realizado pelos algoritmos, que atuam na construção dos atributos e classes para os dados e os classifica em grupos distintos (Goldschmidt & Passos, 2005).

Figura 11.

Fluxo de informação no modelo de classificação.



Nota. Figura adaptada de 'Introdução ao Data Mining. Mineração de Dados' de Kumar et al. (2009).

A transformação digital quando atribuída ao ambiente corporativo, tem permitido a geração de grandes volumes de dados, derivado da utilização de ferramentas tecnológicas em atividades diárias. Conjuntos de informação armazenado em um banco de dados, são possíveis geradores de conhecimento, que podem ser extraídos quando aplicadas técnicas de data mining (Fosso Wamba et al., 2021), para resultar em diferentes tipos de parâmetros de acordo com a tarefa que está sendo empregada (Papanagnou & Matthews-Amune, 2018).

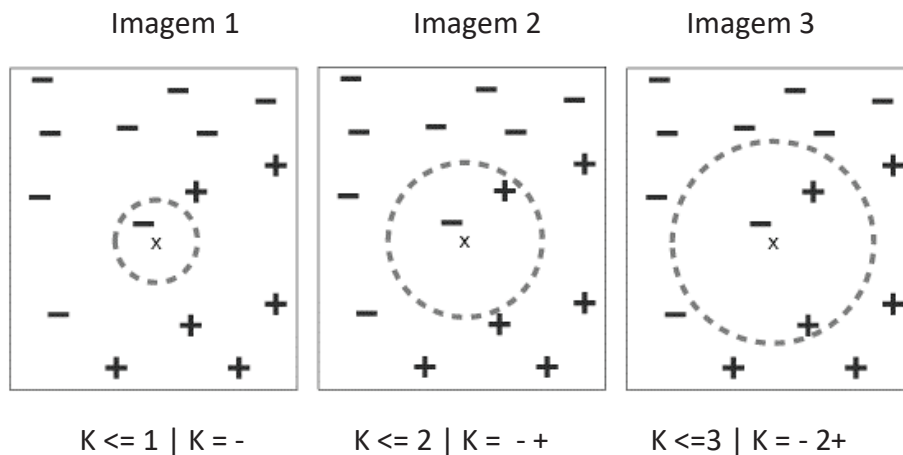
A classificação, segundo Sajid et al. (2021), é composta por diferentes métodos como os classificadores Bayesianos, as árvores de decisão, o vizinho mais próximo e máquina de vetores de suporte, do inglês Support Vector Machine (SVM). Cada método é composto por diferentes algoritmos que permitem o entendimento dos padrões presentes na base de dados, realizando a classificação dos itens que têm forte correlação (Han & Trimi, 2022).

As árvores de decisão são construídas por uma série de questionamentos, como filtro dos critérios que estão sendo avaliados, geralmente é desenvolvida com dados no formato de uma pergunta, com duas possíveis respostas, sim ou não, e esses questionamentos realizam a exclusão ou a permanência do dado analisado, e assim sucessivamente, até encontrar o conjunto de informações que passaram por todos os critérios. Os algoritmos mais utilizados para a construção da árvore são ID3, C4.5, SLIQ e SPRINT (V. Kumar et al., 2009).

O método do vizinho mais próximo entende que "se caminha como um pato, se come comida de pato e se comporta como um pato, então provavelmente é um pato". Este método calcula o ponto de distanciamento entre as informações, e compreende que quanto mais próximo os atributos que a compõe, maior o número de informações similares, proporcionando a formação de famílias por afinidade de informações (Mariano et al., 2020).

Figura 12.

As imagens representam os vizinhos mais próximos de um ponto no centro, o raio do círculo que abrange os vizinhos é definido por k .



Nota. Figura adaptada de 'Introdução ao Data Mining. Mineração de Dados' de Kumar et al. (2009).

Os classificadores Bayesianos são construídos sobre a mensuração via probabilidade das variáveis que a compõe, quanto maior a probabilidade de um item acontecer maior o agrupamento entre o conjunto de informações que tem resultados próximos ou iguais, na base de dados. Para o cálculo é utilizado $P(X, Y) = P(Y|X) \times P(X) = P(X|Y) \times P(Y)$ e os resultados são tratados por algoritmos, como por exemplo, Naive Bayes, Selective naive Bayes, Seminaive Bayes, One dependence Bayesian, k-dependence Bayesian, Bayesian multinets (V. Kumar et al., 2009).

O support Vector Machine (SVM) é composto por um conjunto de equações que resultam vetores, e tem como tarefa definir uma área específica da base de dados e determinar o agrupamento dos componentes Mariano et al. (Mariano et al., 2020), todos os dados que ficarem nas margens estabelecidas, foram classificados pela aplicação dos algoritmos GSVM, FSVM, TWSVMs, VaR-SVM, RSVM (V. Kumar et al., 2009).

2.5.2 Associação

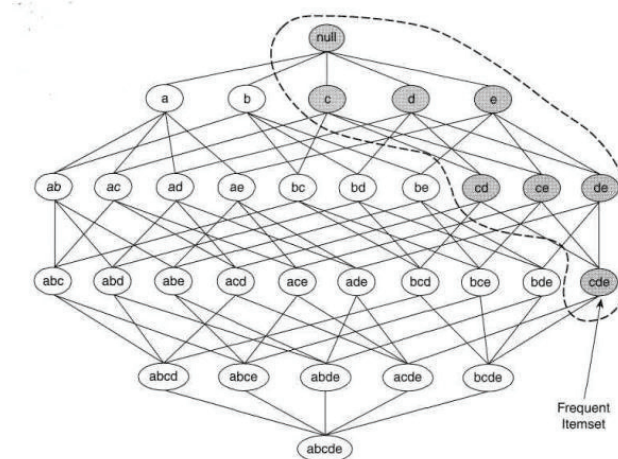
A análise de associação aplica os algoritmos que a compõe para entender os relacionamentos de interesse que existe na base de dados que está sendo estudado. O relacionamento de interesse são informações construídas sobre o suporte que determina a frequência com que tal tarefa acontece no período, e a confiança que determina a frequência que o item A apareceu agrupado na movimentação do item B, como na compra do lápis o cliente também irá comprar a borracha, pois os dois tem um grau associativo dependente com correlação forte (Mariano et al., 2020).

A regra de associação é construída sobre diversos algoritmos como, sequência, sequência temporal, paralelo e outros onde tratam os dados em diferentes formas e analisam o grau de associação das informações (Goldschmidt & Passos, 2005).

O método de associação foi criado para o entendimento de correlações existentes no varejo, a ferramenta proporciona a visualização dos itens que apresentam maior rotatividade no estoque e quais materiais estão diretamente relacionados a eles (Arivazhagan et al., 2022). O cálculo das regras de suporte e confiança são realizadas por algoritmos, sendo eles a metodologia de frequentes que utiliza a Apriori e Pattern growth algorithms, a temporal sequence, parallel e other (F. Chen et al., 2015).

O algoritmo Apriori identifica se um conjunto de itens for frequente então todos os seus subconjuntos também devem ser frequentes, com esse entendimento o método reduz ao máximo o número de conjuntos presentes no banco de dados, analisando apenas os dados com atributos que possuem correlação com o item analisado (V. Kumar et al., 2009).

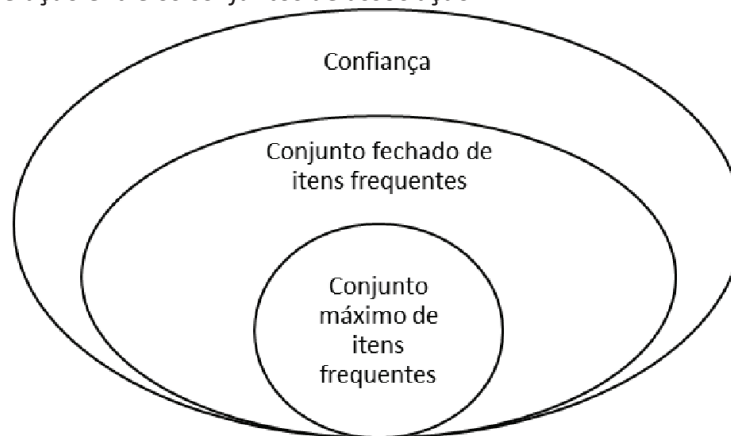
Figura 13.
 Algoritmo Apriori



Nota. Figura de Kumar et al. (2009).

A confiança da associação é arbitrada pela frequência de correlação apresentados, o método realiza a identificação dos pequenos conjuntos que podem ser demonstrados por máximos frequentes e irá possibilitar demais combinações entre os dados, ou pode ser identificado como conjunto fechado de itens frequentes que são utilizados para remover regras de associação similares (Mariano et al., 2020).

Figura 14.
Demonstração da relação entre os conjuntos de associação.



Nota. Autores (2022).

A metodologia paralela, é idealizada sobre dois algoritmos Partition based e FP-Growth, construído para a codificação do conjunto de dados usando uma estrutura de informação compactada denominada de árvore de FP, quando analisada é possível a visualização de transações que se sobrepõe pela frequência dos itens, e o método define que quanto mais sobreposições estiverem presentes maior o grau de associação (Kumar et al., 2009).

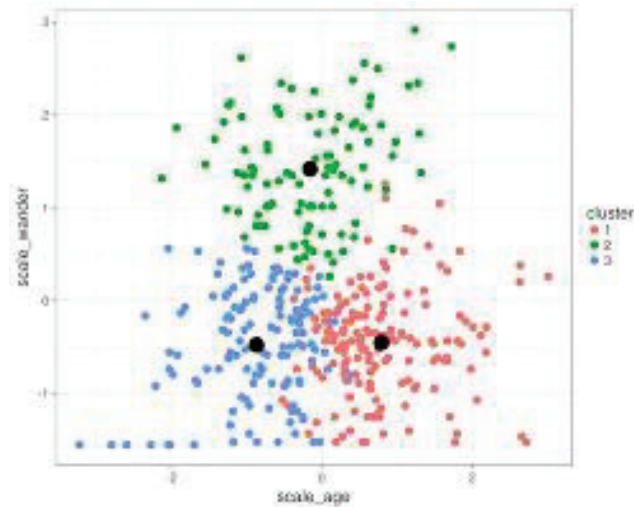
2.5.3 Agrupamento

A tarefa de agrupamento é realizada para aproximar fatores em comum e gerar famílias de conhecimento e informações similares (Kumar et al., 2009) que são identificadas na utilização de algoritmos específicos, e realizam o entendimento dos dados e a construção de agrupamentos ou clusters (Li et al., 2014).

O algoritmo K-means é o algoritmo mais antigo, quando analisado os métodos de agrupamento, o método trabalha com a definição de centros e determina que cada coleção que está na distância k do centro é um grupo, o processo é refeito diversas vezes até que a informação analisada fique estável e não mude de centroide (Mariano et al., 2020).

Figura 15.

Representação da utilização da técnica K-means para localizar informações similares no espaço.

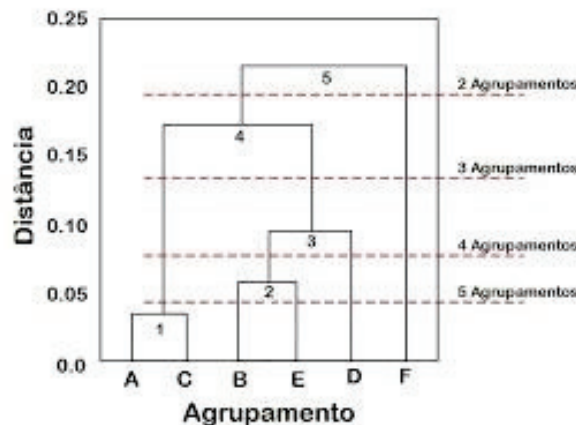


Nota. Figura de Yobero (2018).

O método de agrupamento hierárquico é desenvolvido para a apresentação dos relacionamentos e das dependências entre grupos, demonstrando a ordem com a qual os grupos são fundidos ou divididos e a subdivisão desses grupos são representados por aglomerados, ou seja, quando a um agrupamento dos grupos mais próximos em cada etapa analisada, ou pelo modelo de divisão que é antônimo ao método anterior, ele inicia com um grupo centralizado e a cada etapa se subdivide, até chegar no parâmetro analisado (Kumar et al., 2009).

Figura 16.

Demonstração do agrupamento hierárquico via dendrograma.



Nota. Figura de 'Análise de cluster não supervisionado em R: agrupamento hierárquico' de Ferreira et al. (2020).

O agrupamento hierárquico foi desenvolvido para trabalhar com grandes números de dados, onde é aplicado os algoritmos para realizar os dimensionamentos dos grupos e minimizar o número

de cálculos de proximidade entre os centroides, e possibilitando um agrupamento resumido do alto número de dados (Goldschmidt & Passos, 2005).

O dimensionamento do conteúdo presente na base de dados, trabalha para reduzir o número total de informações em uma menor dimensão, que permite chegar no mesmo resultado que trabalhar com o banco de dados inteiro, porém acaba se tornando melhor o processamento de menos informação. Para esse mecanismo são utilizadas algumas ferramentas como a Análise de valores singulares (PCA) e Decomposição de Valores Singulares (SVD), que desenvolvem o conceito de média sobre o número de dados para análise e a construção de matrizes de informação (V. Kumar et al., 2009).

2.6 ANÁLISE DE VALOR

A análise de valor é predominantemente empregada na área econômica, uma vez que a aplicação correta desse método proporciona uma visão confiável sobre os aspectos analisados. Nos últimos anos, a análise de valor tem sido amplamente utilizada no mercado americano e está ganhando crescente relevância no Brasil. Diversos autores destacam a importância de utilizar a análise de valor na avaliação de diferentes cenários, pois acredita-se que essa abordagem pode aprimorar a execução de projetos e, conseqüentemente, gerar resultados diferenciados para a organização. (Vargas & Abba, 2018).

No contexto dos diferentes ramos empresariais, a análise de valor pode ser aplicada para compreender o comportamento de diversos fatores, não apenas os econômicos, no qual é predominantemente utilizada dentro das empresas. A análise de valor é defendida como uma técnica de análise que pode ser adaptada à ótica do problema a ser avaliado, com o objetivo único de identificar ineficiências no processo e melhorá-lo. (Miles, 2015).

A construção da análise de valor baseia-se em um conjunto de disciplinas que visam alcançar o comportamento desejado pelo cliente. Para isso, o analista deve compreender os procedimentos e processos, além de realizar uma avaliação do ambiente onde a ineficiência se encontra. Somente então será possível identificar os sistemas específicos e aplicar as técnicas apropriadas para a análise da situação determinada. (Miles, 2015).

A análise de valor em uma empresa é conduzida por meio de uma análise estrutural da estratégia de negócios, na qual se identifica um ponto de ineficiência do processo, independentemente do setor. Essa avaliação deve ser realizada conforme a estrutura definida pela organização, respeitando os sistemas que a integram e sempre buscando a visão estratégica do negócio. Isso inclui a análise da concorrência e dos fornecedores, sempre vinculada à interpretação dos sinais do mercado. Quando a análise é realizada sob essa ótica, pode promover a competitividade da organização. (Porter, 2005).

2.7 ESTADO DA ARTE

Nesta pesquisa o estado da arte foi concebido a partir do estudo bibliométrico, seção 2.1, desenvolvida para identificar o comportamento das publicações relacionadas à tomada de decisão utilizando a ciência de dados, na gestão de inventário. Para isso, realizou-se o mapeamento e a análise das publicações correlacionadas ao tema, em diferentes bases de dados científicas.

A bibliometria é uma forma utilizada para tratar um alto volume de pesquisas e extrair informações importantes, além de detectar vieses de estudos (Ciampi et al., 2022), medir a qualidade e a quantidade de publicações referentes a área e visualizar a influência dos autores e das universidades em determinado período (Doğan & Karaca, 2020). A bibliometria é um método estatístico e qualitativo, que permite a observação da evolução da área de estudo (Sá et al., 2020).

Mediante o exposto, o presente estado da arte é composto pela elaboração do estudo bibliométrico e uma revisão das pesquisas realizadas mais recentemente no Brasil e no mundo, considerando-se como critério os cinco estudos mais citados, e a pesquisa com mais similaridade com o presente estudo.

Esta seção da pesquisa, apresenta seis trabalhos correlatos, os quais são apresentados no Quadro 1, seguido pela descrição desses estudos.

Quadro 1.

Apresentação dos seis trabalhos mais citados no estudo bibliométrico

	(Lee et al., 2015)	(Balali et al., 2015)	(Kisilevich et al., 2013)	(Bala, 2012)	(Abbasi et al., 2020)	(Bevilacqua et al., 2019)
Nº de citações	35	42	35	30	27	14
País de origem	Hong Kong	EUA	Alemanha	Índia	Austrália	Itália
Nº de dados	177	10.000	168	10418	180	16317
Origem dos dados	Histórico das bombas de fusão de um hospital	6,2 milhas em dois segmentos das rodovias interestaduais I-57 e I-74 dos EUA	Travel Global Systems (TGS) / quarto standard	pacote de 5 kg de Atta	Estoque de sangue, para o abastecimento de quatro hospitais	Estoque de uma fábrica de sapatos
Período dos dados	2010 a 2020		check-in de 7 a 21 dias	treinamento de 36 meses e um conjunto de dados de teste de 12 meses	A simulação é executada em um período de 18.500 dias,	Avaliou a m ² de armazenamento e agrupamento do estoque
Dados utilizados	Estoque do setor de saúde	Mapeamento de sinais de trânsito e seu comportamento	Histórico de hotéis em um determinado espaço	SKU de varejo	Grupo sanguíneo, O+	Dimensões de bandejas e número de itens que são alocados nelas
Método/Algo ritmo	Rede neural artificial (ANN) / Lógica Fuzzy	Histogram of Oriented Gradients + Color	Isotonic regression / LibSVM epsilon-SVR / LibSVM nu-SVR / Linear regression / Locally weighted learning with linear regression / Additive regression with isotonic regression / Multilayer Perceptron	K-means / agrupamento e ARIMA	ANN / K-NN / CART / RF	FP-Growth
Medidas de qualidade	Mean Absolute Deviation (MAD) / Mean Squared Error (MSE)	Gráfico de histograma	Regressão linear, não linear e não paramétrica	Pearson correlation-based	Programa linear inteiro estocástico	

Validação interna	middleware	Verdadeiro Positivo (TP), Falso Positivo (FP) e Falso Negativo (FN) / SVM linear	Erro Absoluto Médio (MAE) e Erro Quadrado Médio (RMSE)	Training MAPE ou NMSE / Teste set MAPE ou NMSE	MSE / Gráfico de histograma	Ferramenta de simulação
Ferramenta	Matlab	Google Street View / API Google Direction / Google Earth	GIS / Java OpenStreetMap Editor (JOSM) / Software R / PostgreSQL/PostGIS / Weka / MDS	SPSS Modeler		RapidMiner
Resultado	Comparação entre o método ANN e FL	Emissão de boletins para análise dos sinais de trânsito.	Avaliação de hotéis considerando características físicas das instalações e de localização	A utilização do método WMB é eficaz	Cada modelo aplicado gerou um resultado ótimo para somente uma variável analisada	Agrupar os itens para ter uma melhor organização do espaço e evitar grandes deslocamentos do almoxarife
Conclusão	Lógica Fuzzy é mais eficiente	Possibilidade de identificar substituição ou instalação de novos sinais de trânsito	implementação de um sistema de suporte à decisão baseado em GIS para analisar o valor do hotel e estimar tarifas objetivas	o modelo proposto é apropriado para itens para os quais uma proporção substancial de venda é fortemente atribuída a um perfil ou perfis específicos	Possibilidade da utilização de todos os modelos apresentados para o gerenciamento de todas as atividades, porém cada um atua sobre uma variável	Resultado validado
Trabalhos Futuros	Melhorar a produtividade da gestão de estoques e reduzir o custo operacional.	Automatizar o processo de coleta de dados para bancos de dados ESRI ArcView GIS	identificar problemas que ainda não foram cobertos pelo protótipo atual	O uso dos clusters fracos dos registros de vendas para previsão de demanda	abordar o problema do ponto de vista meta-heurístico de otimização baseada na evolução e usar as decisões obtidas a partir de modelos de aprendizado de máquina como soluções iniciais de meta-heurísticas ou algoritmos matemáticos	proporcionar a aplicabilidade do método a demais ramos industriais

Nota. Autores (2022).

Lee et al. (2015) realizaram uma pesquisa sobre o estoque de bombas de fusão hospitalar, no período de 2010 a 2020 em um hospital em Cingapura. Para o entendimento dos dados foram utilizados os métodos Rede Neural Artificial (ANN) e Lógica Fuzzy (FL). A ANN considera a previsão de ativos rotineiros, já a FL trabalha na identificação de ativos não programados, ou seja, operações inesperadas como emergências. O estudo é construído sobre 177 regras, referente ao histórico de bombas de fusão hospitalar coletado de dois dispositivos o identificador de rádio frequência RFID e rede de sensores sem fio WSN, que conectam e enviam os dados brutos para o middleware. O pré-processamento de dados foi realizado com métodos de agregação, filtragem e normalização. A análise da informação foi realizada por meio de redes neurais artificiais e de lógica fuzzy, usando a ferramenta Matlab. A qualidade das análises realizadas, a fim de demonstrar o resultado mais aproximado com o real, bem como comparar os modelos gerados por meio da rede neural e da lógica fuzzy realizou-se por meio das medidas de Mean Absolute Deviation (MAD) / Mean Squared Error (MSE), desta forma o autor conseguiu ver o comportamento entre previsto e a real demanda de um período anterior. A MAD e MSE demonstraram que a lógica fuzzy é o modelo ideal para o trabalho. Para futuras pesquisas os autores destacam a importância de melhorar a produtividade da gestão de estoques e reduzir o custo operacional na área hospitalar.

