

UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS

Letícia Coelho Learth

Revisão sistemática da literatura sobre métodos de *machine learning* aplicados  
para apoiar a gestão de fornecedores

São Carlos  
2024

Letícia Coelho Learth

Revisão sistemática da literatura sobre métodos de *machine learning* aplicados  
para apoiar a gestão de fornecedores

Monografia apresentada ao Curso de  
Engenharia de Produção, da Escola de  
Engenharia de São Carlos da Universidade de  
São Paulo, como parte dos requisitos para  
obtenção do título de Engenheira de Produção.

Orientador: Prof. Dr. Luiz Cesar Ribeiro  
Carpinetti

VERSÃO CORRIGIDA

São Carlos

2024

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP com os dados inseridos pelo(a) autor(a).

Coelho Learth, Letícia  
C438r      Revisão sistemática da literatura sobre métodos de machine learning aplicados para apoiar a gestão de fornecedores / Letícia Coelho Learth; orientador Luiz Cesar Ribeiro Carpinetti. São Carlos, 2024.

Monografia (Graduação em Engenharia de Produção) -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2024.

1. Gestão de fornecedores. 2. machine learning. 3. supplier. I. Título.

Eduardo Graziosi Silva - CRB - 8/8907



## DEDICATÓRIA

*Este trabalho é dedicado aos meus pais Otávio  
Augusto Learth Cunha e Ana Carla de Aragão  
Coelho, aos meus irmãos e amigos que fiz ao longo  
da minha jornada, um sonho sonhado junto se torna  
realidade.*

## FOLHA DE APROVAÇÃO

**Candidato:** Letícia Coelho Learth

**Título do TCC:** REVISÃO SISTEMÁTICA DA LITERATURA SOBRE MÉTODOS DE MACHINE LEARNING APLICADOS PARA APOIAR A GESTÃO DE FORNECEDORES

**Data de defesa:** 13/12/2024

Comissão Julgadora	Resultado
Professor Titular Luiz César Ribeiro Carpinetti (orientador)	APROVADO
Instituição: EESC - SEP	
Professor Doutor Lucas Gabriel Zanon	APROVADO
Instituição: EESC - SEP	
Pesquisador Rafael Ferro Munhoz Arantes	APROVADO
Instituição: EESC - SEP	

Presidente da Banca: Professor Titular Luiz César Ribeiro Carpinetti

## RESUMO

LEARTH, L. C. **Revisão sistemática da literatura sobre métodos de machine learning aplicados para apoiar a gestão de fornecedores.** 2024. Monografia (Trabalho de Conclusão de Curso) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2024.

No contexto atual de cadeias de suprimentos cada vez mais complexas e dinâmicas, a gestão de fornecedores (*Supplier Relationship Management - SRM*) desempenha um papel estratégico na competitividade das empresas e nesse contexto a aplicação de técnicas de machine learning surge como uma abordagem promissora para otimizar os processos envolvidos, como seleção, segmentação, avaliação e desenvolvimento de fornecedores, oferecendo maior precisão e eficiência na tomada de decisões. Esta pesquisa apresenta uma revisão sistemática da literatura sobre o uso de técnicas de aprendizado de máquina (*machine learning*) no suporte à gestão de fornecedores. O estudo seguiu uma metodologia de revisão, contemplando a definição de *strings* de busca, seleção de artigos relevantes e análise detalhada dos conteúdos. Foram analisados 67 artigos classificados como alinhados ao tema em análise, considerando aspectos como algoritmos utilizados, objetivos, desafios e benefícios reportados. Entre os métodos identificados, destacaram-se o uso de redes neurais, árvores de decisão, e técnicas de clusterização, sendo as abordagens supervisionadas as mais frequentes (64%). Os resultados evidenciaram que a etapa de seleção de fornecedores é também a mais estudada (49%), enquanto desenvolvimento e pré-qualificação receberam menos atenção. Além disso, observou-se predominância de aplicações com bases de dados reais (76%). Conclui-se que, apesar dos avanços identificados, desafios como a interpretabilidade dos modelos e a integração de dados heterogêneos ainda são considerados fatores limitantes e que processos de gestão de fornecedores com potencial de exploração são desenvolvimento de fornecedores e pré-qualificação.