Balali et al. (2015) criaram e mapearam sinais de trânsito usando imagens do Google Street View para identificar possíveis substituições ou instalações de sinais de trânsito na rodovia estadual que atravessa o Tennessee nos Estados Unidos. Utilizando o processamento de fotos da API do Google Street View e Google Earth, o método de visão computacional, identifica os sinais de trânsito que são categorizados em: sinais regulamentares, de advertência, de parada e de preferência. A análise de dados é realizada em um trecho de 6,2 milhas em dois segmentos das rodovias interestaduais I-57 e I-74 dos Estados Unidos e captou 10.000 dados provenientes de sinais de trânsito. A estratégia de avaliação das imagens é polilinha para representar a visão do condutor durante o trajeto, e aplicando-se o classificador de máquinas de vetor de suporte (SVM) linear para detecção e classificação de sinais de trânsito e o Histogram of Oriented Gradients (HOG) + Color para a mensuração dos resultados, derivando-se um histograma de distribuição de cores, para caracterizar as cores locais em cada sinal de trânsito. Como resultado foi desenvolvido um boletim de dados, que permite o acompanhamento automático do comportamento dos sinais de trânsito. Para futuras pesquisas citam a automatização do processo de coleta de dados para bancos de dados ESRI ArcView GIS.

Kisilevich et al. (2013) desenvolveram uma aplicação computacional para auxiliar na tomada de decisão referente ao melhor hotel para se realizar uma reserva, considerando critérios de hóspedes e atendimento das instalações. Na demonstração foram utilizados dados de históricos de preços de quartos standard de 168 hotéis que foram divididos em componentes estatísticos e dinâmicos. Os

dados estatísticos incluem os nomes dos hotéis, seus IDs internos, suas coordenadas de localização no World Geodetic System, instalações do hotel, comodidades dos quartos e categorias de hotéis. A componente dinâmica inclui os preços de uma noite que os clientes receberam durante a sua procura de alojamento, a data da procura e a data da reserva. Os dados são utilizados para promover um comparativo entre as instalações e dispor do melhor local de acordo com a preferência do hóspede. Na construção da ferramenta foram utilizados os componentes como Java OpenStreetMap Editor (JOSM), Software R, PostgreSQL/PostGIS, Weka, MDS e as representações dos resultados via gráficos. Para a mensuração dos dados encontrados foi utilizado os critérios de Erro Absoluto Médio (MAE) e Erro Quadrado Médio Médio (RMSE). Os menores valores de MAE e RMSE indicam um menor erro entre os preços reais e os previstos. A construção do modelo, permitiu a implementação de um sistema de suporte à decisão baseado em GIS para analisar o valor do hotel e estimar tarifas objetivas. Para futuras pesquisas os autores destacam a identificação de problemas que ainda não foram cobertos pelo protótipo atual.

Bala (2012) aplica a mineração de dados, para entender o comportamento sazonal do ramo varejista, influenciado pela vinda de turistas para o oeste indiano, em períodos específicos do ano. A proposta do estudo é baseada na aplicação de ferramentas para a previsão de demanda de unidades de manutenção de estoque que é ligada à logística de armazém e designa os diferentes itens do estoque para supermercado. O modelo proposto é baseado no método de agrupamento e ARIMA⁸, e o desempenho foi comparado em dois conjuntos de dados, um de treinamento de 36 meses e outro de teste de 12 meses, onde foi aplicado o modelo de seleção de recursos usando a razão de Pearson com valor-p de 0,05 para encontrar fatores importantes. Como resultado os modelos de previsão semanal, agrupamento baseado em ARIMA, descrito como WM7 e o agrupamento sazonal, chamados de WM8, superaram suas contrapartes em previsão diária sobre o agrupamento do conjunto de dados, chamado de DM7 e o agrupamento sazonal, DM8. No entanto, todos os quatro modelos superaram a prática existente. Com relação ao nível de estoque dado pelos dias de alcance e falha de vendas, o WM8 é ligeiramente melhor do que o WM7. O modelo utilizado no estudo é apropriado para itens para os quais uma proporção substancial de venda é fortemente atribuída a um perfil ou perfis específicos. Para futuras pesquisas o autor frisa a importância de empregar o agrupamento dos registros de vendas para a previsão de demanda.

Abbasi et al. (2020) apresentam a capacidade de simular e prever uma cadeia de suprimentos de sangue especificadamente do grupo sanguíneo O-positivo e propor a tomada de decisão sobre o abastecimento de quatro hospitais australianos ou a reposição do estoque da central de atendimento

⁸ É um modelo empregado com frequência para realizar previsões, a prática possibilita realizar a decomposição de séries temporais, suavização exponencial, regressão de séries temporais, modelos de autorregressivos e de média móvel integrada.

ao cliente. Os dados analisados são provenientes de 44 atributos de entrada e 136 atributos de saída para 18.500 observações ou instâncias diárias. Para a execução do gerenciamento das operações, foi proposto um programa linear estocástico para medir os dados de entrada do sistema. Os métodos de aprendizado de máquina para a arquitetura dos resultados são Artificial Neural Network (ANN), k-Nearest Neighbor (k-NN), Classification and Regression Tree (CART) e Random Forest (RF), que são utilizados para prever soluções de otimização com base no resultado esperado. Na mensuração dos resultados empregaram o MSE que demonstra a frequência em média, para cada conjunto de dados analisados. Como resultado, cada modelo de aprendizado de máquina aplicado nos dados, gerou diferentes parâmetros para cada variável aplicada, como quantidades de pedido e atendimento, custo do estoque, tempo que as unidades constam no estoque, então os autores deixam a critério dos itens que estão mais sendo avaliados para a tomada de decisão e direcionam o melhor método a ser utilizado. Para pesquisas futuras sugerem abordar o problema do ponto de vista meta-heurístico de otimização baseada na evolução e usar as decisões obtidas a partir de modelos de aprendizado de máquina como soluções iniciais de meta-heurísticas ou algoritmos matemáticos.

Bevilacqua et al. (2019) apresenta no estudo a utilização de mineração de dados atrelada a ferramentas de organização para reestruturação de almoxarifado de uma empresa de sapatos. O foco principal da utilização de mineração de dados é para reduzir o número de vezes que o almoxarife se desloca até o estoque para pegar algum item. O processamento da informação será realizado através de gerar atributos, selecioná-los, transformá-los em binominal para a organização no formato correto, e pôr fim a aplicação FP-Growth e regras de associações. O software escolhido foi o RapidMiner, é aberto e será utilizado para avaliação da base de dados e aplicação da mineração de dados. Foram tratados todo os espaços disponíveis nas prateleiras do almoxarifado, que são preenchidas com 777 categorias de itens e possuem 21 atributos para cada um, são registrados o código de identificação, a indicação do estoque máximo e a área ocupada. A frequência de ida do almoxarife ao estoque é mensurada em uma tabela de Excel e correlacionada com a metragem ocupada por esse item no espaço de armazenamento. O método proposto foi validado pelo autor que levanta o questionamento para futuras pesquisas baseadas em proporcionar a aplicabilidade do método a demais ramos industriais.

3 METODOLOGIA DA PESQUISA

A pesquisa é definida como a possibilidade de entender questões presentes no dia a dia das pessoas, que influenciam aspectos naturais e sociais. A análise e a compreensão dos fatos derivam da ciência, tornando possível explicar os fenômenos, desafiando crenças (Júnior & Batista, 2023).

A metodologia da pesquisa é norteadora de um trabalho científico. Nela, é possível definir métodos, regras e instrumentos que serão utilizados durante a construção do estudo. As técnicas apresentadas devem garantir a confiabilidade e qualidade das informações expostas. A metodologia deve ser construída sobre uma ótica crítica que garanta a replicação e compreensão do conhecimento (Kruger, 2023).

Uma estrutura organizada que assegura o entendimento das etapas estabelecidas na pesquisa é definida como metodologia da pesquisa. Além disso, ela não permite a influência de fatores externos na modelagem do conhecimento e possibilita a reprodutibilidade dos estudos demonstrados (Lima & Mito, 2007).

3.1 ABORDAGEM, OBJETIVOS, PROCEDIMENTOS E TÉCNICAS DE PESQUISA

A abordagem da pesquisa é quantitativa, sendo descrita quanto à natureza, aos objetivos e aos procedimentos adotados na elaboração da dissertação. Essa abordagem é construída sobre uma base numérica, utilizando técnicas de tratamentos de dados, como, por exemplo, técnicas estatísticas, para entender o comportamento da informação. O principal foco na condução de uma pesquisa quantitativa é a extração do conhecimento em cada etapa da análise da base de dados trabalhada. Essa análise deve ser elaborada com base em uma amostra ou população concreta e de qualidade (Fávero & Belfiore, 2024). A natureza da pesquisa caracteriza como é a tratativa do conhecimento extraído do estudo, sendo que nesta dissertação tem-se uma pesquisa aplicada, definida por Fleury & Werlang (2017) como o principal meio de descobertas de soluções e aplicação no dia a dia profissional ou acadêmico, pois está diretamente relacionada a geração de diagnósticos e a identificação de problema.

A pesquisa é aplicada e de base tecnológica, assim resulta no conhecimento tecnológico gerado na aplicação em um produto, serviço ou processo, tendo-se a participação ativa das mudanças que estão acontecendo em função da transformação digital e utiliza as inovações, por meio da pesquisa científica, para alavancar os resultados das organizações (Santos & Hashimoto, 2003).

No que se refere aos objetivos é uma pesquisa descritiva e exploratória. A pesquisa descritiva é definida como a busca pela descrição e entendimento do processo que está sendo estudado,

utilizando-se de métodos de tratamento de informações presentes na análise para a extração de possíveis resultados (Ó. Sousa & Rodrigues, 2022). A pesquisa exploratória norteia o tema de estudo, pois visa a descoberta e o entendimento de fenômenos, aplicando-se ferramentas de exploração de conteúdo para a fundamentação da pesquisa que está sendo construída (Jung, 2004). Também, é utilizada para compreender as características presentes em publicações científicas anteriores, que comprovam a utilização de métodos para possíveis soluções e apresenta explicações fundamentadas (Révillion, 2003).

A caracterização da pesquisa em relação aos procedimentos refere-se aos métodos utilizados para a realização da pesquisa. Na presente pesquisa aplica-se a revisão narrativa, o estudo bibliométrico e o estudo de caso.

A revisão bibliográfica é o primeiro contato entre o pesquisador e o tema, tornando-se obrigatória em pesquisa de abordagem exploratória, pois apresenta os conceitos, os métodos e as metodologias mais utilizados pelos autores e permite o mapeamento de fundamentações teóricas ao tema em pauta (A. S. de Sousa et al., 2021).

As pesquisas na área foram analisadas por meio de um estudo bibliométrico no intuito de visualizar publicações sobre o tema definido da presente pesquisa e obedecendo aos critérios estabelecidos (Gaviria-Marin et al., 2019). Como resultado foram encontradas publicações que nortearam o estudo.

O estudo bibliométrico é realizado para o pesquisador entender como o campo da ciência referente a temática estudada está se comportando, utilizando análises e entendimento de fatores que compõe a pesquisa, como o mapeamento de publicações por indivíduo, instituições e países. Os resultados deste tipo de estudo são apresentados de forma qualitativa e quantitativa, o que auxilia no seu entendimento e no norteamento de novos estudos (Doğan & Karaca, 2020).

Com os resultados extraídos do estudo bibliométrico, elaborou-se uma revisão de literatura narrativa do estado da arte, construindo-se uma síntese das pesquisas desenvolvidas a fim de se identificar os problemas de pesquisas estudados e as metodologias utilizadas para a resolução dos problemas (Cordeiro et al., 2007). As pesquisas apresentadas no estado da arte são aplicadas e testadas em situações reais, a fim de aprimorar atividades rotineiras, como é o caso da presente pesquisa, que de acordo com Jung (2004) utiliza o método de estudo de caso para a aplicação das ferramentas desenvolvidas em um sistema de produção, de forma individual ou coletiva.

O procedimento de estudo de caso, expõe a dinâmica atual do processo estudado e descreve a aplicação de métodos junto a mensuração dos resultados, para efetivar a proposta de pesquisa adotada (Ventura, 2007). Na presente pesquisa o estudo de caso é realizado em um ambiente industrial, no sul do estado de Santa Catarina.

3.2 DELIMITAÇÃO DO ESTUDO

No desenvolvimento desta pesquisa, buscou-se o entendimento das atividades relacionadas ao setor produtivo, objeto do presente estudo, do local de realização e da população, no caso, o conjunto de dados empregado para a aplicação das metodologias adotadas. As etapas que compõem a pesquisa são descritas, as quais foram estruturadas por meio de ferramentas e de métodos a fim de permitir a conclusão de cada uma das fases de desenvolvimento.

O estudo delimita-se na integração entre os meios tecnológicos e a tomada de decisão empresarial, voltados ao fortalecimento da sinergia dos principais temas propostos, sendo eles a indústria e a inteligência artificial.

3.2.1 Descrição do local e população em estudo

Este estudo foi realizado na região sul do estado de Santa Catarina, com foco no setor produtivo de extração de carvão mineral. Utilizou-se dados reais, presentes no banco de dados de uma empresa que compõe o núcleo de carboníferas que praticam a lavra de subsolo na região.

O estudo de caso foi aplicado em uma indústria carbonífera de grande porte, sendo a primeira de Santa Catarina a obter as três principais certificações, de gestão da qualidade, de gestão ambiental e de gestão de saúde e segurança ocupacional. A empresa possui um programa de incentivo à implantação de tecnologia e inovação, proporcionando capacitação da mão de obra aos funcionários e atualização das comunidades locais (Mineral, 2021).

O produto desta indústria consiste no beneficiamento do carvão mineral, sendo o seu sistema produtivo⁹ subdividido em setores de transformação mineral e de apoio produtivo. A presente pesquisa relaciona-se com o sistema de apoio produtivo, responsável por manter os equipamentos de lavra e de usina em funcionamento, visando a maior recuperação mineral, conforme estabelecido no projeto de construção da usina.

O setor de manutenção, onde foi aplicada a presente pesquisa, é dividido em cinco categorias, cada uma responsável por atender a diferentes tipos de falha. Todas as peças utilizadas na manutenção de equipamentos e componentes provêm do almoxarifado, disponível na empresa. As peças requisitadas pelos mecânicos, têm as saídas do estoque via sistema, ficando registradas as informações como a data de requisição, o setor responsável e o equipamento ou os componentes em que a peça será utilizada.

⁹ O sistema produtivo é um conjunto de processo e operações que interagem entre si para produzir serviço ou produto, processo contínuo no qual pode ser chamado de produção (Hopp & Spearman, 2012).

O Setor de Tecnologia da Informação da organização disponibilizou o conjunto de dados relacionado aos suprimentos de inventário de manutenção, no período entre os anos de 2019 e 2022, dados estes que foram provenientes do Sistema de Gestão de Tarefas da empresa.

Na amostra encaminhada existiam 14.818 peças ativas no inventário da organização. No entanto, as peças não se limitavam à manutenção; também incluía peças utilizadas durante o processo de lavra, como dinamites, óleo diesel, entre outros.

As etapas de limpeza, tratamento, transformação e seleção do conjunto de dados estão descritas na seção 3.4.1, Pré-processamento de dados, em que se apresentam os métodos e as ferramentas adotados nesta pesquisa, que utilizou um total de 172.691 registros de circulação dessas 14.818 peças.

3.2.2 Etapas da pesquisa

Este estudo investigou a aplicação de ferramentas que incorporam a ciência de dados no processo de tomada de decisão para a gestão de inventário de peças de manutenção. As etapas adotadas no desenvolvimento da pesquisa são apresentadas no quadro 2.

A pesquisa foi desenvolvida com base nos resultados obtidos no estudo bibliométrico e no estado da arte. A análise dos materiais desempenhou um papel fundamental para a compreensão e delimitação dos temas e das ferramentas propostas no trabalho.

Com base nas etapas apresentadas no quadro 2 e aprofundando-se o entendimento sobre o tema, iniciou-se a avaliação das informações disponíveis. Nesse contexto, realizou-se um estudo detalhado sobre o processo de requisição de peças no almoxarifado de manutenção.

O estudo descreve o processo atual da gestão de inventário do setor estudado. Para isso, foi realizado o mapeamento do processo interno e sua apresentação por meio de um fluxograma, figura 17. O fluxograma, conforme Tabile et al. (2015) utiliza gráficos para ilustrar os pontos de entrada e de saída do processo. Posteriormente, os dados relacionados à gestão de inventário no setor de manutenção foram analisados por meio da ciência de dados, utilizando-se de um algoritmo de aprendizado de máquina e métricas de qualidade, a fim de se identificar o modelo com os melhores resultados para apoio à tomada de decisão.

Com base no processo exposto na figura 17 foi possível definir quais conjuntos de dados o Setor de Tecnologia da Informação da organização, encaminharia para o andamento da presente pesquisa. A amostra enviada foi dividida em partes, iniciando-se o processo de pré-processamento.

Durante o pré-processamento de dados, aplicaram-se métodos de limpeza de dados para a remover os dados obsoletos ou irrelevantes, bem como dados que poderiam tendenciar a

interpretações errôneas. Além disso, utilizou-se a técnica de imputação de valores para fortalecer a confiabilidade dos dados disponíveis.

Após a etapa de pré-processamento, iniciou-se a mineração de dados. Antes de aplicar o algoritmo escolhido, foram pré-selecionadas amostras de dados, com base na confiabilidade da informação.

Quadro 2.

Apresentação das etapas da pesquisa.

APLICAÇÃO DA METODOLOGIA DE BUSCA DE CONHECIMENTO EM BASE DE DADOS	Bibliometria	1. Bases de dados Web of Science Scopus Science Direct IEEE Xplore
		2. Palavras chaves Inventory management AND "Data Mining"
		3. Período analisado 2012 a 2022
	Estado da arte	1. Os 5 trabalhos mais citados (Lee et al., 2015) (Balali et al., 2015) (Kisilevich et al., 2013) (Bala, 2012) (Abbasi et al., 2020)
		2. estudo com forte aderência ao tema (Bevilacqua et al., 2019)
DESCRIÇÃO E ENTENDIMENTO DO PROCESSO DE INVENTÁRIO DE MANUTENÇÃO	Descrição do processo avaliado	1. Mapeamento do processo de GI
		2. Construção do fluxograma de GI
		3. Identificação do conjuntos de dados Banco de requisição Banco de Ordens de Serviços Extração do SQL
Acesso ao sistema integrado da empresa		
PRÉ-PROCESSAMENTO DOS DADOS	Exportação dos dados	1. Banco de requisição 18 Atributos 197.959 Registros
		2. Banco Ordens de Serviços 41 Atributos 55.381 Registros
		3. Extração do SQL 2 Atributos 123.169 Registros
	Estruturação dos dados	1 Software Power Query Excel
		2. Métodos Eliminação de valores externos Padronização dos dados Método de imputação Tratamento de classes desbalanceadas Correção de valores ausentes Aplicação de One-Hot Encoding
		3. Estrutura 25 Atributos 50.162 Registros
MINERAÇÃO DOS DADOS	Método	1. Associação Algoritmo <i>Apriori</i>
		1.1 Suporte 40%
		1.2 Confiança 80%
PÓS PROCESSAMENTO DE DADOS	Análise dos resultados por meio da métrica <i>Lift</i>	
	Aplicação da análise de valor	

Nota. Autores (2023).

A mineração de dados é utilizada para interpretar padrões e realizar a descoberta de conhecimento em bases de dados, buscando por padrões e identificando as potencialidades de apoio à tomada de decisão. Nesta pesquisa, no processo de mineração optou-se em utilizar o algoritmo Apriori que é um dos mais empregados na tarefa de associação (V. Kumar et al., 2009), a fim de se relacionar a amostra construída, obtendo-se as regras de associação entre as peças requisitadas no almoxarifado da unidade de manutenção.

Por fim, no pós-processamento empregou-se a métrica Lift para avaliação da qualidade dos modelos gerados pelo algoritmo de associação Apriori, bem como realizou-se a análise de valor sobre os resultados obtidos. A análise de valor, segundo Kassa (2015), é uma ferramenta amplamente utilizada na gestão econômica empresarial e tem como foco entender o comportamento de tendências e permitir a tomada de decisão sobre números concretos.

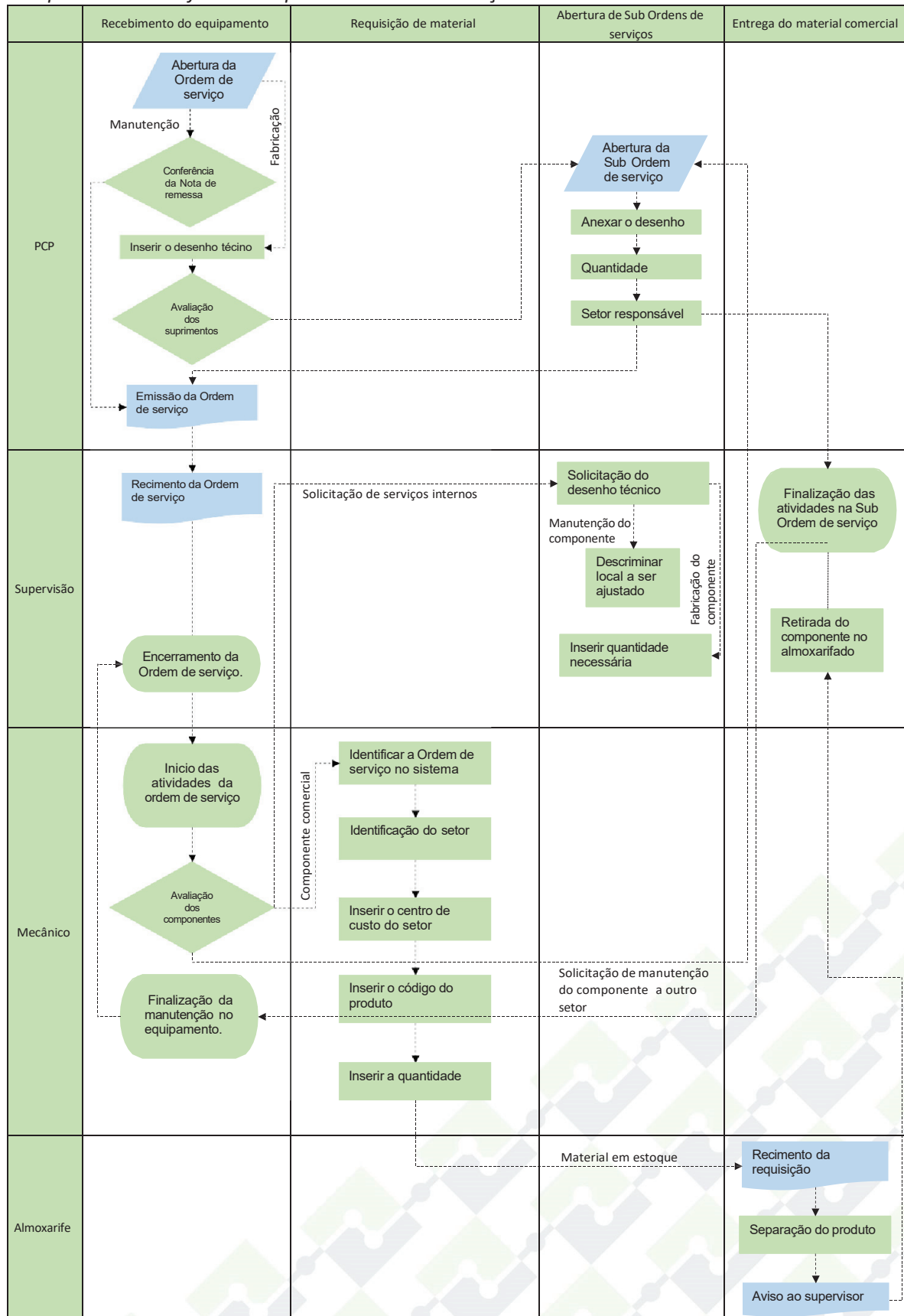
3.3 PROCESSO DE COLETA DE DADOS

A base de dados empregada nesta pesquisa não é de domínio público, assim os dados necessários para a realização do trabalho, foram disponibilizados pelo Setor de Tecnologia da Informação de uma indústria carbonífera. A exposição dos dados na pesquisa será realizada considerando os princípios das diretrizes impostas pela empresa.

A coleta de dados foi realizada no sistema de gerenciamento de tarefas, relacionando-se ao inventário do setor de manutenção da organização, que disponibilizou as bases de dados referente as requisições e as ordens de serviço, as quais foram trabalhadas nesta pesquisa a fim de se realizar a organização e pré-processamento dos dados. O conjunto de dados foi disponibilizado no formato CSV, e foram manuseados no Power Query e Excel.

A saída de peça do estoque de manutenção é realizada via sistema, conforme apresentado no fluxo da figura 17, o processo registra as informações, como data de saída, setor responsável, quantidade de material, o setor que irá utilizá-la, entre outras.

Figura 17.
 Comportamento do fluxo de suprimentos de manutenção.



Nota. Autores (2023).

Na realização do processo de manutenção é necessária a emissão da Ordem de Serviço (OS), realizada pelo Setor de Planejamento e Controle de Produção. A OS possui as informações sobre o componente que foi encaminhado para a manutenção, como o número de patrimônio do equipamento e componente, a unidade que encaminhou, a data de encaminhamento e todos os registros fiscais, classificando em manutenção ou fabricação de componentes novos.

As atividades de manutenção são iniciadas somente com a OS em mãos. Durante o processo, os mecânicos realizam a avaliação das peças dos componentes, caso não esteja viável sua reutilização, é realizada a requisição de uma peça nova para a substituição, preenchida pelos mecânicos digitalmente no sistema. O preenchimento digital informa o código da peça necessária, a quantidade requisitada e o número da OS em que a peça será utilizada.

O fluxo de informações gera no sistema uma população de dados referente ao inventário de manutenção e ao histórico de ordens de serviço. Os dados extraídos do sistema integrado da empresa, contém atributos chaves, como o código identificador do componente e atributos compostos, como a descrição do componente. Além disso, é possível visualizar atributos simples como, histórico de circulação do produto na empresa e o setor responsável em solicitar o material.

Para a construção da base de dados foram disponibilizados relatórios específicos, como os registros de requisição de componentes e de ordens de serviço. Para vincular o número de ordem de serviço ao número de requisição foi exportado via *Structured Query Language* (SQL) a mescla das informações do banco de dados do sistema integrado, obtendo-se os conjuntos apresentados na tabela 07.

Tabela 7.

Estrutura dos relatórios extraído do sistema de gestão da empresa.

	REGISTROS DE REQUISIÇÃO DE COMPONENTES	REGISTROS DE ORDENS DE SERVIÇO	REQUISIÇÃO POR ORDEM DE SERVIÇO
REGISTROS	197.959	55.381	123.169
ATRIBUTOS	18	39	2

Nota. Autores (2023).

A tabela 07 apresenta a população total extraída da base de dados, a fim de se iniciar o processo de análise dados que permitiu a descoberta de conhecimentos nos registros analisados.

3.4 PROCESSO DE ANÁLISE DE DADOS

O processo de análise de dados compreendeu as etapas da pesquisa referente ao pré-processamento dos dados, a mineração de dados e ao pós-processamentos dos dados. Assim, a

análise de dados foi realizada por meio de uma combinação de algoritmo de inteligência artificial e análises quantitativas.

A análise de dados por meio da inteligência artificial possibilita a identificação de relações não definidas *Apriori* e a previsão de demandas futuras, auxiliando no entendimento dos itens e da quantidade que devem ser comprados e armazenados, no gerenciamento de incertezas para evitar erros na tomada de decisão (Núñez-Merino et al., 2022). A leitura sob demanda para entender as tendências futuras, realizada pelo aprendizado de máquina, correlaciona eventos, como por exemplo, os naturais como clima ou pandemias, com fatores econômicos que influenciam diretamente o desempenho da organização (Omar et al., 2020).

A etapa de definição dos métodos de mineração de dados utilizados, fundamenta o objetivo específico, de identificação dos métodos de ciência de dados a serem empregados para a descoberta de conhecimento na base de dados de inventário de manutenção, para a delimitação dos dados realizou-se o pré-processamento da população utilizando, por exemplo, o método de padronização de dados. Este método consiste em transformar dois ou mais objetos similares em um único, reduzindo a população de dados, que ainda assim poderá gerar o mesmo resultado ou resultados mais precisos (Kumar et al., 2009).

Assim, o tamanho da amostra foi definido pelo número de itens resultantes do pré-processamento da base de dados. De acordo com Mariano et al. (2020) uma delimitação muito grande de itens para a execução da mineração de dados pode resultar em falsos resultados ou na sua indução.

3.4.1 Pré-processamento dos dados

O pré-processamento dos dados compreendeu várias atividades que são aplicadas nos dados para que eles se tornem adequados para a execução do processo de mineração de dados, a fim de se corrigir erros e dados inconsistentes que possam interferir na descoberta de conhecimento (Kumar et al., 2009).

O início do processo foi realizado na estrutura da base de dados, definido por alguns autores como D'Castro (2020), a etapa de estruturação e tratamento que subsidia de modo confiável a aplicação dos algoritmos de descoberta de conhecimento, pois é realizada utilizando-se dos princípios da análise de dados que será executada (tarefa e algoritmo).

Na construção da base de dados empregada nesta pesquisa, realizou-se o tratamento dos dados (tabela 07). O primeiro conjunto de dados trabalhado foi o referente a requisições de peças (figura 18), em que são apresentadas todas as peças requisitadas no almoxarifado.

Figura 18.
Tratamento da base de dados de requisições de materiais.

Total de Atributos 18	1 - Remoção dos atributos vazios 16	2 - Remoção dos atributos redundantes 11
Nr. Requisição >	Nr. Requisição >	Nr. Requisição
Empresa >	Empresa	Data
Data >	Data >	Cód. Requisiteiro
Cód. Requisiteiro >	Cód. Requisiteiro >	Material (Codigo)
Situação	Material (Codigo) >	Descricao do Material
Material (Codigo) >	Descricao do Material >	UN
Descricao do Material >	UN >	Qtd. Material
UN >	Qtd. Material >	Qtd. Devolucao
Qtd. Material >	Qtd. Devolucao >	Centro de Custo (Código)
Qtd. Devolucao >	Centro de Custo (Código) >	Item
Centro de Custo (Código) >	Nome Centro de Custo	CD_EQUIPAMENTO_EMPRESA
Nome Centro de Custo >	# Requisição Material	
# Requisição Material >	Item >	
Cód. Empresa	CD_EQUIPAMENTO_EMPRESA >	
Item	CD_EQUIPAMENTO	
CD_EQUIPAMENTO_EMPRESA >	DS_EQUIPAMENTO	
CD_EQUIPAMENTO >		
DS_EQUIPAMENTO >		

Nota. Autores (2023).

O banco possuía em sua estrutura 18 atributos, descritos na coluna Total de atributos. Para aprimorar a extração do conhecimento e entender o comportamento das requisições foram extraídos os registros de ordens de serviço (figura 19).

Figura 19.

Tratamento da base de dados de Ordens de serviço.

Total de Atributos 39	1 - Remoção dos atributos vazios 29	2 - Remoção dos atributos redundantes 24
Código >	Código >	Código
Nr. OS >	Nr. OS >	Nr. OS
Emissão >	Emissão >	Emissão
Nome da Peça >	Nome da Peça >	Nome da Peça
Quantidade >	Quantidade >	Quantidade
Nr. Desenho >	Nr. Desenho >	Nr. Desenho
Data Requerida Entrega >	Data Requerida Entrega >	Data Requerida Entrega
Data Previsão Entrega >	Data Previsão Entrega >	Data Previsão Entrega
Data Devolução	Nr. Empresa >	Nr. Empresa
Faturamento	Nome da Empresa	Nr. Nota Remessa
Nr. Empresa >	Nr. Nota Remessa >	Emissão Remessa
Nome da Empresa >	Emissão Remessa >	Nr. Nota Retorno
Nr. Nota Remessa >	Nr. Nota Retorno >	Emissão Retorno
Emissão Remessa >	Emissão Retorno >	Solicitante (Código)
Nr. Nota Retorno >	Solicitante (Código) >	Supervisor (Código)
Emissão Retorno >	Supervisor (Código) >	Nota Saída Remessa (Código)
Nr. Nota Faturamento	Nota Saída Remessa (Código) >	Encerramento Operacional
Emissão Faturamento	Encerramento Operacional >	Encerramento Custos
Solicitante (Código) >	Encerramento Custos >	Nota Saída Retorno (Código)
Supervisor (Código) >	Nota Saída Retorno (Código) >	Cód. Componente
Nota Saída Faturamento (Código)	Cód. Componente >	Cód. Equipamento
Nota Saída Remessa (Código) >	Componente	Cód. Centro Custo
Encerramento Operacional >	Cód. Equipamento >	Oper. Fiscal
Encerramento Custos >	Equipamento	CD_USUARIO_INCLUSAO
Nota Saída Retorno (Código) >	CD_EMPRESA	
Cód. Componente >	Cód. Centro Custo >	
Componente >	Centro Custo	
Cód. Equipamento >	Oper. Fiscal >	
Equipamento >	CD_USUARIO_INCLUSAO >	
CD_EMPRESA >		
DS_OBSERVACAO_INTERNA		
DS_OBSERVACAO_EXTERNA		
Cód. Centro Custo >		
Centro Custo >		
Oper. Fiscal >		
FG_CONF_MATERIAL		
DS_MOTIVO_BAIXA		
CD_USUARIO_INCLUSAO >		
FG_SITUACAO		

Nota. Autores (2023).

Todos os materiais requisitados, devem por obrigatoriedade estar vinculados em uma ordem de serviço, pois o documento é uma identificação do equipamento encaminhado para manutenção. Dentro da ordem de serviço estão cadastrados todos os itens mencionados na coluna total de atributos da figura 18.

Para tratar as diferentes bases de dados, usou-se a mesma lógica, onde iniciamos pela exclusão de atributos nulos contidos nos dois relatórios, na figura 18 e 19, é possível identificar os atributos que permaneceram, os atributos desconsiderados estão em negrito. O segundo critério foi

excluir atributos redundantes, os registros possuíam atributo “Centro de custo (Código)”, seguido pelo atributo “Nome do centro de custo”. Para padrão de tratamento de informações, optou-se em trabalhar apenas com códigos do sistema e utilizar o dicionário de dados para interpretações.

Com os dois conjuntos de dados tratados, utilizou-se o software Power Query para realizar a mesclagem de informações e obter a estrutura da base de dados demonstrada na figura 20. O conjunto construído possui 29 atributos e 172.691 registros e tem como objetivo extrair via ciência de dados todo o conhecimento presente no conjunto de dados afim de estudar os modelos gerados pela ciência de dados aplicados à transformação digital e apoio à tomada de decisão do inventário de manutenção.

Figura 20.

Estrutura da base de dados.

Código
Nr. OS
Emissão
Quantidade
Nr. Desenho
Data Requerida Entrega
Data Previsão Entrega
Nr. Empresa
Nr. Nota Remessa
Emissão Remessa
Nr. Nota Retorno
Emissão Retorno
Solicitante (Código)
Supervisor (Código)
Encerramento Operacional
Encerramento Custos
Cód. Componente
Cód. Equipamento
Cód. Centro Custo
Nr. Requisição
Requisições.Data
Requisições_Cod_Responsáveis
Requisições.Material (Codigo)
Requisições.UN
Requisições.Qtd. Material
Requisições.Qtd. Devolucao
Requisições.Centro de Custo (Código)
Requisições.Item
Requisições.CD_EQUIPAMENTO_EMPRESA

Nota. Autores (2023).

Com a conclusão da estruturação da base de dados, identificou-se uma estrutura baseada em atributos que são responsáveis em descrever as características das entidades que estão sendo apresentadas. Os atributos estão subdivididos em simples, compostos, chaves (não únicos) e multivalorados, conforme apresentado no Quadro 3.

A tipologia dos atributos representa como podem ser tratados e interpretados, possuindo 18 atributos vinculados a tipologia simples, pois não possuem características especiais em seus registros e não é possível desdobrar esses registros em mais de um atributo, por se tratar de códigos e número de notas fiscais.

No atributo composto é possível o desmembramento para encontrar um ou mais atributos derivados da operação. Assim, os atributos classificados no quadro 3 como compostos, podem ser desmembrados em mais atributos. Um cadastro recorrente são as datas, já que se observar o comportamento diário, mensal ou anual dos registros.

A base de dados possui o atributo multivalorado, representados por “OS.Código”, que caracteriza determinado componente, porém não se repete em outros registros. A cada recebimento de componentes para manutenção é gerado um código novo.

Quadro 3.

Representação dos atributos e suas tipologias.

TIPOLOGIA	ATRIBUTOS
SIMPLES	OS.Quantidade OS.Nr. Empresa OS.Nr. Nota Remessa OS.Nr. Nota Retorno OS.Solicitante (Código) OS.Supervisor (Código) OS.Cód. Componente OS.Cód. Equipamento OS.Cód. Centro Custo Exportação.Nr. Requisição Tabela_Código_Supervisor Exportação.Material (Codigo) Exportação.UN Exportação.Qtd. Material Exportação.Qtd. Devolucao Exportação.Centro de Custo (Código) Exportação.Item Exportação.CD_EQUIPAMENTO_EMPRESA
COMPOSTOS	OS.Nr.OS OS.Emissão OS.Data Requerida Entrega OS.Data Previsão Entrega OS.Emissão Remessa OS.Emissão Retorno OS.Encerramento Operacional OS.Encerramento Custos Exportação.Data OS.Nr. Desenho
MULTIVALORADOS	OS.Código

Nota. Autores (2023).

Considerando-se que a base de dados construída pode estar contaminada por itens duplicados, componentes cadastrados que não são utilizados para manutenção, entre outros fatores, realizou-se o pré-processamento dos dados. O pré-processamento compreendeu várias atividades, dentre as quais a limpeza dos dados que consiste na remoção dos registros duplicados e não confiáveis que podem comprometer a qualidade do resultado da mineração de dados.

Durante as etapas de pré-processamento de dados, está a verificação da confiabilidade dos dados que estão sendo empregados. A confiabilidade é a capacidade de não permitir que exista dados

incoerentes no banco, de forma que resulte em uma distorção das análises realizadas. A confiabilidade é definida como a capacidade de não variar ou influenciar de modo errôneo a interpretação dos dados (Veras & Martins, 1994).

A análise de valor no campo empresarial pode ser limitada, pela baixa capacidade de processamento e armazenamento de informações referentes a atividades realizadas periodicamente. As organizações estão em processos de digitalização de algumas atividades, deste modo antes da aplicação de métodos de extração de conhecimento sobre bases de dados, deve ser averiguada a sua confiabilidade (Queiroz & Pereira, 2019).

A aplicação da mineração em dados com baixa confiabilidade direciona os algoritmos para a produção de resultados tendenciosos ou específicos, caracterizando uma distorção no conhecimento extraído (Ferreira et al., 2020).

A confiabilidade de registros é definida como a capacidade dos dados serem avaliados por pessoas distintas em diferentes momentos, sem variação por erros ou falhas de registros. Quando se aplicam as mesmas tarefas, algoritmos e parâmetros a uma base de dados confiável, ela apresenta o mesmo resultado (Veras & Martins, 1994). Assim, a confiabilidade está diretamente relacionada com a qualidade dos registros observados, pois quando se tem uma alta confiabilidade é possível realizar uma tomada de decisão segura (Hora et al., 2010).

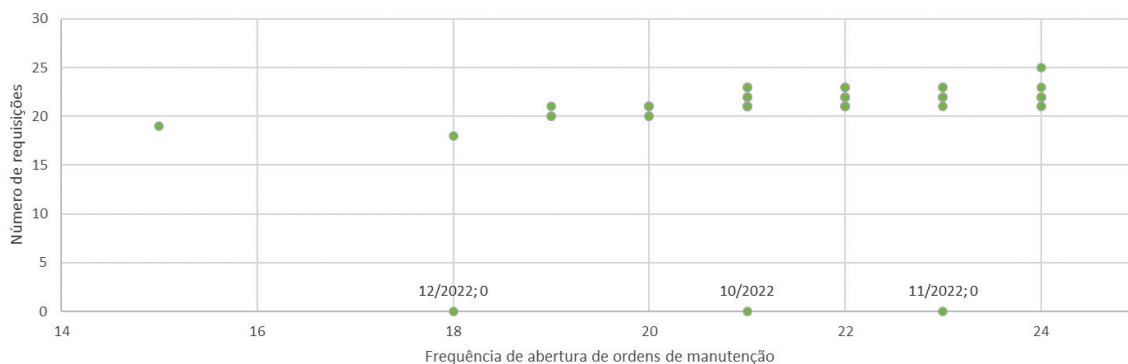
As técnicas de pré-processamento aplicadas na base de dados foram de remoção de valores extremos, tratamento de classes desbalanceadas, imputação de valores, eliminação de valores incoerentes e classificação via *one-hot encoding*.

A remoção de valores extremos consiste na aplicação de ferramentas estatísticas para realizar a leitura da base de dados e identificar conjuntos fora da curva que podem influenciar o resultado ou ocasionar campos vazios (Alves, 2003). No reconhecimento dos valores extremos empregou-se o gráfico de dispersão sobre todos os registros de Ordens de serviço emitidas e das requisições de componentes. De acordo com Lins (1993), o gráfico de dispersão permite visualizar grandezas com o comportamento atípico ao longo da uma reta, quando os registros ficam muito isolados dos demais são considerados atípicos e, portanto, desconsiderados para o processo de mineração de dados.

A remoção de valores extremos foi aplicada ao atributo período, abrangendo todos os registros cadastrados no sistema entre 2019 e 2022. No entanto, alguns relatórios foram enviados para análise sem os dados de outubro, novembro e dezembro de 2022, resultando na exclusão desses meses da análise.

Figura 21.

Gráfico de dispersão do conjunto de dados.



Nota. Autores 2023.

Na figura 21 é possível ver três cadastros com valores 0, os quais se referem aos meses de outubro, novembro e dezembro de 2022 para registros de requisições de materiais. Para evitar o campo nulo na base de dados, foram desconsiderados os registros de cadastramento de ordem de serviço no período. A análise da informação tratada deve ser extraída de dados com alta confiabilidade, sendo ela definida como fator principal na alteração da forma de tomada de decisão, pois os dados derivam conhecimento, porém devem ser embasado sobre registros confiáveis, portando a qualidade do material deve sempre ser conferida (Costa et al., 2022).

A imputação de valores tem como objetivo substituir valores incorretos ou desconhecidos por valores verdadeiros, a técnica pode ser utilizada quando há mais de um registro para um único item e um deles está incompleto, desta forma deve haver a imputação da descrição correta, substituindo os valores incoerentes (Alves, 2003).