Palavras-chave: 1. Gestão de fornecedores 2. *Machine Learning* 3. *Supplier*



## ABSTRACT

**LEARTH, L. C. Systematic Literature Review on Machine Learning Methods Applied to Support Supplier Management.** 2024. Monografia (Trabalho de Conclusão de Curso) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2024.

In the current context of increasingly complex and dynamic supply chains, Supplier Relationship Management (SRM) plays a strategic role in companies' competitiveness. In this scenario, the application of machine learning techniques emerges as a promising approach to optimize the involved processes, such as supplier selection, segmentation, evaluation, and development, offering greater precision and efficiency in decision-making. This research presents a systematic literature review on the use of machine learning techniques to support supplier management. The study followed a review methodology, including the definition of search strings, selection of relevant articles, and detailed content analysis. A total of 67 articles aligned with the research topic were analyzed, considering aspects such as the algorithms used, objectives, challenges, and reported benefits. Among the identified methods, the use of neural networks, decision trees, and clustering techniques stood out, with supervised approaches being the most frequent (64%). The results showed that the supplier selection stage is also the most studied (49%), while development and pre-qualification received less attention. Furthermore, there was a predominance of applications using real databases (76%). It is concluded that, despite the identified advances, challenges such as model interpretability and the integration of heterogeneous data are still considered limiting factors. Supplier management processes with potential for further exploration include supplier development and pre-qualification.

Keywords: 1. Supplier Management 2. Machine Learning 3. Supplier



## **LISTA DE FIGURAS**

Figura 1 – Estrutura SRM.....	32
Figura 2 – Matriz de avaliação de materiais e fornecedores.....	33
Figura 3 – <i>Framework</i> conceitual das atividades SEM.....	34
Figura 4 – Esquemático das etapas de realização da pesquisa.....	43
Figura 5 – <i>Strings</i> de busca dos artigos.....	45
Figura 6 – Esquematização da extração dos artigos.....	47
Figura 7 – Volume de publicações obtidos ao longo da pesquisa.....	49
Figura 8 – Distribuição do número de publicações ao longo dos anos.....	50



## LISTA DE TABELAS

Tabela 1 – Países que publicaram sobre o tema e seus números de publicações.....	51
Tabela 2 – Instituições que mais publicaram sobre o tema.....	52
Tabela 3 – Recorrência de uso de cada técnica de <i>Machine Learning</i> .....	53
Tabela 4 – Segmentação por abordagens supervisionadas ou não supervisionadas.....	55
Tabela 5 – Tipos de data set utilizados.....	56
Tabela 6 – Recorrência de tarefas de <i>Machine Learning</i> .....	56
Tabela 7 – Distribuição das aplicações em relação à etapa do SEM.....	57
Tabela 8 – Distribuição dos quantidade de referente ao uso do tipo de linguagem e abordagens de <i>Machine Learning</i> .....	58
Tabela 9 – Distribuição dos artigos sobre as linguagens utilizadas e as etapas do processo de SRM.....	58
Tabela 10 – Recorrência de uso de cada técnica MCDM.....	59
Tabela 11 – Distribuição do uso das técnicas MCDM.....	60
Tabela 12 – Distribuição dos artigos entre as abordagens de machine learning e as etapas do processo SEM.....	61
Tabela 13 – Divisão dos artigos entre etapas do processo SRM e tarefas de machine learning com quantidade.....	75