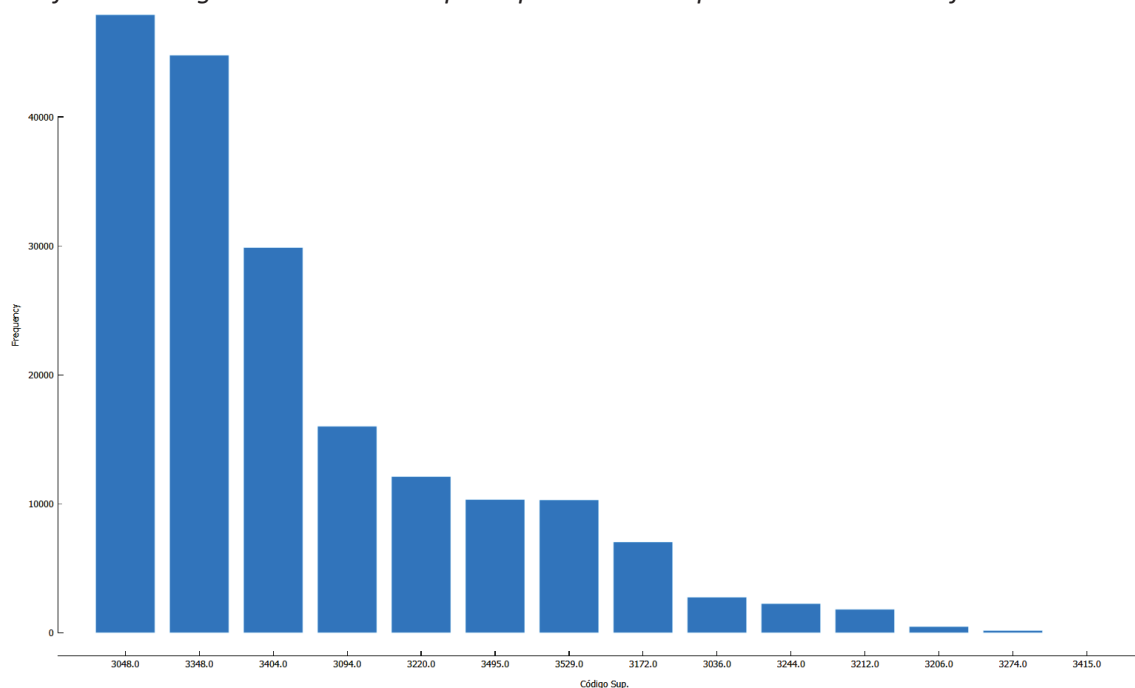
A técnica de imputação de dados, pode ser utilizada de diversas formas, tendo-se a substituição de casos, que busca substituir valores nulos ou incoerentes nos registros por dados reais e com qualidade, a imputação via média ou moda é utilizada para completar registros nulos por meio do levantamento de mediana ou moda dos valores anteriores a imputação de novos valores (Araújo et al., 2015) e a imputação via conhecimento de domínio que ocorre por meio do conhecimento empírico do cientista sobre a base de dados trabalhada, sendo geralmente realizada por um analista que possui familiaridade com os atributos envolvidos (Alves, 2003).

A etapa de pré-processamento nesta pesquisa, seguiu com a aplicação do método de imputação dos dados, analisando-se o atributo "Requisições_Cod_Responsáveis", que possui os registros de todos os códigos responsáveis pela requisição de peças no almoxarifado da organização. No entanto, como os registros são oriundos de um determinado período, entende-se que há alterações de códigos responsáveis pela requisição de materiais, logo o sistema captura o novo código

requisitante, porém alguns setores encontravam-se com mais de um código cadastrado, apresentando o comportamento exposto na figura 22.

Figura 22.

Gráfico de histograma de cadastros que requisitavam componentes no almoxarifado.



Nota. Autores (2023).

A figura 22 demonstra a disparidade sobre o comportamento de alguns valores na base de dados, onde o eixo x apresenta os códigos responsáveis e o eixo y a quantidade de peças solicitadas no almoxarifado da empresa. O setor de manutenção analisado possui 6 subáreas, desta forma a figura 22 deveria demonstrar apenas 6 colunas. Na figura também é possível identificar que existe mais de um código para cada setor, visto a incoerência aplicou-se a imputação de valores com maiores frequências, de códigos para setores, conforme demonstrado no quadro 4.

Quadro 4.

Imputação de valores na base de dados.

Cód. Cadastrado	%Cód. Cadastrado		Cód. Único		Cód. Único	%Cód. Único
3094	8,61%	>	3094			
3404	16,07%	>	3404			
3686	0,00%	>	3404			
3348	24,10%	>	3348			
3048	25,79%	>	3048	3094	10,90%	
3529	5,54%	>	3529	3404	16,33%	
3274	0,09%	>	3094	3348	24,10%	
3495	5,56%	>	3036	3048	25,79%	
3709	0,00%	>	3529	3529	12,06%	
3036	1,48%	>	3036	3036	10,82%	
3244	1,21%	>	3094			
3220	6,52%	>	3529			
3285	0,00%	>	3036			
3172	3,78%	>	3036			
3049	0,00%	>	3036			
3415	0,00%	>	3094			
3212	0,97%	>	3094			
3206	0,26%	>	3404			

Nota. Autores (2023).

No quadro 4 pode-se observar o agrupamento dos códigos de responsáveis de requisições de materiais, utilizando o critério de substituição por frequência, onde optou-se em considerar como fixo o código cadastrado mais frequente em cada setor, como resultados foram definidos 6 códigos, 3094, 3404, 3348, 3048, 3529 e 3036, em que cada um deles define um conjunto de manutenção.

Seguiu-se com a tarefa de imputação, porém com a aplicação sobre os cadastros das peças presentes no almoxarifado, ao todo, foram verificados os 13.442 cadastros, onde foram avaliadas as descrições de todos as peças presentes. A análise objetiva substituir ou excluir nomes genéricos, que não levariam a uma tomada de decisão eficiente, o interesse principal da pesquisa é entender o inventário de peças de manutenção e não prever itens obrigatórios aos mecânicos.

Quadro 5.

Famílias de itens que foram removidas da base de dados.

FAMÍLIA	REGISTROS
CHAVE	Todos os registros de chave de fendas e ferramentas para realizar a manutenção.
TINTA	Registros de tintas utilizadas para pintar os equipamentos finalizados.
COMPONENTES DE LIMPEZA DAS PEÇAS	Desengraxastes e óleo diesel, são utilizados para limpeza das peças antes da manutenção.

MADEIRAS

Itens de madeira que foram requisitados na Ordem de serviço também foram desconsiderados.

Nota. Autores (2023).

O quadro 5 apresenta os conjuntos de registros excluídos da base de dados, tais instâncias não somariam aos resultados da presente pesquisa, pois trata-se apenas de utensílios utilizados pelos mecânicos nos serviços de manutenção, ou na limpeza e entrega dos componentes e não influenciam diretamente no processo de manutenção. Em todo o processo não são reconhecidas peças mecânicas que necessitam de partes de madeira em sua montagem, logo foram desconsiderados todos os registros de madeira.

Durante o pré-processamento, foram identificados e substituídos itens redundantes (figura 23), que possuíam dois ou mais cadastros para o mesmo componente, e foram ajustados apenas para um único cadastro, transparecendo o método de imputação.

Figura 23.

Exemplos de imputação de dados

Exemplo 1 =	ANEL "O" RING 2019 PARKER (UN)	ANEL "O" RING 2019 PARKER
Exemplo 2 =	ANEL O'RING 2129 DUREZA 90 SHORE	ANEL O'RING 2129 DUREZA 90 SHORE

Nota. Autores (2022).

O componente ANEL "O" RING 2019 PARKER (UN) – Cód. 7728, possui a mesma descrição do item ANEL "O" RING 2019 PARKER – Cód. 12289, sendo que ambos são tratados como unidade, então foi identificado qual tem mais frequência de uso e substituído nos registros, a mesma lógica foi aplicada ao exemplo 2 e a outros itens da base de dados que apresentavam problemas similares.

As atividades de padronização retiraram 3.577 peças da base de dados, resultando em 9.565 peças que contabilizam 161.609 registros.

Seguindo o processo de imputação das informações presentes na base de dados, foram verificados os equipamentos e componentes registrados na Ordem de serviço. Entende-se que com a análise das manutenções realizadas sobre os equipamentos e componentes e quais as peças trocadas com mais frequência no processo de manutenção, pode-se levantar um possível cenário de análise de falha para entendimento dos problemas mecânicos apresentados.

Foram conferidos se todos os equipamentos estavam cadastrados, identificando-se que havia campos nulos, realizou-se a imputação de dados e os cadastramentos do código do equipamento em registros nulos. A identificação dos equipamento, se deu por conta da descrição do atributo "OS.Nome da Peça", que demonstra como padrão da seguinte forma "BLOCO DA VALVULA

INTELIGENTE ENROLADOR DE CABO Nº 01 – SC 01”, nesse descritivo, conseguiu-se identificar dois pontos, a primeira parte da nomenclatura “BLOCO DA VALVULA INTELIGENTE ENROLADOR DE CABO Nº 01” significa o nome do componente que está em manutenção, e que estava instalado no equipamento “SC 01”, descrito na segunda parte.

A imputação de dados foi necessária para os registros dos equipamentos e dos componentes, criando-se um dicionário de dados. Os componentes que não foram registrados no sistema foram considerados como não rotineiros e foram desconsiderados na base de dados.

O tratamento de classe desbalanceada, trata-se de um balanceamento efetivo entre os conjuntos de dados, a fim de assegurar que os atributos possuam números próximos de registros. Pode-se referir a esse tipo de método como a avaliação entre um domínio majoritário e um minoritário. Segundo Araújo et al. (2015), a maioria dos algoritmos de inteligência artificial possuem dificuldade em extrair informações de conjuntos desbalanceados.

A contagem de registros em cada um dos atributos foi realizada (quadro 6), encontrando-se atributos que não estão com todos os campos preenchidos.

Entende-se que os componentes que não foram registrados no sistema, não são componentes rotineiros de manutenção, desta forma, podem ser desconsiderados da pesquisa construída. Seguindo o fluxo da construção do banco de dados, foram desconsiderados os registros nulos do atributo “OS.Equipamento” e “OS.Componente”, tendo como resultado 58.504 registros.

Quadro 6.

Apresentação do percentual de preenchimento dos atributos.

ATRIBUTO	CONFIABILIDADE
OS.CÓDIGO	100,000%
OS.EMISSION	100,000%
OS.NOME DA PEÇA	100,000%
OS.QUANTIDADE	100,000%
OS.DATA REQUERIDA ENTREGA	100,000%
OS.NR. EMPRESA	100,000%
OS.SOLICITANTE (CÓDIGO)	100,000%
OS.CÓD. COMPONENTE	100,000%
OS.CÓD. EQUIPAMENTO	100,000%
OS.CÓD. CENTRO CUSTO	100,000%
OS.OPER. FISCAL	100,000%
OS.CD_USUARIO_INCLUSAO	100,000%
EXPORTAÇÃO.NR. REQUISIÇÃO	100,000%
EXPORTAÇÃO.DATA	100,000%
TABELA_CÓDIGO_SUPERVISOR9.NOME SUPERVISOR	100,000%
EXPORTAÇÃO.MATERIAL (CODIGO)	100,000%
EXPORTAÇÃO.DESCRICAO DO MATERIAL	100,000%
EXPORTAÇÃO.UN	100,000%
EXPORTAÇÃO.QTD. MATERIAL	100,000%
EXPORTAÇÃO.QTD. DEVOLUCAO	100,000%
EXPORTAÇÃO.CENTRO DE CUSTO (CÓDIGO)	100,000%
EXPORTAÇÃO.ITEM	100,000%
OS.DATA PREVISÃO ENTREGA	100,000%

OS.SUPERVISOR (CÓDIGO)	99,986%
OS.ENCERRAMENTO CUSTOS	99,978%
OS.NR. NOTA REMESSA	82,692%
OS.EMIÇÃO REMESSA	82,692%
OS.NOTA SAÍDA REMESSA (CÓDIGO)	82,692%
EXPORTAÇÃO.CD_EQUIPAMENTO_EMPRESA	59,208%
OS.NR. NOTA RETORNO	24,207%
OS.EMIÇÃO RETORNO	24,207%
OS.NOTA SAIDA RETORNO (CÓDIGO)	24,207%
OS.NR. DESENHO	6,153%

Nota. Autores (2023).

Apresentados no quadro 6. O quadro apresenta as porcentagens de preenchimento, desta forma optou-se em avaliar os atributos acima de 99%, tratando-se de OS.Data Previsão Entrega, OS.Supervisor (Código), OS.Encerramento Custos.

Seguindo a lógica de avaliação, foram desconsiderados da base os atributos “OS.Nr. Nota Remessa”, “OS.Emissão Remessa”, “OS.Oper.Fiscal”, “OS.Nota Saída Remessa (Código)”, “Exportação.CD_EQUIPAMENTO_EMPRESA”, “OS.Nr. Nota Retorno”, “OS.Emissão Retorno”, “OS.Nota Saida Retorno (Código)” e “OS.Nr. Desenho”.

Ao avaliar os atributos foi utilizado o método de imputação por frequência para completar o registros de “OS.Supervisor (Código)”. Para o atributo “OS. Encerramento custos” foi realizado a média do tempo entre a entrada e saída dos componentes similares no processo, e imputado o valor de finalização de custo, o método foi aplicado, pois trata-se de poucos registros.

No tratamento do atributo “OS.Encerramento Custos”, desenvolveu-se a tarefa correção de valores ausentes, ou seja, identificou os componentes que ficaram com a data em aberto. Para realizar a imputação dos valores faltantes, foi realizado o seguinte cálculo:

$$\text{OS.Encerramento Custos} - \text{OS.Emissão}$$

Porém, com a aplicação da fórmula observou-se valores negativos, representando que a data de encerramento está anterior a data de emissão, demonstrando informação errônea em 8.341 registros que estavam com as informações incoerentes, representado 14,26% dos registros presentes na base de dados. Desta forma, foram retirados todos os registros que apresentavam esse erro, tendo-se para a pesquisa um montante de 50.162 registros.

O *one-hot encoding*, outra técnica empregada no pré-processamento dos dados, diferencia comportamentos no banco de dados por 0 ou 1, assim facilita aos algoritmos de mineração de dados a realizarem leituras e entender padrões, para extrair conhecimento com maior clareza (Rodrigues et al., 2021).

Repetimos o processo de *One-Hot encoding* para a identificação de tipo de serviço, que será realizado no componente encaminhado a manutenção, onde discriminamos “Manutenção encaminhada de cliente” = 1, “Manutenção interna” = 0. A classificação ocorreu utilizando a regra, se tivermos “OS.Nr. Nota Remessa” então “Manutenção encaminhada de cliente” se não “Manutenção

interna”. A presente pesquisa não está estruturada para estudar as interdependências entre Ordens de Serviço e sim o comportamento do almoxarifado, logo foi criado *via One-Hot encoding* o atributo “OS.Escalonamento.Sub” que identifica como 1 os registros que são Sub-OS e 0 os registros que não são. Os primeiros dígitos que distinguem a unidade que encaminhou, podem ser identificados no atributo “OS.Nr. Empresa”, desta forma considerou-se o “OS.Nr.OS” um atributo redundante na base de dados, optando-se pela retirada do mesmo. O processo também foi realizado para a identificação de tipo de serviço, que será realizado no componente encaminhado a manutenção, discriminando-se como “Manutenção encaminhada de cliente” = 1, “Manutenção interna” = 0, considerando-se que se constar “OS.Nr. Nota Remessa” então “Manutenção encaminhada de cliente” se não “Manutenção interna”.

A eliminação de valores incoerentes no pré processamento foi construída por meio de uma correlação dos atributos e suas interdependências, considerando-se o quanto um atributo é importante para efetivar o conhecimento extraído dos registros analisados. Nessa etapa o cientista constrói uma matriz de dependência, em que é possível e identificar a importância dos atributos analisados, desconsiderando-se aqueles atributos com pouca dependência (D’Castro, 2020).

Após a aplicação das técnicas de pré processamento, apresentadas anteriormente, o processo foi finalizado com a base de dados apresentando 24 atributos e 50.162 registros (quadro 7).

Quadro 7.

Atributos finais da base de dados

ATRIBUTOS

OS.CÓDIGO
OS.ESCALONAMENTO.SUB
TIPO DE SERVIÇO
OS.EMIÇÃO.DATA
OS.EMIÇÃO.HORA
OS.QUANTIDADE
OS.DATA REQUERIDA ENTREGA
OS.DATA PREVISÃO ENTREGA
OS.NR. EMPRESA
OS.SOLICITANTE (CÓDIGO)
OS.SUPERVISOR (CÓDIGO)
OS.ENCERRAMENTO CUSTOS
OS.CÓD. COMPONENTE
OS.CÓD. EQUIPAMENTO
OS.CÓD. CENTRO CUSTO
OS.CD_USUARIO_INCLUSAO
EXPORTAÇÃO.NR. REQUISIÇÃO
EXPORTAÇÃO.DATA
EXPORTAÇÃO.HORA
TABELA_CÓDIGO_SUPERVISOR9.NOME SUPERVISOR
EXPORTAÇÃO.MATERIAL (CODIGO)

EXPORTAÇÃO.UN

EXPORTAÇÃO.QTD. MATERIAL

EXPORTAÇÃO.CENTRO DE CUSTO (CÓDIGO)

Nota. Autores (2023).

O pré-processamento dos dados deve adequar os registros, de forma a não permitir possíveis falhas à mineração de dados que irá resultar na extração do conhecimento, identificando-se possíveis desvio de padrões e corrigindo-os (A. Kumar et al., 2020).

3.4.2 Mineração de dados

A etapa de mineração de dados é responsável pela transformação da informação resultante do pré-processamento, realizando-se conforme o objetivo da análise por meio de diferentes tarefas, métodos e algoritmos.

Na descoberta do conhecimento os métodos empregados irão mapear as relações existentes entre os dados, identificando aqueles que são fundamentais ao estoque e quais componentes têm o uso associado. Na execução da mineração de dados se trabalhou diretamente com o setor de manutenção da indústria carbonífera, pois é o responsável por definir os objetivos e os requisitos do negócio, estabelecendo-se as diretrizes para o modelo de dados a ser gerado por meio da aplicação da mineração de dados.

Os modelos foram construídos a partir da definição das tarefas, dos métodos e dos algoritmos a serem aplicados à base de dados e a definição da ferramenta computacional a ser usada, entre as disponíveis gratuitamente. As ferramentas computacionais para mineração de dados suportam diferentes tipos de dados e algoritmos e demandam de diversos recursos de hardware e conhecimentos computacionais.

A mineração de dados é composta por diversas tarefas, sendo que nesta pesquisa optou-se pela tarefa de associação, pois segundo Librelotto & Mozzaquatro (2013) ela permite a identificação de padrões que acontecem juntos em uma base de dados. A identificação de regras de associação no processo de mineração de dados ocorre por meio de uma análise do conjunto de dados de forma transacional, deste modo, examina a existência de associação ou correlação entre os atributos avaliados, em diferentes níveis hierárquicos da árvore de associações (Silva et al., 2016).

Na tratativa da base de dados de forma transacional, foi definido como um conjunto de itens do domínio $I = \{X145263, X2427, \dots\}$ onde cada item apresentado é uma peça do almoxarifado. A tratativa transacional demonstra o número de ocorrências dos itens em $T = \{i_1, \dots, i_l\}$, tal que $i_l \subset I$ e $|T| \leq m$. Logo, segundo Silva et al. (2016) pode-se dizer que $T \subseteq I$ e o conjunto de transações é representada por TID, formalizando a criação de uma base de dados transacional.

Na base de dados transacional desenvolvida para a aplicação na presente pesquisa, define que se o registro estiver denominado por “t” então aconteceram, mas se os registros estiverem vazios, então não ocorreram.

Dentre os algoritmos para associação, nesta pesquisa empregou-se o algoritmo *Apriori*, a fim de se determinar a frequência e a associação dos componentes analisados. Assim, pode-se visualizar a dependência dos itens registrados, demonstrando o relacionamento entre o conjunto de peças que compõe o almoxarifado, podendo, por meio das análises, permitir um estoque mais eficiente e a construção de uma gestão de estoque efetiva.

O algoritmo *Apriori*, desenvolvido por Agrawal e Srikant (1994), é um dos mais usados quando se trata da tarefa de associação (Ramalho et al., 2020). O *Apriori* identifica a quantidade de ocorrências de cada item na base de dados, essa etapa é chamada de cálculo de suporte, definida pela equação 1.

$$Supp(LHS) = p(LHS) = \frac{n_{LHS}}{N}$$

Equação 1

A equação demonstra a frequência relativa de um conjunto LHS (probabilidade incondicional), em que n_{LHS} são o total de vezes que o conjunto LHS é observado e N é o total de transações presentes na base de dados.

Na sequência do processo o algoritmo trabalha para satisfazer a duas regras principais, sendo elas: ocorrer dentro de um suporte mínimo e com confiança, tornando possível a demonstração de dados com correlação forte. A confiança é definida pela equação 2.

$$conf(LHS \rightarrow RHS) = p(RHS|LHS) = \frac{p(LHS, RHS)}{p(LHS)}$$

Equação 2

A equação de confiança demonstra a probabilidade dos registros RHS participarem das transações dos registros LHS. Nesta pesquisa, definiu-se como suporte = 0.4 e confiança = 0.8, definido como média de frequência das análises realizadas. Esses valores foram considerados, conforme a literatura disponível na área (L. A. da Silva et al., 2016).

A existência de associações ou a correlação entre os atributos implica que eles frequentemente aparecem juntos em uma transação ou que uma variação na frequência de observação de um atributo num conjunto de transações ocorre com uma variação na frequência de observação de um segundo atributo nesse mesmo conjunto de transações.

A ferramenta computacional empregada para a realização da mineração de dados foi o software R, uma ferramenta gratuita que utiliza programação Java para desenvolver a inteligência de máquina e guiar a série de análises realizadas no decorrer da execução do estudo.

O presente estudo teve objetivo analisar os modelos gerados pela ciência de dados para apoio à tomada de decisão na gestão de inventário em uma indústria carbonífera. A primeira pergunta definida pela pesquisa foi elaborada em virtude de entender, quais as associações das transações realizadas no almoxarifado. Para a aplicação da primeira regra de mineração no banco de dados, foi evidenciado para a IA analisar dois grupos, sendo o grupo 1 a associação de registros de transações por componentes e o grupo 2, que trata da regra de associação de registros de transações por equipamento., realizou-se a análise do grupo 1 referente aos registros de transações por componentes e do grupo 2 que trata dos registros de transações por equipamento.

Na análise da base de dados, optou-se por separá-la por tipos de manutenção, a decisão foi estabelecida pelo poder computacional da máquina empregada na execução da pesquisa. Inicialmente, minerou-se a unidade mais requisitada no conjunto de dados (quadro 8).

A análise dos conjuntos de dados, foi realizado para identificar o maior grupo de manutenção presente, então foi avaliado primeiramente a unidade mais requisitada no almoxarifado.

Quadro 8.

Unidades presentes na base de dados

Exportação.Un	CONTAGEM	% CONTAGEM
10001	45239	90,19%
10010	1095	2,18%
10100	241	0,48%
11000	236	0,47%
1000	150	0,30%
100	1883	3,75%
10	302	0,60%
1	118	0,24%
1010	322	0,64%
110	389	0,78%
1110	17	0,03%
11111	108	0,22%
11011	32	0,06%
10101	17	0,03%
11001	1	0,00%
10000	8	0,02%
1011	3	0,01%

Nota. Autores (2023).

A unidade escolhida para análise foi a 10001, que representa a unidade por valor unitário requisitado no almoxarifado, para cada registro desse código foram requisitadas peças com valores inteiros do almoxarifado. Também foi realizada a análise por tipo de manutenção (quadro 9), por meio do atributo “Tabela_Código_supervisor.Nome Setor”.

Quadro 9.

Tipos de manutenção presentes na base de dados

TABELA_CÓDIGO_SUPERVISOR.NOME	CONTAGEM	% CONTAGEM
SETOR		
3404	3914	7,80%
3348	24895	49,63%
3048	10756	21,44%
3529	3942	7,86%
3036	6011	11,98%
3094	643	1,28%

Nota. Autores (2023).

Considerando-se os dados levantados, a pesquisa se direcionou para a manutenção 3348, que representa o conjunto de manutenção hidráulica, tendo-se o total de 23.483 dados para aplicação do método.

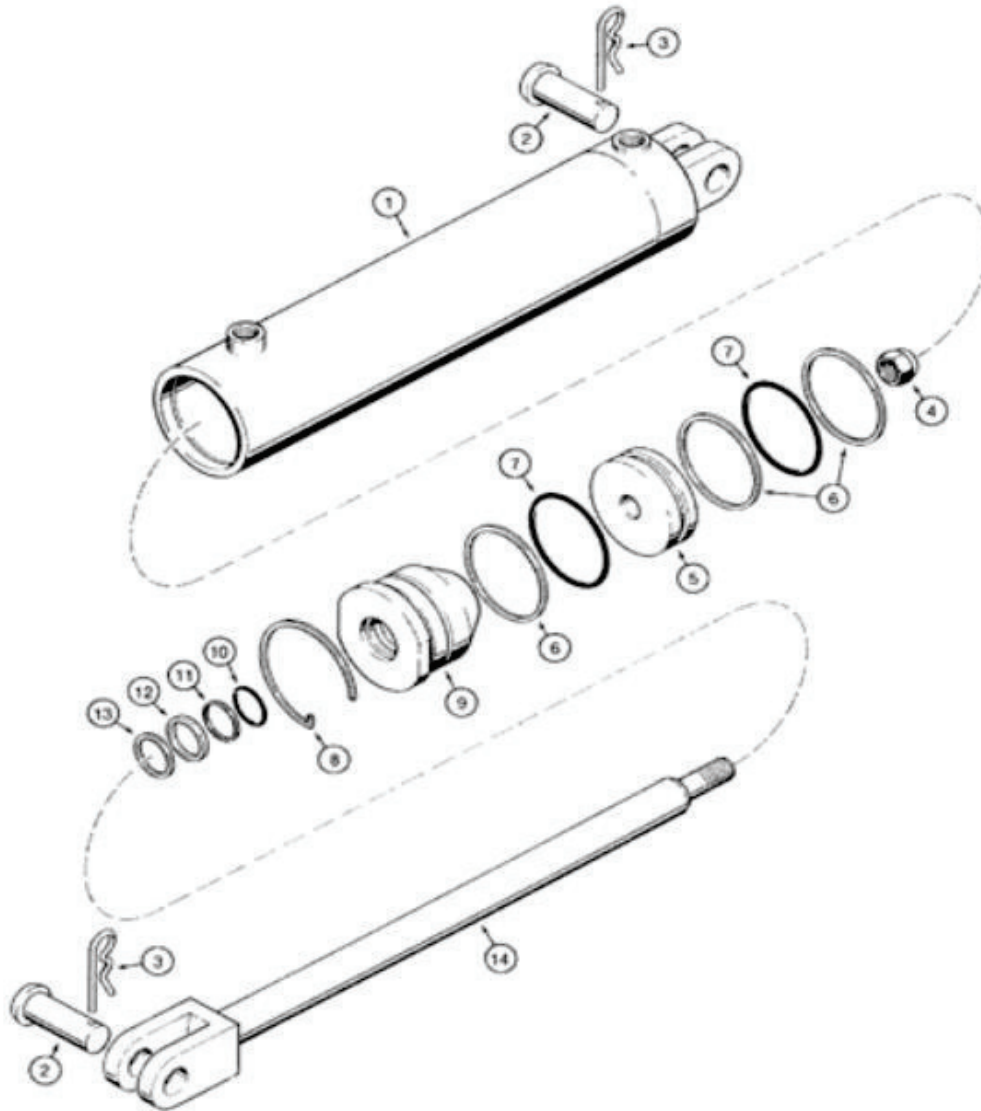
A manutenção hidráulica assegura o funcionamento adequado dos ativos. Quando executada e monitorada, permite a redução de vazamentos, que podem representar riscos ambientais, dependendo do fluido que estiver sendo liberado. Além disso, a manutenção hidráulica proporciona maior autonomia aos equipamentos em operação. A atuação da equipe hidráulica na carbonífera em estudo é essencial para garantir a manutenção dos componentes que utilizam fluidos como meio de ativação, incluindo-se os cilindros hidráulicos e bombas. Além disso, o setor é responsável pela fabricação de conexões que permitem o fluxo de fluidos nos equipamentos, com o objetivo de evitar possíveis vazamentos, constituindo-se em um importante setor desta indústria carbonífera.

Os dados em questão referem-se à movimentação de peças utilizadas no processo de manutenção de componentes hidráulico, como por exemplo, o anel Oring 14 – representado na posição 10 da figura 14 - que é montado no pistão do cilindro hidráulico de elevação da cabeça de corte do minerador JOY. Neste exemplo, buscou-se compreender a quantidade de anel Oring 14 necessários em estoque para garantir a disponibilidade ágil do cilindro hidráulico de elevação da cabeça de corte para o Minerador JOY. Assim, objetiva-se evitar falhas, como a falta desse item em

estoque, ao mesmo tempo em que se evita manter um volume elevado de anel, o que pode ocupar uma parcela significativo da curva estratégica de estoque¹⁰.

Figura 24.

Cilindro hidráulico de elevação da cabeça de corte do minerador JOY



Nota. (Agriculture, 2024).

A organização da base de dados foi realizada no PowerQuery, selecionando-se os quatro atributos *OS.Código*, *OS.Cód. Componente*, *Exportação.Material (Codigo)*, *Exportação.Qtd. Material*. Os atributos *Exportação.Material (Codigo)* e *Exportação.Qtd. Material* foram alterados de colunas para linhas, desta forma visualizou-se a base de dados por transações. A *OS.Código* representa o número da ordem de serviço, ou seja, a quantidade de vezes que a peça foi encaminhada, já *OS.Cód.*

¹⁰ A curva estratégica de estoque, defini via análise ABC as peças mais estratégicas para compor o estoque da empresa, esses itens podem ser avaliado pela ótica de custo, frequência e escassez, com a avaliação eficaz desses pontos, é possível definir a curva de estoque ideal (B. W. Silva, 2019).

Componente apresenta que em uma ordem de serviço foi encaminhado um determinado componente.

Com a base de dados organizada no formato CSV, iniciou-se o processo de aplicação do algoritmo *Apriori* no software R (figura 24).

Figura 25.

Algoritmo Apriori aplicado para associação de peças no almoxarifado e componentes.

```
> base <- read.table ("C:\\Users\\maria\\Desktop\\bast_hidr.csv",header=TRUE,sep=";")
> class (base)
[1] "data.frame"
> library ("arules")
Carregando pacotes exigidos: Matrix

Attaching package: 'arules'

The following objects are masked from 'package:base':

  abbreviate, write

> base_convertida <- as (base, "transactions")
Warning message:
Column(s) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 2
4, 25, 26, 27, 28, 29, 30, 31, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49,
50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 7
3, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96,
97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 1
16, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134,
135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153,
154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172
, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 19
1, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 2
11, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 2 [... truncated]
> regra_componentes <- apriori (data=base_convertida, parameter=list(minlen=2, supp=0.4, conf=
0.8)
+ )
Apriori

Parameter specification:
confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target  ext
0.8      0.1    1 none FALSE          TRUE         5     0.4      2    10 rules TRUE

Algorithmic control:
filter tree heap memopt load sort verbose
0.1 TRUE TRUE  FALSE TRUE    2    TRUE

Absolute minimum support count: 1403

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[2450 item(s), 3509 transaction(s)] done [1.02s].
sorting and recoding items ... [1225 item(s)] done [0.13s].
creating transaction tree ... done [0.04s].
checking subsets of size 1 2 done [6.86s].
writing ... [1482446 rule(s)] done [0.68s].
creating S4 object ... done [0.56s].
Warning message:
In apriori(data = base_convertida, parameter = list(minlen = 2, :
Mining stopped (time limit reached). Only patterns up to a length of 2 returned!
>
```

Nota. Autores (2023).

A função `inspectCom` as regras geradas, solicitou-se a leitura utilizando a função `inspect(head())`, que faz parte do pacote `dplyr` do software R, que tem como objetivo trazer mais eficiência ao processamento e análise dos resultados. Neste pacote estão disponíveis várias funções de agrupamento, filtros `ADDIN CSL_CITATION`

```
{
  "citationItems": [
```

```
{
  "id": "ITEM-1",
  "itemData": {
    "author": [
      {
        "dropping-particle": "da",
        "family": "Silva",
        "given": "Leandro Augusto",
        "non-dropping-particle": "",
        "parse-names": false,
        "suffix": ""
      },
      {
        "dropping-particle": "",
        "family": "Peres",
        "given": "Sarajane Marques",
        "non-dropping-particle": "",
        "parse-names": false,
        "suffix": ""
      },
      {
        "dropping-particle": "",
        "family": "Boscarioli",
        "given": "Clodis",
        "non-dropping-particle": "",
        "parse-names": false,
        "suffix": ""
      }
    ],
    "id": "ITEM-1",
    "issued": {
      "date-parts": [
        [
          2016
```

```

    ]
  ]
},
"publisher": "Elsevier",
"title": "Introdução a mineração de dados com aplicação em R",
"type": "book"
},
"uris": [
  "http://www.mendeley.com/documents/?uuid=fb33bf6b-f0d1-44ec-8efb-ec60c6396880"
]
}
],
"mendeley": {
  "formattedCitation": "(L. A. da Silva et al., 2016)",
  "plainTextFormattedCitation": "(L. A. da Silva et al., 2016)",
  "previouslyFormattedCitation": "(L. A. da Silva et al., 2016)"
},
"properties": {
  "noteIndex": 0
},
"schema": "https://github.com/citation-style-language/schema/raw/master/csl-citation.json"
}

```

(*head()*), que retorna os seis primeiros registros encontrados na regras de associações, identificadas pelo maior lift (Silva et al., 2016).

Na análise de associação entre as peças do almoxarifado e os equipamentos hidráulicos, estruturou-se a base utilizando a mesma metodologia, porém substituíram-se os componentes por equipamentos na transformação da base em transacional. Foram selecionados os quatro atributos *OS.Código*, *OS.Cód. Equipamento*, *Exportação.Material (Codigo)* e *Exportação.Qtd. Material*, realizando-se as análises pelo algoritmo *Apriori* referente as peças requisitadas para os equipamentos (figura 26).

Figura 26.

Algoritmo Apriori aplicado para associação de peças no almoxarifado e equipamentos.

```
> base <- read.table("C:\\Users\\maria\\Desktop\\Dissertação\\16 - Etapa Mineração\\Equipamen
tos.csv", header=TRUE, sep=";")
> class(base)
[1] "data.frame"
> library("arules")
Carregando pacotes exigidos: Matrix

Attaching package: 'arules'

The following objects are masked from 'package:base':

  abbreviate, write

> base_convertida <- as(base, "transactions")
Warning message:
Column(s) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 2
4, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47,
 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 7
1, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94,
 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114
, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 13
3, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 1
52, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170,
 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189,
 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208
, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220 [... truncated]
> regra_2 <- apriori(data=base_convertida, parameter=list(minlen=2, supp=0.4, conf=0.8))
Apriori

Parameter specification:
 confidence minval smax arem aval originalSupport maxtime support minlen maxlen target ext
           0.8   0.1   1 none FALSE           TRUE         5     0.4     2    10 rules TRUE

Algorithmic control:
 filter tree heap memopt load sort verbose
  0.1 TRUE TRUE FALSE TRUE     2 TRUE

Absolute minimum support count: 1402

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[2848 item(s), 3507 transaction(s)] done [1.72s].
sorting and recoding items ... [1424 item(s)] done [0.22s].
creating transaction tree ... done [0.06s].
checking subsets of size 1 2 done [5.93s].
writing ... [2026352 rule(s)] done [0.75s].
creating S4 object ... done [0.47s].
Warning message:
In apriori(data = base_convertida, parameter = list(minlen = 2, :
  Mining stopped (time limit reached). Only patterns up to a length of 2 returned!
>
```

Nota. Autores (2023).

3.4.3 Pós-processamento dos dados

A etapa de pós-processamento dos dados compreende a avaliação dos padrões descobertos, empregando-se para isso medidas descritas como *user-driven* e *data-driven*. As medidas *user-driven* são aquelas em que os usuários especificam as suas crenças, baseando-se no conhecimento que têm sobre o domínio de aplicação, enquanto as medidas *data-driven* estimam de forma objetiva e automática a qualidade dos modelos.

O pós-processamento dos dados nesta pesquisa consistiu na aplicação das medidas de qualidade, também denominadas de índices de validação, que determinam a qualidade dos modelos de dados construídos. Estes índices variam conforme a tarefa de mineração de dados empregada. No caso desta pesquisa, que empregou a tarefa de associação, utilizaram-se as medidas de suporte, confiança e lift. Além disso, realizou-se uma análise de valor sobre os resultados obtidos na etapa de mineração de dados.

O suporte na regra de associação está ligado diretamente com o número de transações que contém o conjunto analisado, considerando-se a frequência, portanto é viável somente aqueles em que o suporte estiver acima do suporte mínimo informado na aplicação do algoritmo *Apriori* (Vasconcelos & Carvalho, 2004). No presente estudo foi definido como padrão o suporte mínimo de 40%, desta forma o algoritmo analisa somente as transações que estiverem com frequência superior a 40% na base de dados.

No que se refere ao suporte, a base de dados é analisada de forma integrativa, porém na regra de associação também se avalia a confiança dos registros transacionais. A confiança permite avaliar quantas vezes o item X aparece junto ao item Y (Cunha Filho & Rocha-Junior, 2022). Nesta pesquisa, na aplicação do algoritmo *Apriori* na base de dados utilizou-se como confiança mínima 80%, demonstrando que o algoritmo deve considerar apenas os registros que possuem a confiança maior que 80% para a construção da regra.

A confiabilidade das informações obtidas, conforme os valores mínimos definidos, são calculadas pelo valor de *lift*, definido na equação 1.2 (Nomelini et al., 2010).

$$Lift (lhs \rightarrow rhs) = \frac{Supp (lhs \rightarrow rhs)}{Supp(lhs) \times Supp (rhs)} \quad (1.2)$$

Com a aplicação dos valores na equação, é possível a avaliação da qualidade dos resultados para a regra, definindo-se que: se *lift* =1 então LHS \Rightarrow RHS são independentes; se *lift* > 1 então LHS \Rightarrow RHS são positivamente dependentes; e, Se *lift* < 1, então LHS \Rightarrow RHS são negativamente dependentes (Schonhorst, 2010).

Na regra de associação entende-se que quanto maior o *lift*, maior a chance das associações analisadas acontecerem. Por exemplo, se $LHS \Rightarrow RHS$ possuir um *lift* = 4 então, tem-se 4 vezes mais chance de ocorrer RHS, se LHS ocorreu.

A análise de valor é conduzida com base nos resultados obtidos na etapa de mineração de dados. Dessa forma, foram avaliados os dois cenários propostos: equipamentos e componentes. A demonstração da análise utiliza ferramentas como o histograma, que é responsável por apresentar a frequência, média e mediana de um determinado conjunto de dados, mostrando a frequência de ocorrência de cada valor em uma determinada classe (Morenttin & Bussab, 2017).

Após a obtenção dos resultados da etapa de mineração de dados, foram identificados os registros no banco de dados que continham determinada peça. Com isso, foi possível identificar com quais componentes essa peça estava correlacionada. Em seguida, procedeu-se ao cálculo de frequência, avaliando quantas vezes determinados componentes eram encaminhados para manutenção. A avaliação foi realizada de forma retroativa, mas está projetada para fornecer visões sazonais futuras.

Com uma visão do comportamento retroativo, foi identificada a relação entre a quantidade total de peças requisitadas por componente encaminhado para manutenção, permitindo a construção dessa relação entre componentes, peças e suas associações. Com os valores obtidos nas análises, é possível determinar a quantidade de peças necessárias em estoque para atender à demanda prevista mensalmente. O pós-processamento dos dados envolve a visualização, a análise e a interpretação do modelo final identificado pela aplicação das medidas de qualidade. Assim, entende-se os padrões extraídos por meio da interpretação pelos analistas que possuem conhecimento do negócio. No entanto, de acordo com Silva et al. (2016), a tarefa de associação disponibiliza informações com fácil entendimento ao cientista, não sendo obrigatória a utilização do pós-processamento para realizar a leitura e interpretação dos resultados obtidos.

3.5 PROCEDIMENTOS ÉTICOS DA PESQUISA

Os critérios de inclusão dos dados a serem coletados referem-se aos suprimentos de inventário na unidade de manutenção de uma indústria carbonífera, no período compreendido entre 2019 e 2022. Os dados foram extraídos do sistema de gestão da empresa pelos funcionários do Setor de Tecnologia e Informação e o analista de suprimentos, os quais posteriormente foram disponibilizados aos pesquisadores.

Esta pesquisa não precisou ser submetida ao Comitê de Ética em Pesquisa, visto que os dados são industriais e oriundos do sistema de gestão da organização, não tendo-se informações referente a indivíduos.

No que se refere aos riscos desta pesquisa tem-se a perda da confidencialidade dos dados e este risco será amenizado pela privacidade mantida, não sendo divulgado os dados da pessoa jurídica. Nesta pesquisa a aplicação da Lei Geral de Proteção de Dados não é necessária, pois aplica-se somente em informações pessoais, como número de identificação, e-mail, entre outros atributos de caráter particular do indivíduo. As informações manuseadas no presente estudo, trata-se de inventário de manutenção e não irá trabalhar com dados de pessoa física.

4 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados provenientes da pesquisa, que consiste na identificação de um modelo de ciência de dados por meio da tarefa de associação pelo algoritmo *Apriori* para apoio à tomada de decisão na gestão de inventário, no que se refere a manutenção hidráulica. O conhecimento descoberto nos dados foi combinado com a análise de risco a fim de poder auxiliar na tomada de decisão eficaz na gestão de inventários de manutenção hidráulica. Também será demonstrado o impacto ao mundo acadêmico e a possibilidade de exploração do tema proposto.

4.1. ACHADOS CIENTÍFICOS

A pesquisa é elaborada no âmbito interdisciplinar, assim pode oferecer novas diretrizes nas teorias do tema proposto, possibilitando uma avaliação sob diferentes perspectivas (Huang et al., 2018). O trabalho adjunto entre cientistas e especialistas da área do domínio de aplicação, assegura a interdisciplinaridade, pois permite a troca de informações e conhecimento de diferentes perspectivas (Claude Raynaut, (2014).

A extração do conhecimento, quando realizada no âmbito interdisciplinar, proporciona a revelação de conhecimentos não explorados. Isso ocorre porque é feita sob perspectivas diferentes daquelas frequentemente utilizadas no processo. A visão interdisciplinar permite trabalhar assuntos não explorados e possibilita a extração do conhecimento por vários ângulos distintos (Fazenda, 2017).

De acordo com Costa et al. (2023) a combinação de métodos de inteligência artificial com análise de valor possibilita a compreensão dos resultados, contribuindo para a melhoria do sistema avaliado. Esse potencial é especialmente relevante quando se trata de ganhos financeiros ou de eficiência no processo produtivo.

Considerando-se os nichos empresariais, a IA é utilizada para mapear e mensurar resultados relacionados a comportamentos ou atividades padrões. Ela realiza um controle sobre variáveis e constrói bancos de dados robustos, dos quais é possível extrair informações cruciais para o ramo empresarial. Além disso, os gestores utilizam métodos de análise de valor para fomentar a estratégia real da empresa, permitindo um melhor comportamento estratégico do negócio (Avelar et al., 2021). O desenvolvimento de modelos de análise de valor adjunto a inteligência artificial, permite um entendimento mais apurado sobre as análises (Sucena & Cury, 2023).

A presente pesquisa apresenta os resultados obtidos no processo de mineração de dados e de avaliação dos modelos gerados pela ciência de dados, descrevendo o conhecimento descoberto, combinados com a análise de valor, para a tomada de decisão eficaz na gestão de inventários de manutenção.