## LISTA DE ABREVIATURAS E SIGLAS

SRM	Supplier Relationship Management
ERP	Enterprise Resource Planning
SCM	Supply Chain Management
SVM	Support Vector Machine
MSE	Mean Squared Error
PCA	Principal Component Analysis
DTs	Decision Trees
RF	Random Forest
SQR	Soma dos Quadrados Residuais
FCM	Fuzzy C-Means
MCDM	Multi-Criteria Decision Making
KNN	K-Nearest Neighbours
NLP	Natural Language Processing
RNN	Recurrent Neural Network
ARIMA	Autoregressive Integrated Moving Average
HLT	Holt's Linear Trend
LS-SVM	Least Squares Support Vector Machine
PGPM	Prediction Generation Probability Model
RRC	Relational Regressor Chain
RIPPER	Repeated Incremental Pruning to Produce Error Reduction
SOM	Self-Organizing Maps
AHP	Analytical Hierarchy Process
DEA	Data Envelopment Analysis
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
ELECTRE	ELimination Et Choix Traduisant la REalité
MULTIMOORA	Multi-Objective Optimization by Ratio Analysis
SWOT	Strengths, Weaknesses, Opportunities, and Threats
TOPPRA	Technique for Order of Preference by Similarity to Ideal Solution
WSM	Weighted Sum Model



## **LISTA DE SÍMBOLOS**

- $\beta$  Letra grega beta
- $\varepsilon$  Letra grega épsilon



## SUMÁRIO

1	INTRODUÇÃO.....	29
1.1	Contextualização.....	29
2	REVISÃO BIBLIOGRÁFICA.....	31
2.1	Gestão de fornecedores.....	31
2.2	Aprendizado de máquina.....	31
2.2.1	Aprendizado de máquina supervisionado.....	35
2.2.1.1	Classificação.....	36
2.2.1.2	Régressão.....	37
2.2.2	Aprendizado de máquina não supervisionado.....	37
2.2.2.1	Clusterização.....	38
2.2.2.2	Redução de dimensionalidade.....	38
2.2.3	Algoritmos de machine learning.....	35
2.2.3.1	Decision tree.....	39
2.2.3.2	Random Forest.....	39
2.2.3.3	Régressão linear.....	40
2.2.3.4	K-Means.....	41
2.2.3.5	Fuzzy C-Means.....	41
3	METODOLOGIA.....	42
4	DESENVOLVIMENTO.....	44
4.1	Definição do protocolo de revisão.....	44
4.2	Busca sistemática por estudos.....	46
4.3	Seleção dos estudos.....	47
4.4	Definição dos campos para análise detalhada.....	47
4.5	Extração dos dados.....	48
5	RESULTADOS E DISCUSSÃO.....	49
5.1	Análises bibliométricas.....	49
5.2	Análises quantitativas.....	53
5.3	Discussão dos resultados.....	76
6	CONCLUSÃO.....	79
7	REFERÊNCIAS.....	80



## 1 INTRODUÇÃO

### 1.1 Contextualização

A Gestão da Cadeia de Suprimentos, ou *Supply Relationship Management*, é compreendida na literatura como um gerenciamento que vai além das fronteiras da empresa, com integração de todos os processos e funções de negócios de toda a cadeia de suprimentos, incluindo marketing, gestão, produção, distribuição e demais processos envolvidos na agregação de valor (RUDBERG; OLHAGER, 2003). A gestão de fornecedores em específico é um campo de estudo importante para a área de *Supply Chain*, pois possui um grande potencial de fornecer vantagens competitivas para as companhias.

Uma boa gestão de fornecedores pode melhorar a eficiência dos processos, estando diretamente envolvida na redução de custos, visto que compras de materiais e componentes representam de 50% a 70% do total dos custos de produção (PRAJOGO; OLHAGER, 2012). Além disso, pode estar relacionada à melhoria e eficiência dos processos pela garantia da qualidade e melhoria da produtividade. Todos esses aspectos trazem grande relevância para o tema devido ao contexto de grande competitividade do mercado atual.