Com base nos métodos de mineração de dados empregados na pesquisa (seção 3.4.2), iniciou-se o entendimento dos resultados gerados pelo algoritmo *Apriori*. Como primeiro passo, o algoritmo foi instruído a criar regras de associação entre as peças requisitadas no almoxarifado, visando a manutenção de componentes hidráulicos. O processo gerou um total de 1.482.446 regras de associação.

Tabela 8.

Apresentação das seis associações com os maiores lift

	LHS	RHS	SUPP	CONF	COVERAGE	LIFT	Nº
[1]	{X135072=}	=> {X2712=}	0,6617270	0,978920	0,6759761	1,4318600	2322
[2]	{X2712=}	=> {X135072=}	0,6617270	0,967903	0,6836706	1,4318600	2322
[3]	{X135072=}	=> {X105881=}	0,6546025	0,968381	0,6759761	1,4029930	2297
[4]	{X155881=}	=> {X135072=}	0,6546025	0,948389	0,6902251	1,4029930	2297
[5]	{X135072=}	=> {X9815=}	0,6736962	0,996627	0,6759761	1,4112850	2364
[6]	{X9815=}	=> {X135072=}	0,6736962	0,953995	0,7061810	1,4112850	2364

Nota. Autores (2023).

A tabela 08 apresenta as seis regras com maiores lift, definidos pela função *inspect(head())*, que demonstra a probabilidade do LHS ser requisitado no almoxarifado com o RHS é alta. Como resultados, pode-se ter Lift (LHS → RHS) > 1, então há uma alta probabilidade de dependência das duas variáveis, representando que ao requisitar a peça LHS, também se solicita a peça RHS, caso contrário, quando o Lift (LHS → RHS) ≤ 0, então é provável que RHS não seja requisitado junto a peça LHS.

O resultado da aplicação do algoritmo *Apriori* neste cenário, permitiu avaliar as peças requisitadas no almoxarifado para componentes hidráulicos, tendo-se, por exemplo, que a peça X135072 é associada a peça X2712. Assim, consultando-se o dicionário de dados, identificaram-se as seguintes associações:

- 135072 - ANEL "O" RING 2014 DUREZA 90 SHORE é requisitado junto com o 2712 ANEL "O" RING 2243 PARKER.
- 135072 - ANEL "O" RING 2014 DUREZA 90 SHORE é requisitado junto com o 105881 ANEL "O" RING 2227 PARKER
- 135072 - ANEL "O" RING 2014 DUREZA 90 SHORE é requisitado junto com o 9815 ARRUELA LISA 3/8".

Assim, o algoritmo estabeleceu as regras de associação entre as peças requisitadas no almoxarifado e os equipamentos em que são realizadas a manutenção hidráulica.

A aplicação do algoritmo *Apriori* com relação aos equipamentos (figura 25) gerou 2.026.352 regras de associação, 543.906 regras a mais que a associação aplicada em componentes hidráulicos.

Tabela 9.

Apresentação das seis associações com os maiores lift

	LHS	RHS	SUPP	CONF	COVERAGE	LIFT	Nº
[1]	{X48735=}	=> {X149937=}	0,9124608	1,0000	0,9124608	1,095938	3200
[2]	{X149937=}	=> {X48735=}	0,9124608	1,0000	0,9124608	1,095938	3200
[3]	{X48735=}	=> {X1298=}	0,9056173	0,9925	0,9124608	1,087718	3176
[4]	{X1298=}	=> {X48735=}	0,9056173	0,9925	0,9124608	1,087718	3176
[5]	{X48735=}	=> {X145263=}	0,9056173	0,9925	0,9124608	1,087718	3176
[6]	{X145263=}	=> {X48735=}	0,9056173	0,9925	0,9124608	1,087718	3176

Nota. Autores (2023).

A tabela 09 apresenta os seis maiores lift encontrados durante a aplicação da regra de associação. Como resultado da regra, aplicando o algoritmo *Apriori* nas transações para equipamentos hidráulicos, tem-se que X48735 é associado a X149937, com um lift de 1,95938, consultando-se o dicionário de dados tem-se que:

- X48735 - 'PARAFUSO ALLEN C/C 3/4" X 1.1/2" UNC – METALAC, é requisitado junto com X149937 - 'TIRANTE PARA COMANDO CH100 3F ROSCA M10 PASSO 1,0MM X 250MM.
- X48735 - 'PARAFUSO ALLEN C/C 3/4" X 1.1/2" UNC – METALAC, é requisitado junto com X1298 - 'BROCA 5,0 MM AR – DORMER.
- X48735 - 'PARAFUSO ALLEN C/C 3/4" X 1.1/2" UNC – METALAC, é requisitado junto com X145263 - 'ANEL PARBAK 8344 PARKER.

A partir das regras identificadas pela IA, por meio da tarefa de associação e do algoritmo *Apriori*, aplicou-se aos resultados obtidos o conceito de análise de valor. Segundo (Battaglia & Bergamo (2010), a análise de valor visa entender os produtos que estão presentes no processo produtivo e analisá-los de forma quantitativa em prol de reduzir os gargalos dos processos e diminuir o custo de produção.

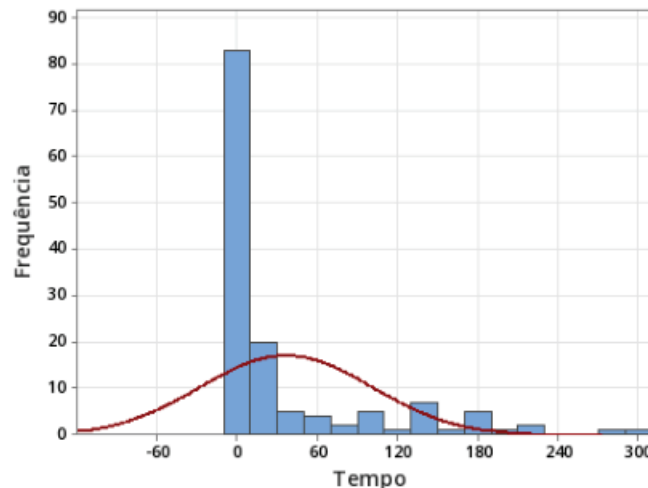
A identificação de possíveis oportunidades no sistema produtivo analisado, junto ao racionamento de produtos e a permissão de uma maior eficiência do processo, está relacionada inteiramente com a análise de valor, definida por Piran (2018) como a ciência que possibilita o entendimento dos sistemas produtivos a partir da análise do comportamentos dos dados, para o direcionamento da tomada de decisão.

Na obtenção dos resultados referente a análise de valor, foram consideradas as peças encontradas no modelo e apresentadas na tabela 14, verificando-se os componentes que são

requisitados. Ao se analisar a peça 135072 ('ANEL "O" RING 2014 DUREZA 90 SHORE), identificou-se que no período avaliado, a mesma foi requisitada no almoxarifado 138 vezes, frequência apresentada na figura 27.

Figura 27.

Tempo entre início da manutenção x requisição da peça 135072 no almoxarifado



Nota. Autores (2024).

O tempo médio entre a chegada do componente para manutenção e a retirada da peça 135072 no almoxarifado, é de 37 dias, com uma mediana de 7 dias, sua maior frequência é vista no intervalo [0,43] dias.

O setor recebeu 293 ordens de manutenção destes componentes, em que foram requisitadas 592 peças 135072, porém a frequência de requisição de peça é diferente em cada componente, variando de [1/4 a 13]. Portanto, a cada quatro vezes que o componente 3248 é encaminhado para a unidade, é utilizada uma peça 135072, porém quando se avalia o componente 4263, observa-se que sempre que solicitado manutenção do mesmo são necessárias 13 peças 135072.

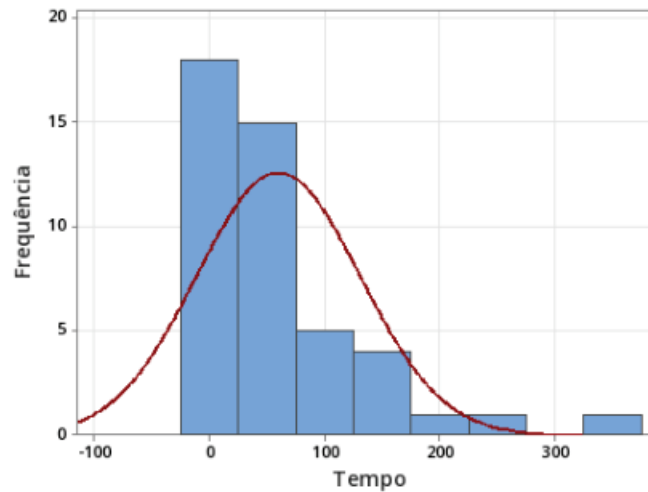
Após a avaliação da probabilidade de requisição da peça nos componentes, encontrou-se uma relação 2/1, ou seja, a cada componente encaminhado, são solicitadas duas peças 135072. Avaliando-se o comportamento no período de um mês, tem-se a probabilidade de receber 46 componentes, representando a probabilidade de requisitar 92 peças. Assim, pode-se definir como o valor mínimo em estoque para atendimento da demanda mensal, 92 peças 135072 no almoxarifado, assim como a peça 2712, identificado pela tarefa de associação com o lift 1,4318600, representando uma correlação forte entre as duas peças.

No que se refere ao cenário das peças requisitadas com frequência em equipamentos, também se realizou a análise de valor. Por exemplo, utilizando-se como base a peça 48735 ('PARAFUSO ALLEN C/C 3/4" X 1.1/2" UNC – METALAC), identificada pela tarefa de associação com o lift de 1,0959, que foi requisitada no almoxarifado 45 vezes durante o período de avaliação, associada

a 9 equipamentos distintos. O intervalo entre a entrada do equipamento para manutenção e a requisição da peça no almoxarifado está demonstrado na figura 28.

Figura 28.

Tempo entre início da manutenção x requisição da peça 48735 no almoxarifado



Nota. Autores (2024).

O tempo médio entre a chegada da peça e a requisição dela no almoxarifado é de 60,2 dias, com uma mediana de 32 dias, porém como demonstrado na figura 28, o intervalo frequente da mesma está entre [1,71] dias.

A peça 48735 foi requisitada para 9 equipamentos hidráulicos, que foram encaminhados 1.948 vezes para manutenção durante o período, sendo requisitado o total de 323 peças. Entre os 9 equipamentos hidráulicos avaliados, pode-se observar diferença entre a frequência de requisição das peças, variando de [1/234, 8/9]. Assim, a cada 234 manutenções realizadas no equipamento 547, requisitou-se 1 peça, porém a cada 9 manutenções realizadas no equipamento 774, solicitaram 8 peças. Fundamentando as relações apresentadas, tem-se que a cada 6 manutenções realizadas no conjunto de equipamentos avaliado, tem-se que é requisitada 1 peça 48735 no almoxarifado.

Realizando a análise do comportamento no período de 1 mês, tem-se a probabilidade de receber 43 equipamentos para manutenção, em que serão requisitadas 7 peças 48735. Logo, o valor mínimo em estoque é de 8 peças 48735 e 8 peças 149937 ('TIRANTE PARA COMANDO CH100 3F ROSCA M10 PASSO 1,0MM X 250MM), identificados pela tarefa de associação com o lift 1,095938, representando uma correlação forte entre as duas peças.

O intuito principal da análise de valor dos dados extraídos pelas regras de associação é diminuir o *lead time* do processo de manutenção, condicionado a chegada do material e a sua requisição no almoxarifado.

Nesta pesquisa, ao todo foram identificadas, pelas regras de associação, 24 peças que podem ter sua curva estratégica de estoque ajustada, a fim de atender a demanda de manutenção. No entanto, como abordado anteriormente, a análise de valor permite o ajuste do estoque, logo com o

objetivo de se aumentar a eficiência do processo, deve-se analisar também as peças que não são requisitadas com muita frequência e que têm um armazenamento excessivo em estoque, além de outras variáveis que podem servir como base.

A análise de valor possui um direcionamento fundamental para a estruturação da estratégia da empresa, pois nela é possível a leitura de controles que permitem a eficiência do processo, junto a movimentação dos custos ou lucros da empresa (Jae-Young Oh Yong-Geon Lee, 2020). O estudo de análise de valor realizado na presente pesquisa, não tem como foco o custo do produto e sim o entendimento da variabilidade do processo.

Durante a elaboração do estudo bibliométrico, para direcionamento da presente pesquisa, identificou-se uma escassez de trabalhos que relacionam a cadeia de suprimentos de manutenção com mineração de dados, portanto a presente pesquisa é de caráter exploratório e permite ao meio científico um novo olhar no âmbito deste setor do sistema produtivo da indústria, em especial da carbonífera, em que se aplicou a pesquisa.

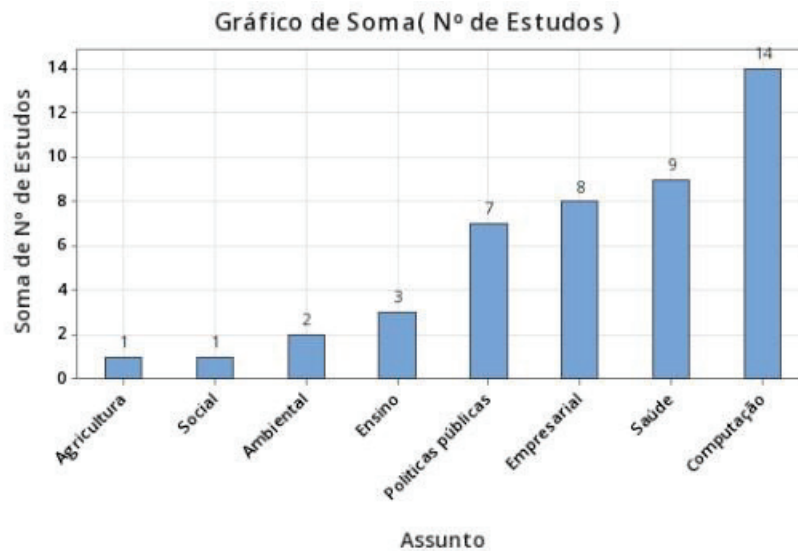
O setor de manutenção no século XXI tem grande importância no sistema produtivo, pois as fábricas encontram-se cada vez mais tecnológicas e robotizadas, portanto, deve-se estimular a atualização do sistema produtivo por meio das novas tecnologias, como as de inteligência artificial.

Na presente pesquisa, em que os resultados se originaram da aplicação da mineração de dados e da análise de valor, realizou-se uma busca nos periódicos CAPES, referente ao número de publicações relacionadas com o tema de estudo. Para isso, utilizou-se o seguinte critério "Value Analysis" AND "Artificial Intelligence", no qual foram encontrados 45 artigos, publicados entre 1987 e 2024.

Os estudos encontrados são referentes a oito domínios de aplicação (figura 29) e têm como tema a aplicação da inteligência artificial e a utilização do raciocínio de análise de valor como modelo matemático, para auxiliar no tratamento das informações processadas pela IA.

Figura 29.

Número de estudos por domínio de aplicação.



Na Figura 29 observa-se que a área da computação se destaca, com foco no aprimoramento de ferramentas existentes, combinando inteligência artificial e análise de valor para tornar as análises mais eficientes e extrair conhecimento. Na área da saúde, essas ferramentas são aplicadas para auxiliar em atendimentos médicos e diagnósticos. O campo empresarial, o foco principal, também está presente, com dois artigos relacionados ao tema da pesquisa.

O estudo, *Supply Chain Based Demand Analysis of Different Deep Learning Methodologies for Effective Covid-19 Detection* (Srujana et al., 2022), concentra-se em compreender o comportamento das cadeias de suprimentos durante a pandemia. O principal objetivo é investigar como a inteligência artificial está sendo utilizada para aprimorar a organização das demandas empresariais, aumentando a visibilidade e reduzindo riscos nas variáveis que impactam a distribuição de suprimentos no processo produtivo. Além disso, para o período pós-pandemia, muitas empresas desenvolveram planos de gestão com foco na transformação da cadeia de suprimentos, e a gestão de valor desempenha um papel crucial nesse contexto.

O estudo, *Advanced Anomaly Detection in Manufacturing Processes: Leveraging Feature Value Analysis for Normalizing Anomalous Data* (Kim et al., 2024), consiste na compreensão por meio da inteligência artificial das principais tendências de falhas nos componentes do processo produtivo industrial, incluindo a medição de vibração e outras variáveis. A IA pode sinalizar a necessidade de manutenção preventiva antes que uma falha significativa no equipamento mecânico ocorra, resultando em ganhos de produtividade.

Embora os outros artigos encontrados na busca sejam relevantes, eles não estão diretamente relacionados ao tema proposto nesta pesquisa. Observa-se que os dois estudos que possuem alguma relação com o tópico abordado, não se concentram diretamente no foco principal da presente

pesquisa que consiste em analisar os modelos gerados pela ciência de dados para apoiar a tomada de decisão na gestão do inventário de manutenção. Portanto, considera-se que esta pesquisa contribui para a aplicação da inteligência artificial em uma área ainda não explorada em estudos anteriores.

4.2 APLICABILIDADE DO ESTUDO PARA SUA ÁREA DE CONHECIMENTO, SETOR/SEGMENTO OU REGIÃO

Nos resultados apresentados, pode-se observar por meio da tarefa de associação pelo algoritmo *Apriori* as regras de associação correlacionadas aos equipamentos e aos componentes hidráulicos na base de dados transacional, entendendo-se a associação entre as mais diversas variáveis presentes no almoxarifado de manutenção carbonífera.

A aplicação da ciência de dados na área carbonífera, proporciona o avanço tecnológico neste sistema produtivo, que é fortemente apoiado pelo governo federal para a evolução do parque fabril de extração mineral do sul do Brasil. Esta tecnologia pode trazer benefícios econômicos, como a redução de itens no estoque com menor necessidade e o fortalecimento daqueles que têm demanda produtiva. O ciclo ininterrupto do processo de carvão mineral via quebra de máquinas, diminui os impactos ambientais ocasionados, por exemplo, pelo derramamento de óleo, queima de energia com o deslocamento de equipamento entre longas distâncias e redução de resíduos.

A pesquisa foi desenvolvida compreendendo a aplicação da ciência de dados na gestão de inventário, podendo fomentar uma visão estratégica para as indústrias da área na região sul de Santa Catarina, como também em outras regiões do país, podendo a abordagem ser replicada para outros domínios de aplicação na área industrial, por exemplo. Assim, por meio de estudos como este, pode-se incentivar a modernização do parque fabril, o aprimoramento da mão de obra e o fortalecimento do vínculo universitário, social, empresarial e político.

O incremento da mão de obra está associado a capacidade de dominar ferramentas tecnológicas e poder analisá-las, extraíndo-se conhecimento a partir de informações inativas no processo. Como resultado, a pesquisa permite a replicação do estudo para diversas áreas, não somente as voltadas a extração mineral e ao setor de manutenção, com o intuito de alavancar o parque fabril e oportunizar as pequenas e médias empresas a aplicação de tecnologias atuais que proporcionam a inovação e a transformação digital nas organizações.

O impacto legal está associado a lei de transição energética justa, que justifica e direciona o presente estudo, fomentando novas ideias para o cumprimento dos artigos sancionados pela lei.

5 CONSIDERAÇÕES FINAIS

A presente pesquisa consistiu na análise dos modelos gerados pela ciência de dados para apoio à tomada de decisão na gestão de inventário em uma indústria carbonífera. Para a construção do modelo, foram aplicados diversos métodos e metodologias que permitiram o desenvolvimento do estudo, a organização de uma base de dados adequada as necessidades do algoritmo *Apriori*, descobrindo-se as relações entre os dados por meio da tarefa de associação.

A organização da base de dados, por meio das atividades envolvidas no pré-processamento dos dados, demonstrou a importância da confiabilidade dos dados, pois pode inutilizar ou banalizar as relações descobertas na execução da mineração de dados. No pré-processamento foram avaliados vários pontos de baixa confiabilidade dos dados, fazendo com que esta etapa compreendesse um processo extenso e rigoroso.

A aplicabilidade da ciência de dados na base referente a manutenção foi observada pelo entendimento dos resultados identificados pelo algoritmo *Apriori*, que permitiu o entendimento das regras de associação entre as peças requisitadas para os componentes e os equipamentos hidráulicos.

A base de dados construída constitui-se por diversos atributos, empregando-se nesta pesquisa as peças requisitadas pelo setor hidráulico no almoxarifado, com valores unitário e considerando-se a avaliação para dois conjuntos diferentes, sendo eles os componentes e os equipamentos.

Os resultados obtidos pela geração das regras de associação na aplicação da mineração de dados, foram submetidos a uma análise de valor, construída no intuito de fomentar a utilização híbrida da ciência de dados e da análise de valor no processo de análise da gestão de inventários da cadeia de suprimentos. Assim, pode-se entender a dinâmica de recebimento de tais componentes para a manutenção, estipulando-se os limites mínimos de peças no almoxarifado para atender a demanda do setor de manutenção.

Mediante a pesquisa realizada, concluiu-se que a aplicação do método concebido no modelo de gestão de uma indústria carbonífera é viável, podendo complementar a estratégia da empresa, considerando-se as diretrizes traçadas pelo time de gestão empresarial, a fim de auxiliar nas movimentações financeiras, referente as prioridades de aquisição, bem como pode apoiar o processo ativo de tomada de decisão. Também, os métodos empregados podem ser replicados nos mais diferentes nichos industriais e do conhecimento, podendo permitir a integração entre sociedade, universidade e ambiente empresarial, fomentando os pilares do mestrado de sistemas produtivos e podendo ocasionar impactos sociais, tecnológicos e econômicos.

Como sugestão para pesquisas futuras, tem-se por exemplo, a construção de um modelo de análise de valor, utilizando ferramentas da ciência de dados na gestão de inventário; a aplicação do modelo a outros setores da indústria carbonífera, além da manutenção hidráulica; e, a aplicação de

diferentes algoritmos de associação e de outras tarefas de mineração de dados, a fim de identificar relações presentes na base de dados.