A gestão de fornecedores, também citada como *Supplier Relationship Management* (SRM), é um conjunto de práticas e processos de colaboração com os fornecedores, possuindo por objetivo cumprir os requisitos estabelecidos pela cadeia de suprimentos em questão (DUBEY; GUNASEKARAN; CHILDE, 2018). Uma das etapas mais cruciais para a plena eficiência da cadeia de suprimentos é a avaliação de fornecedores (GHORABAEE; ARIMI; ZAVADSKAS, 2017). Essa etapa envolve a consideração de múltiplos critérios, muitas vezes complexos de definir, que costumam ser avaliados por um grupo de especialistas, considerando várias opções de fornecedores.

Além da avaliação, outras etapas do SRM são a seleção, a segmentação e o desenvolvimento desses fornecedores, como descrito no modelo elaborado por Razei (2012). Em linhas gerais, a avaliação consiste na análise das organizações em relação aos critérios estabelecidos, dando aos tomadores de decisão as informações necessárias para realizarem as escolhas que se alinham com seus objetivos estratégicos. Os pesos estabelecidos a esses critérios podem ser distintos, variando com o grau de relevância do mesmo no processo de decisão.

A segmentação consiste na alocação dos fornecedores em diferentes categorias, levando em conta parâmetros que podem variar de acordo com o modelo de segmentação adotado. Um

dos modelos mais influentes foi o proposto por Kraljic (1983), que utiliza para análise os parâmetros de importância da compra e complexidade do mercado, fornecendo quatro grandes categorias para os itens: Não-crítico, Alavancado, Estratégico e Gargalo.

O processo de seleção consiste na escolha entre as opções já avaliadas e segmentadas nas demais etapas, considerando os fatores desejados, como capacidade, custo, qualidade ou outros específicos a cada processo de tomada de decisão. Segundo Boer (2001), essa etapa não deve ser considerada apenas como a decisão final do processo, mas deve incluir vários passos de tomada de decisão anteriores, como a formulação dos critérios de seleção e uma pré-qualificação dos potenciais fornecedores. Além disso, essa etapa costuma utilizar modelos de decisão de ranqueamento das alternativas para auxílio ao processo de seleção.

Por último, a etapa de desenvolvimento com o fornecedor envolve qualquer iniciativa tomada pela organização compradora para a melhora do desempenho do fornecedor. Essa etapa é bastante estratégica e envolve duas das mais evidentes características do capital social: conhecimento compartilhado e o investimento em ativos compartilhados (KRAUSE; HANDFIELD, 2007).

Santos, Osiro e Lima (2017) destacam que, com o uso de tecnologias de informação, como *big data* e sistemas de *Enterprise Resource Planning* (ERP), diversas empresas possuem bancos de dados históricos sobre transações anteriores com fornecedores e esses dados podem ser utilizados para melhorar o processo de gestão de fornecedores. Os autores também apontam que, dentre os diversos critérios que a organização pode escolher para avaliar seus fornecedores, existem os critérios quantitativos, que devem ser analisados por meio de dados numéricos provenientes de seu histórico de compras, e os critérios qualitativos, os quais devem ser avaliados por meio de julgamentos baseados na experiência dos especialistas. Os dados referentes ao histórico de fornecimento podem ser compilados em indicadores quantitativos de desempenho e os especialistas em compras devem combiná-los com avaliações qualitativas para apoiar as decisões de gerenciamento de fornecedores de forma mais ampla (SANTOS; OSIRO; LIMA, 2017).

Para lidar com situações nas quais uma grande quantidade de dados deve ser processada, é comum encontrar na literatura aplicações de *machine learning* em contextos dentro de *Supply Chain Management* (SCM), tais como: desenvolvimento de fornecedores, seleção de fornecedores, avaliação de fornecedores, previsão de demanda, entre outros (JUNIOR; CARPINETTI, 2019). Além disso, segundo Hullermeiter (2021), as técnicas de *machine learning* têm apresentado aplicações promissoras na área dedicada à tomada de decisão automatizada.