REFERÊNCIAS

- Abbasi, B., Babaei, T., Hosseinifard, Z., Smith-Miles, K., & Dehghani, M. (2020). Predicting solutions of large-scale optimization problems via machine learning: A case study in blood supply chain management. *Computers and Operations Research*, *119*, 104941.
<https://doi.org/10.1016/j.cor.2020.104941>
- Abreu, A., & Gomes, J. Á. P. (2013). *Aplicação do Desenho de Experiências na Definição de Estratégias Operacionais*. October 2016.
- Agostinelli, S., & Cumo, F. (2017). MACHINE LEARNING APPROACH FOR PREDICTIVE MAINTENANCE IN AN ADVANCED BUILDING. *Energy Production and Management in the 21st Century*, *131*, nan. <https://doi.org/10.2495/EPM220111>
- Agrawal, R., & Srikant, R. (1994). Fast Algorithms for Mining Association Rules. *Paper Presented at the Proceedings of the 20th International Conference on Very Large Data Bases, Santiago*.
- Agriculture, C. (2024). *CASE IH DISK HARROW (1/96-12/02) (08-21) - TRANSPORT LIFT CYLINDER - 548004R92, WELDED TYPE, 4 INCH ID*.
- Alt, P. R. C., & Martins, P. G. (2009). *Administração de materiais e recursos patrimoniais* (3rd ed.). Saraiva.
- Alves, G. E. A. B. P. De. (2003). Pré-processamento de Dados em Aprendizado de Máquina Supervisionado. *American Journal of Distance Education*, 1–204.
- Angeles, R. (2016). STEADYSERV BEER: IOT-ENABLED PRODUCT MONITORING USING RFID Rebecca. In *Proceedings of the International Conferences on ICT, Society, and Human Beings 2016, Web Based Communities and Social Media 2016, Big Data Analytics, Data Mining and Computational Intelligence 2016 and Theory and Practice in Modern Computing 2016 - Part o*.
- Araújo, F. H. D. de, Santana, A. M., & Santos Neto, P. de A. dos. (2015). Uma Abordagem Influenciada por Pré-processamento para Aprendizagem do Processo de Regulação Médica. *Journal of Health Informatics*, *7*(1), 8–15. <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/317>
- Arivazhagan, B., Pandikumar, S., Sethupandian, S. B., & Subramanian, R. S. (2022). Pattern Discovery and Analysis of Customer Buying Behavior Using Association Rules Mining Algorithm in E-Commerce. *2022 1st International Conference on Electrical, Electronics, Information and Communication Technologies, ICEEICT 2022*.
<https://doi.org/10.1109/ICEEICT53079.2022.9768473>
- Avelar, E. A., Vinícius Dias Jordão, R., Maria Couto Ferreira, G., & Najela Ekaterina Ribeiro da Silva, B. (2021). Inteligência artificial como suporte a sistemas de controle e contabilidade gerencial.

- Sistemas & Gestão*, 16(1), 57–64. <https://doi.org/10.20985/1980-5160.2021.v16n1.1668>
- Bala, P. K. (2012). Improving inventory performance with clustering based demand forecasts. *Journal of Modelling in Management*, 7(1), 23–37. <https://doi.org/10.1108/17465661211208794>
- Balali, V., Ashouri Rad, A., & Golparvar-Fard, M. (2015). Detection, classification, and mapping of U.S. traffic signs using google street view images for roadway inventory management. *Visualization in Engineering*, 3(1), 1–18. <https://doi.org/10.1186/s40327-015-0027-1>
- Bandaru, S., Aslam, T., Ng, A. H. C., & Deb, K. (2015). Generalized higher-level automated innovization with application to inventory management. *European Journal of Operational Research*, 243(2), 480–496. <https://doi.org/10.1016/j.ejor.2014.11.015>
- Battaglia, D., & Bergamo, E. S. (2010). Análise de valor e engenharia de valor: uma ferramenta de redução de custos em um projeto. *P&D Em Engenharia de Produção*, January 2010, 102–115.
- Bellomarini, L., Fayzrakhmanov, R. R., Gottlob, G., Kravchenko, A., Laurenza, E., Nenov, Y., Reissfelder, S., Sallinger, E., Sherkhonov, E., Vahdati, S., & Wu, L. (2022). Data science with Vadalog: Knowledge Graphs with machine learning and reasoning in practice. *Future Generation Computer Systems*, 129, 407–422. <https://doi.org/10.1016/j.future.2021.10.021>
- Benabdellah, A. C., Benghabrit, A., & Bouhaddou, I. (2019). A survey of clustering algorithms for an industrial context. *Procedia Computer Science*, 148, 291–302. <https://doi.org/10.1016/j.procs.2019.01.022>
- Bevilacqua, M., Ciarapica, F. E., & Antomarioni, S. (2019). Lean principles for organizing items in an automated storage and retrieval system: An association rule mining – Based approach. *Management and Production Engineering Review*, 10(1), 29–36. <https://doi.org/10.24425/mper.2019.128241>
- Boehmke, B., Hazen, B., Boone, C. A., & Robinson, J. L. (2020). A data science and open source software approach to analytics for strategic sourcing. *International Journal of Information Management*, 54(July 2019), 102167. <https://doi.org/10.1016/j.ijinfomgt.2020.102167>
- Brock, L. G., & Davis, L. B. (2015). Estimating available supermarket commodities for food bank collection in the absence of information. *Expert Systems with Applications*, 42(7), 3450–3461. <https://doi.org/10.1016/j.eswa.2014.11.068>
- Carneiro, S. M. M. (1994). *Interdisciplinaridade: um novo paradigma do conhecimento?*
- Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A. V., & Rong, X. (2015). Data Mining for the Internet of Things: Literature Review and Challenges. *International Journal of Distributed Sensor Networks*, 2015. <https://doi.org/10.1155/2015/431047>
- Chen, J., Gusikhin, O., Finkenstaedt, W., & Liu, Y. N. (2019). Maintenance, repair, and operations parts inventory management in the era of industry 4.0. *IFAC-PapersOnLine*, 52(13), 171–176.

<https://doi.org/10.1016/j.ifacol.2019.11.171>

- Ciampi, F., Faraoni, M., Ballerini, J., & Meli, F. (2022). The co-evolutionary relationship between digitalization and organizational agility: Ongoing debates, theoretical developments and future research perspectives. *Technological Forecasting and Social Change*, 176(December 2021), 121383. <https://doi.org/10.1016/j.techfore.2021.121383>
- Cin, P. D., Joyce, S., Jurgens, J., & Tuteja, A. (2022). Earning Digital Trust : Decision-Making for Trustworthy Technologies. *World Economic Forum*, November.
- Coelho, L. C., Follmann, N., & Rodriguez, C. M. T. (2009). O impacto do compartilhamento de informações na redução do efeito chicote na cadeia de abastecimento. *Gestão & Produção*, 16(4), 571–583. <https://doi.org/10.1590/s0104-530x2009000400007>
- Conceição Farias do Egito Costa, S., Elisio de Figueiredo Lopes Lucena, A., & Paiva, W. de. (2023). Análise e predição do comportamento resiliente de solos com adição de cimento empregando-se técnicas de inteligência artificial. *Sistemas & Gestão*, 18(1), 54–64. <https://doi.org/10.20985/1980-5160.2023.v18n1.1856>
- Cordeiro, A. M., de Oliveira, G. M., Rentería, J. M., & Guimarães, C. A. (2007). Systematic review: A narrative review. *Revista Do Colegio Brasileiro de Cirurgioes*, 34(6), 428–431. <https://doi.org/10.1590/s0100-69912007000600012>
- Corrêa, H. L. (2019). *Administração de Cadeias de Suprimentos e Logística - Integração na Era da Indústria 4.0* (4th ed.). Administração de Cadeias de Suprimentos e Logística - Integração na Era da Indústria 4.0.
- Côrtes, S. D. C., Porcaro, R. M., & Lifschitz, S. (2002). Mineração de Dados – Funcionalidades, Técnicas e Abordagens. *PUC-Rio Informática*, 35. ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02_10_cortes.pdf
- Cosmo, B., Galeriani, T., Patricia Novakoski, F., & Martins Ricini, B. (2020). Carvão mineral. *Revista Agronomia Brasileira*, 4(1), 1–10. <https://doi.org/10.29372/rab202001>
- Costa, W. da S. da, Dallacort, R., Carvalho, M. A. C. de, & Santos, S. B. dos. (2022). Sistema Web para pré-processamento e análise de dados meteorológicos. *Revista Brasileira de Climatologia*, 30, 591–610. <https://doi.org/10.55761/abclima.v30i18.15079>
- Cunha Filho, O. A. L., & Rocha-Junior, J. B. (2022). *Encontrando Regras de Associação sem Especificar Suporte Mínimo e Confiança Mínima*. 168–174. https://doi.org/10.5753/sbbd_estendido.2022.21860
- D’Castro, R. J. (2020). *Pré-Processamento Para Mineração De Processos: Técnicas Para Simplificação Automática De Logs De Eventos*.
- Dawar, S. hybrid framework for mining high-utility itemsets in a sparse transaction database, Goyal,

- V., & Bera, D. (2017). A hybrid framework for mining high-utility itemsets in a sparse transaction database. *Applied Intelligence*, 47(3), 809–827. <https://doi.org/10.1007/s10489-017-0932-1>
- Demeter, K., Szász, L., & Kó, A. (2019). A text mining based overview of inventory research in the ISIR special issues 1994–2016. *International Journal of Production Economics*, 209(January 2017), 134–146. <https://doi.org/10.1016/j.ijpe.2018.06.006>
- Deng, C., & Liu, Y. (2021). A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection. *Wireless Communications and Mobile Computing*, 2021(1m). <https://doi.org/10.1155/2021/9969357>
- Diez-Olivan, A., Del Ser, J., Galar, D., & Sierra, B. (2019). Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Information Fusion*, 50(July 2018), 92–111. <https://doi.org/10.1016/j.inffus.2018.10.005>
- Diogo, R. A., Kolbe Junior, A., & Santos, N. (2019). a Transformação Digital E a Gestão Do Conhecimento: Contribuições Para a Melhoria Dos Processos Produtivos E Organizacionais. *P2P E Inovação*, 5(2), 154–175. <https://doi.org/10.21721/p2p.2019v5n2.p154-175>
- Doğan, G., & Karaca, O. (2020). A bibliometric analysis of the field of anesthesia during 2009–2018. *Brazilian Journal of Anesthesiology (English Edition)*, 70(2), 140–152. <https://doi.org/10.1016/j.bjane.2020.04.013>
- Dwivedi, Y. K., Hughes, D. L., Coombs, C., Constantiou, I., Duan, Y., Edwards, J. S., Gupta, B., Lal, B., Misra, S., Prashant, P., Raman, R., Rana, N. P., Sharma, S. K., & Upadhyay, N. (2020). Impact of COVID-19 pandemic on information management research and practice: Transforming education, work and life. *International Journal of Information Management*, 55(July), 102211. <https://doi.org/10.1016/j.ijinfomgt.2020.102211>
- ElMaraghy, H., Monostori, L., Schuh, G., & ElMaraghy, W. (2021). Evolution and future of manufacturing systems. *CIRP Annals*, 70(2), 635–658. <https://doi.org/10.1016/j.cirp.2021.05.008>
- Faceli, K., Lorena, A. C., Gama, J., & Carvalho, A. C. P. de L. F. de. (2011). *Inteligência artificial: uma abordagem de aprendizado de máquina*. LTC.
- Fahimnia, B., Pournader, M., Siemsen, E., Bendoly, E., & Wang, C. (2019). Behavioral Operations and Supply Chain Management—A Review and Literature Mapping. *Decision Sciences*, 50(6), 1127–1183. <https://doi.org/10.1111/deci.12369>
- Fávero, L. P., & Belfiore, P. (2024). *Manual de Análise de dados* (LTC - LIVROS (ed.)).
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). *From Data Mining to Knowledge Discovery in Databases*. AI Magazine. <https://doi.org/https://doi.org/10.1609/aimag.v17i3.1230>

- Fazenda, I. (2017). *Interdisciplinaridade: História, teoria e pesquisa* (Papirus Editora (ed.); 1ª).
- Felsberger, A., & Reiner, G. (2020). Sustainable industry 4.0 in production and operations management: A systematic literature review. *Sustainability (Switzerland)*, 12(19), 1–39.
<https://doi.org/10.3390/su12197982>
- Ferraço, C. E., & de Farias, I. M. S. (2021). Social Insertion: in search of senses and indicators for the evaluation of post-graduation in the Education area. *Ensaio*, 29(111), 420–440.
<https://doi.org/10.1590/S0104-40362020002802413>
- Ferreira, R. R. M., Paim, F. A. de P., Rodrigues, V. G. S., & Castro, G. S. A. (2020). Análise de cluster não supervisionado em R: agrupamento hierárquico. *Embrapa Territorial*, 0103–7811.
- Fleury, M. T. L., & Werlang, S. R. da C. (2017). Pesquisa aplicada: conceitos e abordagens. *Anuário de Pesquisa 2016-2017*, 10–15.
<https://bibliotecadigital.fgv.br/ojs/index.php/apgvpesquisa/article/view/72796>
- Forum, W. E. (2022a). *SDG 12: Responsible Consumption and Production*. World Economic Forum.
<https://intelligence.weforum.org/topics/a1G0X0000057N18UAE>
- Forum, W. E. (2022b). *The Digital Transformation of Business*. World Economic Forum.
- Fosso Wamba, S., Bawack, R. E., Guthrie, C., Queiroz, M. M., & Carillo, K. D. A. (2021). Are we preparing for a good AI society? A bibliometric review and research agenda. *Technological Forecasting and Social Change*, 164(December 2020).
<https://doi.org/10.1016/j.techfore.2020.120482>
- Gaviria-Marin, M., Merigó, J. M., & Baier-Fuentes, H. (2019). Knowledge management: A global examination based on bibliometric analysis. *Technological Forecasting and Social Change*, 140(May 2018), 194–220. <https://doi.org/10.1016/j.techfore.2018.07.006>
- George, B., Loo, J., & Jie, W. (2022). Recent advances and future trends on maintenance strategies and optimisation solution techniques for offshore sector. *Ocean Engineering*, 250(October 2021), 110986. <https://doi.org/10.1016/j.oceaneng.2022.110986>
- Giri, C., Jain, S., Zeng, X., & Bruniaux, P. (2019). A Detailed Review of Artificial Intelligence Applied in the Fashion and Apparel Industry. *IEEE Access*, 7, 95376–95396.
<https://doi.org/10.1109/ACCESS.2019.2928979>
- Goldschmidt, R., & Passos, E. (2005). *Data mining um guia prático*. Elsevier.
- Govindasamy, C., & Antonidoss, A. (2021). Comparative Analysis of Colliding Bodies Optimization-based Inventory Management over other Heuristics: Cloud with blockchain criteria. *Proceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineering*, 244–251. <https://doi.org/10.1109/Confluence51648.2021.9377157>
- Granillo-Macías, R. (2020). INVENTORY MANAGEMENT AND LOGISTICS OPTIMIZATION: A DATA

MINING PRACTICAL APPROACH. *Logforum*, 16(4), 535–547.

<https://doi.org/10.17270/J.LOG.2020.512>

Groover, M. P. (2017). *Fundamentos da Moderna Manufatura Versão SI* (5ª Edição). Grupo GEN.

Gunduz, M. A., Demir, S., & Paksoy, T. (2021). Matching functions of supply chain management with smart and sustainable Tools: A novel hybrid BWM-QFD based method. *Computers and Industrial Engineering*, 162(September), 107676. <https://doi.org/10.1016/j.cie.2021.107676>

Guo, X., Liu, C., Xu, W., Yuan, H., & Wang, M. (2014). A prediction-based inventory optimization using data mining models. *Proceedings - 2014 7th International Joint Conference on Computational Sciences and Optimization, CSO 2014*, 611–615.

<https://doi.org/10.1109/CSO.2014.118>

Hamister, J. W., Magazine, M. J., & Polak, G. G. (2018). Integrating Analytics Through the Big Data Information Chain: A Case From Supply Chain Management. *Journal of Business Logistics*, 39(3), 220–230. <https://doi.org/10.1111/jbl.12192>

Han, H., & Trimi, S. (2022). Towards a data science platform for improving SME collaboration through Industry 4.0 technologies. *Technological Forecasting and Social Change*, 174(January 2021), 121242. <https://doi.org/10.1016/j.techfore.2021.121242>

Hassija, V., Chamola, V., Gupta, V., Jain, S., & Guizani, N. (2021). A Survey on Supply Chain Security: Application Areas, Security Threats, and Solution Architectures. *IEEE Internet of Things Journal*, 8(8), 6222–6246. <https://doi.org/10.1109/JIOT.2020.3025775>

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154, 72–80. <https://doi.org/10.1016/j.ijpe.2014.04.018>

Hopp, W. J., & Spearman, M. L. (2012). *A Ciência da Fábrica* (Bookman (ed.); 3ª).

Hora, H. R. M. da, Monteiro, G. T. R., & Arica, J. (2010). Confiabilidade em Questionários para Qualidade: Um Estudo com o Coeficiente Alfa de Cronbach. *Produto & Produção*, 11(2), 85–103.

Huang, B., Gan, W., & Li, Z. (2021). Application of Medical Material Inventory Model under Deep Learning in Supply Planning of Public Emergency. *IEEE Access*, 9, 44128–44138. <https://doi.org/10.1109/ACCESS.2021.3057869>

Huang, L., Wu, C., Wang, B., & Ouyang, Q. (2018). Big-data-driven safety decision-making: A conceptual framework and its influencing factors. *Safety Science*, 109(January), 46–56. <https://doi.org/10.1016/j.ssci.2018.05.012>

Hussain, M., Javed, W., Hakeem, O., Yousafzai, A., Younas, A., Awan, M. J., Nobanee, H., & Zain, A.

- M. (2021). Blockchain-Based IoT Devices in Supply Chain Management: A Systematic Literature Review. *Sustainability (Switzerland)*, 13(24), 1–23. <https://doi.org/10.3390/su132413646>
- Ijadi, A., Kavian, A., Khalilzadeh, M., & Brauers, W. K. M. (2018). CLUS-MCDA : A novel framework based on cluster analysis and multiple criteria decision theory in a supplier selection problem. *Computers & Industrial Engineering*, 118(November 2017), 409–422. <https://doi.org/10.1016/j.cie.2018.03.011>
- Initiative, H. D. S. (2022). *Data Science*. World Economic Forum.
- Institution, S. (2022). *Diversity and Inclusion: Inclusive Education Design*. World Economic Forum. <https://intelligence.weforum.org/topics/a1G0X0000057IniUAE/key-issues/a1G0X0000057N7FUAU>
- Jae-Young Oh Yong-Geon Lee, G. K. (2020). *Improvement of Solar Power Forecasting Using Interpretation of Artificial Intelligence* (Vol. 69, Issue 7, p. 1111).
- Jeble, S., Dubey, R., Childe, S. J., Papadopoulos, T., Roubaud, D., & Prakash, A. (2018). Impact of big data and predictive analytics capability on supply chain sustainability. *International Journal of Logistics Management*, 29(2), 513–538. <https://doi.org/10.1108/IJLM-05-2017-0134>
- Jung, C. F. (2004). *Metodologia para Pesquisa & Desenvolvimento* (Axcel (ed.)).
- Júnior, C. A. de O. M., & Batista, M. C. (2023). *Metodologia da pesquisa em educação e ensino de ciências*.
- Kassa, A. O. (2015). *Value Analysis and Engineering Reengineered* (CRC Press (ed.)).
- Kim, S., Seo, H., & Lee, E. C. (2024). Advanced Anomaly Detection in Manufacturing Processes: Leveraging Feature Value Analysis for Normalizing Anomalous Data. *Electronics (Switzerland)*, 13(7). <https://doi.org/10.3390/electronics13071384>
- Kisilevich, S., Keim, D., & Rokach, L. (2013). A GIS-based decision support system for hotel room rate estimation and temporal price prediction: The hotel brokers' context. *Decision Support Systems*, 54(2), 1119–1133. <https://doi.org/10.1016/j.dss.2012.10.038>
- Kruger, J. M. (2023). *Metodologia da Pesquisa em Administração: Em Linguagem descomplicada* (BAGAI (ed.)).
- Kumar, A., Shankar, R., & Radi, N. (2020). A big data driven framework for demand-driven forecasting with effects of marketing-mix variables. *Industrial Marketing Management*, 90(May 2019), 493–507. <https://doi.org/10.1016/j.indmarman.2019.05.003>
- Kumar, H., Singh, M. K., Gupta, M. P., & Madaan, J. (2020). Moving towards smart cities: Solutions that lead to the Smart City Transformation Framework. *Technological Forecasting and Social Change*, 153(April 2018), 119281. <https://doi.org/10.1016/j.techfore.2018.04.024>
- Kumar, V., Steinbach, M., & Tan, P.-N. (2009). *Introdução ao Data Mining. Mineração de Dados*.

Ciência Moderna.

Labucay, I. (2022). Is There a Smart Sustainability Transition in Manufacturing? Tracking Externalities in Machine Tools over Three Decades. *Sustainability (Switzerland)*, 14(2).

<https://doi.org/10.3390/su14020838>

Lee, C. H., Liu, C. L., Trappey, A. J. C., Mo, J. P. T., & Desouza, K. C. (2021). Understanding digital transformation in advanced manufacturing and engineering: A bibliometric analysis, topic modeling and research trend discovery. *Advanced Engineering Informatics*, 50(August), 101428.

<https://doi.org/10.1016/j.aei.2021.101428>

Lee, C. K. M., Na, C. M., & Kit, N. C. (2015). IoT-based asset management system for healthcare-related industries. *International Journal of Engineering Business Management*, 7.

<https://doi.org/10.5772/61821>

Lee, C. K. M., & Palaniappan, S. (2014). Effective asset management for hospitals with RFID. *2014 IEEE International Technology Management Conference, ITMC 2014*.

<https://doi.org/10.1109/ITMC.2014.6918596>

Li, L., Shen, C., Wang, L., Zheng, L., Jiang, Y., Tang, L., Li, H., Zhang, L., Zeng, C., Li, T., Tang, J., & Liu, D. (2014). *iMiner: Mining Inventory Data for Intelligent Management*. 2057–2059.

<https://doi.org/10.1145/2661829.2661848>

Li, Y., Jiang, W., Yang, L., & Wu, T. (2018). On neural networks and learning systems for business computing. *Neurocomputing*, 275, 1150–1159.

<https://doi.org/10.1016/j.neucom.2017.09.054>
Librelotto, S. R., & Mozzaquatro, P. M. (2013). Análise dos algoritmos de mineração J48 e apriori aplicados na detecção de indicadores da qualidade de vida e saúde. *Revista Interdisciplinar de Ensino, Pesquisa e Extensão*, 1(1), 26–37.

Lima, T. C. S. de, & Mioto, R. C. T. (2007). Procedimentos metodológicos na construção do conhecimento científico: a pesquisa bibliográfica. *Revista Katálysis*, 10(spe), 37–45.

<https://doi.org/10.1590/s1414-49802007000300004>

Lins, B. F. E. (1993). *Ferramentas básicas da qualidade*.

Liou, J. J. H., Chang, M. H., Lo, H. W., & Hsu, M. H. (2021). Application of an MCDM model with data mining techniques for green supplier evaluation and selection. *Applied Soft Computing*, 109, 107534.

<https://doi.org/10.1016/j.asoc.2021.107534>

López-Soto, D., Angel-Bello, F., Yacout, S., & Alvarez, A. (2017). A multi-start algorithm to design a multi-class classifier for a multi-criteria ABC inventory classification problem. *Expert Systems with Applications*, 81, 12–21.

<https://doi.org/10.1016/j.eswa.2017.02.048>

Maass, W., & Storey, V. C. (2021). Pairing conceptual modeling with machine learning. *Data and Knowledge Engineering*, 134(January), 101909.

<https://doi.org/10.1016/j.datak.2021.101909>

- Mariani, M. M., & Fosso Wamba, S. (2020). Exploring how consumer goods companies innovate in the digital age: The role of big data analytics companies. *Journal of Business Research*, 121(February), 338–352. <https://doi.org/10.1016/j.jbusres.2020.09.012>
- Mariano, D. C. B., Marques, L. torres, Silva, M. S., & Aleixo, J. F. meireles. (2020). *Data mining*. [Digitalo Local da Editora]: Grupo A.
- Miles, L. (2015). *Techniques of Value Analysis and Engineering* (3rd Editio). Lawrence D. Miles Value Foundation.
- Mineral, B. (2021). *Carbonifera Rio Deserto | AS MAIORES EMPRESAS DO SETOR MINERAL*. <https://www.brasilmineral.com.br/maiores/riodeserto>
- Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. (2022). IoT, Big Data, and Artificial Intelligence in Agriculture and Food Industry. *IEEE Internet of Things Journal*, 9(9), 6305–6324. <https://doi.org/10.1109/JIOT.2020.2998584>
- Morenttin, P. A., & Bussab, W. de O. (2017). *Estatística básica* (Saraiva (ed.); 9ª edição).
- Muhuri, P. K., Shukla, A. K., & Abraham, A. (2019a). Industry 4.0: A bibliometric analysis and detailed overview. *Engineering Applications of Artificial Intelligence*, 78(December 2018), 218–235. <https://doi.org/10.1016/j.engappai.2018.11.007>
- Muhuri, P. K., Shukla, A. K., & Abraham, A. (2019b). Industry 4.0: A bibliometric analysis and detailed overview. *Engineering Applications of Artificial Intelligence*, 78(November 2017), 218–235. <https://doi.org/10.1016/j.engappai.2018.11.007>
- Namir, K., Labriji, H., & Ben Lahmar, E. H. (2021). Decision Support Tool for Dynamic Inventory Management using Machine Learning, Time Series and Combinatorial Optimization. *Procedia Computer Science*, 198(2021), 423–428. <https://doi.org/10.1016/j.procs.2021.12.264>
- Ng, T. C., Lau, S. Y., Ghobakhloo, M., & Fathi, M. (2022). The Application of Industry 4 . 0 Technological Constituents for Sustainable Manufacturing : A Content-Centric Review. *Sustainability Review*. <https://doi.org/https://doi.org/10.3390/su14074327> Academic
- Nomelini, J., Rezende, S. O., Yamamoto, C. H., Bezerra, L. A. F., Vozi, P. A., & Lôbo, R. B. (2010). Emprego de regras de associação para extração de padrões mercadológicos de touros Nelore com avaliação genética. *Revista Brasileira de Zootecnia*, 39(12), 2632–2639. <https://doi.org/10.1590/S1516-35982010001200011>
- Núñez-Merino, M., Maqueira-Marín, J., Moyano-Fuentes, J., & Castano-Moraga, C. (2022). Industry 4 . 0 and supply chain . A Systematic Science Mapping analysis. *Technological Forecasting & Social Change*, 181(June). <https://doi.org/10.1016/j.techfore.2022.121788>
- Omar, I. A., Jayaraman, R., Salah, K., Debe, M., & Omar, M. (2020). Enhancing vendor managed inventory supply chain operations using blockchain smart contracts. *IEEE Access*, 8, 182704–

182719. <https://doi.org/10.1109/ACCESS.2020.3028031>

Papanagnou, C. I., & Matthews-Amune, O. (2018). Coping with demand volatility in retail pharmacies with the aid of big data exploration. *Computers and Operations Research*, 98, 343–354.

<https://doi.org/10.1016/j.cor.2017.08.009>

Pátaro, R. F., & Bovo, M. C. (2017). A interdisciplinaridade como possibilidade de diálogo e trabalho coletivo no campo da pesquisa e da educação. *Revista NUPEM*, 4(6), 45–63.

<https://doi.org/10.33871/nupem.v4i6.96>

Pejic-Bach, M., Bertancel, T., Meško, M., & Krstić, Ž. (2020). Text mining of industry 4.0 job advertisements. *International Journal of Information Management*, 50(August 2019), 416–431.

<https://doi.org/10.1016/j.ijinfomgt.2019.07.014>

Philip Chen, C. L., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314–347.

<https://doi.org/10.1016/j.ins.2014.01.015>

Philippi, A., & Fernandes, V. (2021). Ciência e tecnologia à luz da interdisciplinaridade. *Tecendo Redes e Conexões Para a Sustentabilidade*, 189–200.

<https://atual.sistemafaep.org.br/boletim/tecendo-redes-e-conexoes-para-a-sustentabilidade/>

Pinto, A. C., & Andrade, J. B. de. (1999). Fator de impacto de revistas científicas: qual o significado deste parâmetro? *Química Nova*, 22(3), 448–453. <https://doi.org/10.1590/s0100-40421999000300026>

Piran, F. S., Lacerda, D. P., & Camargo, L. F. R. (2018). *Análise e Gestão da Eficiência: Aplicação em Sistemas Produtivos de Bens e Serviços*. Elsevier.

Porter, M. (2005). *Estratégia Competitiva - Técnicas Para Análise de Indústrias e da Concorrência* (G. Atlas (ed.); 1ª edição).

Portugal, M., Branca, S., & Rodrigues, M. (2011). Dados de medida de fator de impacto das revistas científicas. *Revista de Enfermagem Referência, III Série*(nº 5), 211–215.

<https://doi.org/10.12707/riii11ui4>

Qazi, A., & Al-Mhdawi, M. K. S. (2023). Exploring dependencies among global environmental, socioeconomic, and technological risks. *Environmental Impact Assessment Review*,

98(September 2022), 106912. <https://doi.org/10.1016/j.eiar.2022.106912>

Queiroz, M. M., & Pereira, S. C. F. (2019). INTENÇÃO DE ADOÇÃO DE BIG DATA NA CADEIA DE SUPRIMENTOS: UMA PERSPECTIVA BRASILEIRA. *Revista de Administração de Empresas*, 59(6), 389–

401. <https://doi.org/10.1590/s0034-759020190605>

Queiroz, M. M., Pereira, S. C. F., Telles, R., & Machado, M. C. (2019). Industry 4.0 and digital supply chain capabilities: A framework for understanding digitalisation challenges and opportunities.

Benchmarking, 28(5), 1761–1782. <https://doi.org/10.1108/BIJ-12-2018-0435>

Ramalho, H. M. de C. S. de C. em L. P. uma A. do A. A. priori de A. de M. para o E. da P. B., Almeida, A. T. C. de, & Fraga, A. A. (2020). Detecção de Casos Suspeitos de Conluio em Licitações Públicas: uma Aplicação do Algoritmo A priori de Aprendizado de Máquina para o Estado da Paraíba. *Teoria e Prática Em Administração*, 10(2), 5–22. <https://doi.org/10.21714/2238-104x2020v10i2-51526>

Rautenberg, S., & Carmo, P. R. V. do. (2019). BIG DATA E CIÊNCIA DE DADOS: COMPLEMENTARIEDADE CONCEITUAL NO PROCESSO DE TOMADA DE DECISÃO. *Brazilian Journal of Information Science: Research Trends*, 13(10.36311/1981-1640.2019.v13n1.06.p56), 56–67. <https://doi.org/10.36311/1981-1640.2019.v13n1.06.p56>

Raynaut, C. (2014). OS DESAFIOS CONTEMPORÂNEOS DA PRODUÇÃO DO CONHECIMENTO: O APELO PARA INTERDISCIPLINARIDADE. *Interthesis*, 11, 103–111.

Révillion, A. S. P. (2003). A Utilização de Pesquisas Exploratórias na Área de Marketing. *Revista Interdisciplinar de Marketing*, 21–37.

Ridzuan, F., & Wan Zainon, W. M. N. (2021). Diagnostic analysis for outlier detection in big data analytics. *Procedia Computer Science*, 197, 685–692. <https://doi.org/10.1016/j.procs.2021.12.189>

Rodrigues, A., Rhodes, B., Cambinda, D., Lourenço, P., Mota, R., Aires, T., Sofia, A., & Rodrigues, O. (2021). *Classification of Handwritten Digits*. 9(4), 44–57.

Rogers, D. L., & Serra, A. C. da C. (2017). *Transformação Digital: repensando o seu negócio para a era digital*. Autêntica Business; 1ª edição.

Roßmann, B., Canzaniello, A., von der Gracht, H., & Hartmann, E. (2018). The future and social impact of Big Data Analytics in Supply Chain Management: Results from a Delphi study. *Technological Forecasting and Social Change*, 130(November 2017), 135–149. <https://doi.org/10.1016/j.techfore.2017.10.005>

Rudnitckaia, J., Venkatachalam, H. S., Essmann, R., Hruska, T., & Colombo, A. W. (2022). Screening Process Mining and Value Stream Techniques on Industrial Manufacturing Processes: Process Modelling and Bottleneck Analysis. *IEEE Access*, 10, 24203–24214. <https://doi.org/10.1109/ACCESS.2022.3152211>

Sá, G. de, Michels, C., Negrelli, A. Q., Gesuino, D. B., Martins, P. J., Bristot, V. M., Filho, L. P. G., Madeira, K., & Yamaguchi, C. K. (2020). Analytic Hierarchy Process in Production Engineering: A Bibliometry Analysis. *International Journal of Advanced Engineering Research and Science*, 7(7), 209–219. <https://doi.org/10.22161/ijaers.77.25>

Sajid, S., Haleem, A., Bahl, S., Javaid, M., Goyal, T., & Mittal, M. (2021). Data science applications for

- predictive maintenance and materials science in context to Industry 4.0. *Materials Today: Proceedings*, 45, 4898–4905. <https://doi.org/10.1016/j.matpr.2021.01.357>
- Salamai, A., Hussain, O. K., Saberi, M., Chang, E., & Hussain, F. K. (2019). Highlighting the Importance of Considering the Impacts of Both External and Internal Risk Factors on Operational Parameters to Improve Supply Chain Risk Management. *IEEE Access*, 7, 49297–49315. <https://doi.org/10.1109/ACCESS.2019.2902191>
- Sangkhasuk, R., Rianthong, S., & Arunrat, K. (2018). Automation managed inventory system (AMI) creation for OTOP distribution center. *ACM International Conference Proceeding Series*, 112–116. <https://doi.org/10.1145/3206098.3206119>
- Santos, A. dos, & Hashimoto, H. (2003). Demonstração do valor adicionado: algumas considerações sobre carga tributária. *Revista de Administração Da Universidade de São Paulo*, 38(2), 153–164.
- Saritha, B., Bonagiri, R., & Deepika, R. (2021). Open source technologies in data science and big data analytics. *Materials Today: Proceedings*, xxxx, 1–5. <https://doi.org/10.1016/j.matpr.2021.01.610>
- Sarmah, S. P., & Moharana, U. C. (2017). Spare Parts Inventory Management Literature and Direction Towards the Use of Data Mining Technique: A Review. *Handbook of Research on Promoting Business Process Improvement Through Inventory Control Techniques*.
- Schmitt, M. (2023). International Journal of Information Management Data Insights Deep learning in business analytics : A clash of expectations and reality. *International Journal of Information Management Data Insights*, 3(1), 100146. <https://doi.org/10.1016/j.ijime.2022.100146>
- Schoenherr, T., & Speier-Pero, C. (2015). Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential. *Journal of Business Logistics*, 36(1), 120–132. <https://doi.org/10.1111/jbl.12082>
- Schonhorst, G. B. (2010). *Mineração de Regras de Associação Aplicada à Modelagem dos Dados Transacionais de um Supermercado*. <http://goo.gl/0ZQIPj>
- Schwandt, D. R., & Marquardt, M. J. (1999). *Organizational Learning*. CRC Press;
- Seeger, P. M., Yahouni, Z., & Yahouni, Z. (2022). Literature review on using data mining in production planning and scheduling within the context of cyber physical systems. *Journal of Industrial Information Integration*. <https://doi.org/10.1016/j.jii.2022.100371>
- Senthilkumaran, U., Manikandan, N., & Senthilkumar, M. (2016). Role of data mining on pharmaceutical industry-a survey. *International Journal of Pharmacy and Technology*, 8(3), 16100–16106.
- Sharma, S. (2016). Expanded cloud plumes hiding Big Data ecosystem. *Future Generation Computer Systems*, 59, 63–92. <https://doi.org/10.1016/j.future.2016.01.003>

- SIESESC. (2022). *CARVÃO MINERAL ESTÁ ASSOCIADO AO DESENVOLVIMENTO ECONÔMICO DE MUNICÍPIOS*. SIESESC.
- Silva, B. W. (2019). *Gestão de Estoques: Planejamento, Execução e Controle* (BWS CONSULTORIA (ed.)).
- Silva, L. A. da, Peres, S. M., & Boscaroli, C. (2016). *Introdução a mineração de dados com aplicação em R*. Elsevier.
- Slack, N., Chambers, S., Johnston, R., & Betts, A. (2013). *Gerenciamento de Operações e de Processos* (2nd ed.). Bookman.
- SOARES, G. O. (2022). A complexidade econômica e a distribuição de renda na economia catarinense entre 2000- 2018. *Material Safety Data Sheet*, 33(1), 1–12.
http://www2.warwick.ac.uk/fac/sci/whri/research/mushroomresearch/mushroomquality/fungi/environment%0Ahttps://us.vwr.com/assetsvc/asset/en_US/id/16490607/contents%0Ahttp://www.hse.gov.uk/pubns/indg373hp.pdf
- Sousa, A. S. de, Oliveira, G. S. de, & Alves, L. H. (2021). A Pesquisa Bibliográfica: Princípios E Fundamentos. *Cadernos Da FUCAMP*, 20(43), 64–83.
- Sousa, Ó., & Rodrigues, D. R. A. M. (2022). *INCLUSÃO DIGITAL : UMA REFLEXÃO SOBRE AS PRÁTICAS PARA ATINGIR A INCLUSÃO SOCIAL*.
- Sousa, P. H. T. de O. (2016). *Classificação utilizando técnicas de aprendizado estatístico: Estudo de caso*.
- Souza, G. M. de, Abreu, V. S. de, & D’Agosto, M. de A. (2021). Indústria 4.0 Aplicada À Gestão Da Cadeia De Suprimentos: Uma Revisão Da Literatura. *Brazilian Journal of Production Engineering - BJPE*, 128–142. <https://doi.org/10.47456/bjpe.v7i2.35009>
- Srujana, K. S., Sukruta Nagraj Kashyap, Shrividhiya, G., Gururaj, C., & Induja, K. S. (2022). Supply Chain Based Demand Analysis of Different Deep Learning Methodologies for Effective Covid-19 Detection. *Springer International Publishing*.
- Stadnicka, D., & Ratnayake, R. M. C. (2017). Enhancing Aircraft Maintenance Services: A VSM Based Case Study. *Procedia Engineering*, 182, 665–672. <https://doi.org/10.1016/j.proeng.2017.03.177>
- Stefanovic, N. (2015). Collaborative predictive business intelligence model for spare parts inventory replenishment. *Computer Science and Information Systems*, 12(3), 911–930.
<https://doi.org/10.2298/CSIS141101034S>
- Strauhs, F. do R., Pietrovski, E. F., Santos, G. D., Carvalho, H. G. de, Pimenta, R. B., & Penteado, R. S. (2012). *Gestão do conhecimento nas organizações*. Aymarã Educação.
- Sucena, M. P., & Cury, M. V. Q. (2023). Artificial Intelligence Applied to Assess Perceptions of the Quality of E-Commerce Logistics: Case Study of Rio de Janeiro. *Brazilian Business Review*, 20(4).

<https://doi.org/10.15728/bbr.2021.1177.en>

- Sundarakani, B., Ajaykumar, A., & Gunasekaran, A. (2021). Big data driven supply chain design and applications for blockchain: An action research using case study approach. *Omega (United Kingdom)*, 102. <https://doi.org/10.1016/j.omega.2021.102452>
- Tabile, P. M., Bernhard, T. W., Dihel, D., & Koepp, J. (2015). A importância do fluxograma para o trabalho da saúde da família na visão do projeto PET-Saúde. *Revista Eletrônica de Gestão & Saúde*, 06(1), 680–690.
- Tilimbe, J. (2019). Ethical Implications of Predictive Risk Intelligence. *Orbit Journal*, 2(2), 1–28. <https://doi.org/10.29297/orbit.v2i2.112>
- Tirkel, I. (2013). Forecasting flow time in semiconductor manufacturing using knowledge discovery in databases. *INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH*.
- Tomaz, P. L. (2022). ANÁLISE DE POSICIONAMENTOS PÚBLICOS SOBRE A LICENÇA SOCIAL PARA OPERAR DA MINA GUAÍBA. 1–23.
- Tran-Dang, H., & Kim, D. S. (2021). The Physical Internet in the Era of Digital Transformation: Perspectives and Open Issues. *IEEE Access*, 9, 164613–164631. <https://doi.org/10.1109/ACCESS.2021.3131562>
- Tsou, C. M. (2013). On the strategy of supply chain collaboration based on dynamic inventory target level management: A theory of constraint perspective. *Applied Mathematical Modelling*, 37(7), 5204–5214. <https://doi.org/10.1016/j.apm.2012.10.031>
- Tsui, K. L., Zhao, Y., & Wang, D. (2019). Big data opportunities: System health monitoring and management. *IEEE Access*, 7, 68853–68867. <https://doi.org/10.1109/ACCESS.2019.2917891>
- Vargas, R. V., & Abba, W. (2018). *Análise de Valor Agregado* (Brasport (ed.); 7º).
- Vasconcelos, L. M. R. de, & Carvalho, C. L. de. (2004). Aplicação de Regras de Associação para Mineração de Dados na Web. *Instituto de Informática Da Universidade Federal de Goiás*, 20. <http://goo.gl/Htb9xp>
- Ventura, M. M. (2007). O Estudo de Caso como Modalidade de Pesquisa. *Rev SOCERJ*, 20(5), 383–386. http://www.polo.unisc.br/portal/upload/com_arquivo/o_estudo_de_caso_como_modalidade_de_pesquisa.pdf
- Veras, C. M. T., & Martins, M. S. (1994). A confiabilidade dos dados nos formulários de Autorização de Internação Hospitalar (AIH), Rio de Janeiro, Brasil. *Cadernos de Saúde Pública*, 10(3), 339–355. <https://doi.org/10.1590/s0102-311x1994000300014>
- Voese, S. B., & Mello, R. J. G. (2013). Análise bibliométrica sobre gestão estratégica de custos no Congresso Brasileiro de Custos: aplicação da lei de Lotka. *Revista Capital Científico - Eletrônica*,

11(1), 1–19. <http://www.spell.org.br/documentos/ver/33063/analise-bibliometrica-sobre-gestao-estrategica-de-custos-no-congresso-brasileiro-de-custos--aplicacao-da-lei-de-lotka>

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84. <https://doi.org/10.1111/jbl.12010>

Wang, Y., Geng, X., Zhang, F., & Ruan, J. (2018). An Immune Genetic Algorithm for Multi-Echelon Inventory Cost Control of IOT Based Supply Chains. *IEEE Access*, 6, 8547–8555. <https://doi.org/10.1109/ACCESS.2018.2799306>

Wijffels, L., Giannikas, V., Woodall, P., McFarlane, D., & Lu, W. (2016). An enhanced cycle counting approach utilising historical inventory data. *IFAC-PapersOnLine*, 49(12), 1347–1352. <https://doi.org/10.1016/j.ifacol.2016.07.748>

Yobero, C. (2018). *K-Means Clustering Tutorial*. RPubs.

Zhang, Y., Montenegro-Marin, C. E., & Díaz, V. G. (2021). Holistic cognitive conflict chain management framework in supply chain management. *Environmental Impact Assessment Review*, 88(February), 106564. <https://doi.org/10.1016/j.eiar.2021.106564>