Baseado no contexto geral apresentado, é importante buscar responder à seguinte pergunta: como as técnicas de *machine learning* têm sido aplicadas nos processos de *Supplier Relationship Management* (SRM)? Para isso, é essencial investigar e analisar as propostas de modelos de SRM que utilizam abordagens de *machine learning*, a fim de identificar o estado atual da arte e fornecer direcionamentos para futuras pesquisas e aplicações. Destaca-se que, até o momento, não foram encontradas revisões sistemáticas da literatura que explorem especificamente o uso de *machine learning* nos processos de SRM, o que reforça a relevância de compreender de que forma essas técnicas estão sendo aplicadas no contexto da gestão de fornecedores.

## 1.2 Objetivo

O objetivo desta pesquisa é identificar o estado da arte da aplicação de técnicas de aprendizado de máquina em problemas relacionados à seleção, segmentação, avaliação e desenvolvimento de fornecedores. Em linhas gerais, esse objetivo consiste em avaliar e comparar as diferentes técnicas de *machine learning* quanto aos seus objetivos, aplicabilidade, benefícios e dificuldades para apoio à decisão nas atividades de gestão de fornecedores. Esse objetivo geral pode ser subdividido nos seguintes objetivos específicos:

1. Realizar revisão sistemática da literatura sobre aplicação de técnicas de *machine learning* em seleção, segmentação e avaliação e desenvolvimento de fornecedores;
2. Analisar as abordagens de *machine learning* aplicadas nos processos de SRM, identificando os principais algoritmos, modelos e técnicas utilizados;
3. Investigar os desafios e limitações associados à aplicação de técnicas de *machine learning*;
4. Síntese dos estudos analisados, principais descobertas e sugestões para pesquisas futuras.

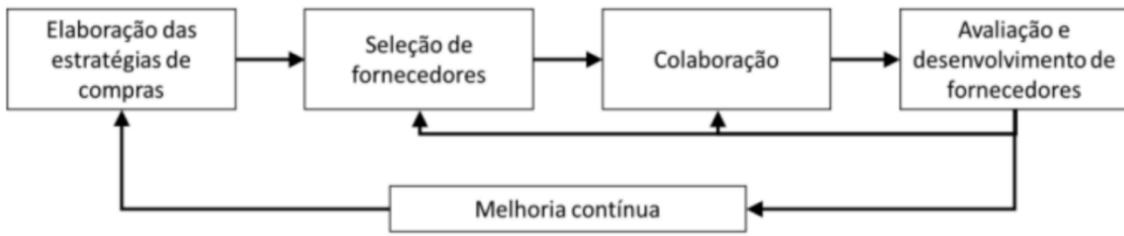
## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 Gestão de fornecedores

A gestão de fornecedores é a área de estudo que visa estudar as melhores práticas da relação entre uma empresa e seus fornecedores no contexto de *Supply Chain*, sendo conhecida na literatura também como *Supplier Relationship Management* (SRM). Dito isso, podemos

definir o SRM como um conjunto de práticas e processos para colaboração com os fornecedores e para o cumprimento de requisitos da cadeia de suprimentos em questão (DUBEY et al., 2018). Segundo Park et al. (2010), um framework que sumariza as principais atividades e etapas relacionadas a essa área de estudo está ilustrado na Figura 1.

Figura 1 – Estrutura SRM



Fonte: Adaptado de Park et al. (2010).

A elaboração das estratégias de compra consiste na classificação dos materiais de compra pelo risco de fornecimento (*supply risk*), sendo o modelo utilizado como base para essa análise o proposto por Kraljic (1983). O segundo passo dessa etapa é a análise da relação com o fornecedor, que consiste na recategorização por um gráfico de relacionamento com fornecedores (*supplier relationship chart*), proposto por Olsen e Ellram (1997). Por fim, a última etapa desse passo é o estabelecimento do plano de ação, que deve seguir a estratégia do portfólio em questão (PARK et al., 2010).

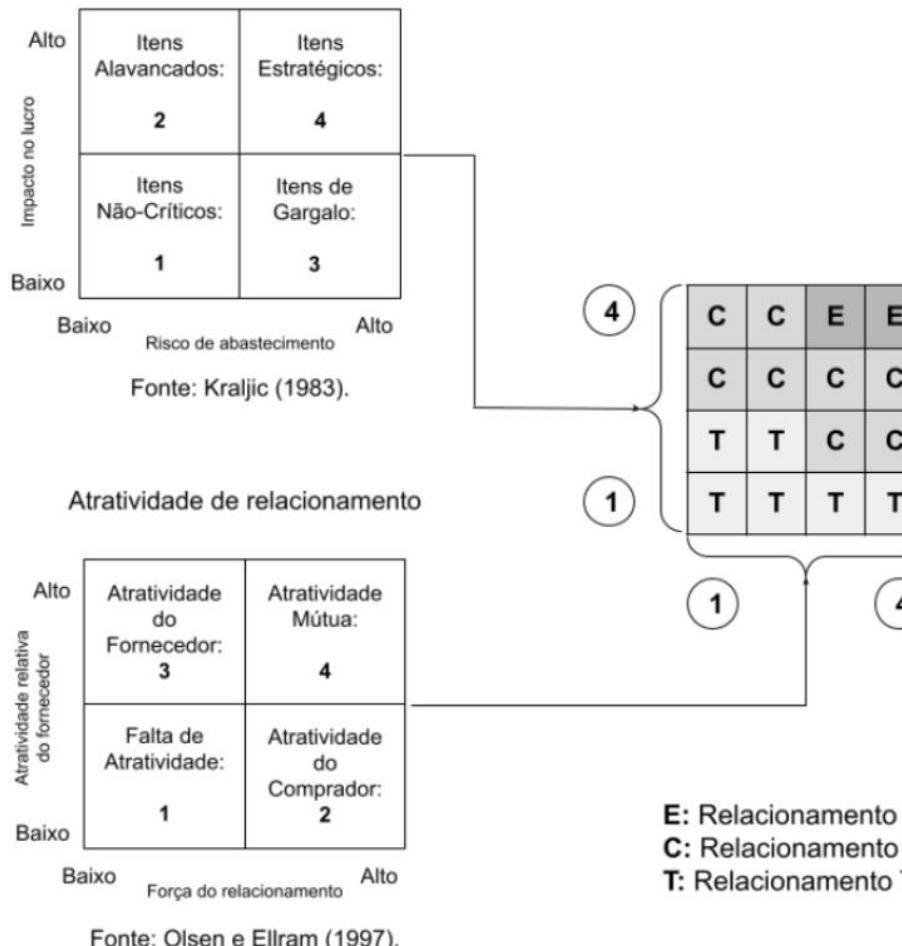
A seleção dos fornecedores é feita em duas etapas. A primeira delas é a pré-seleção dos fornecedores para construção de um conjunto maior de fornecedores (*pool of suppliers*), sendo esta feita por meio da utilização de critérios adequados para cada grupo específico (pool) e da avaliação de cada fornecedor em cada critério. Essa primeira etapa pode ser comparada com a segmentação dos fornecedores, comumente presente em outros modelos presentes na literatura. A segunda etapa é a seleção de parceiros para o fornecimento dos materiais, por meio dos grupos de fornecedores ou por meio dos parceiros já existentes (PARK et al., 2010).

A colaboração pode ser dividida em dois estágios: o de desenvolvimento do produto e o de produção. Basicamente, a informação deve ser compartilhada para ganhos mútuos, o fornecedor deve estar incluído em etapas de elaboração do produto e uma boa arquitetura do sistema de gestão de fornecedores é necessária para guiar esse processo (PARK et al., 2010).

A avaliação e o desenvolvimento de fornecedores seguem três etapas. A primeira é a avaliação da estratégia do material a ser fornecido, a segunda é a avaliação do fornecedor, que

forma a matriz de avaliação apresentada na Figura 2, e, por último, a terceira etapa é a avaliação da relação com o fornecedor (PARK et al., 2010).

Figura 2 – Matriz de avaliação de materiais e fornecedores  
Importância estratégica



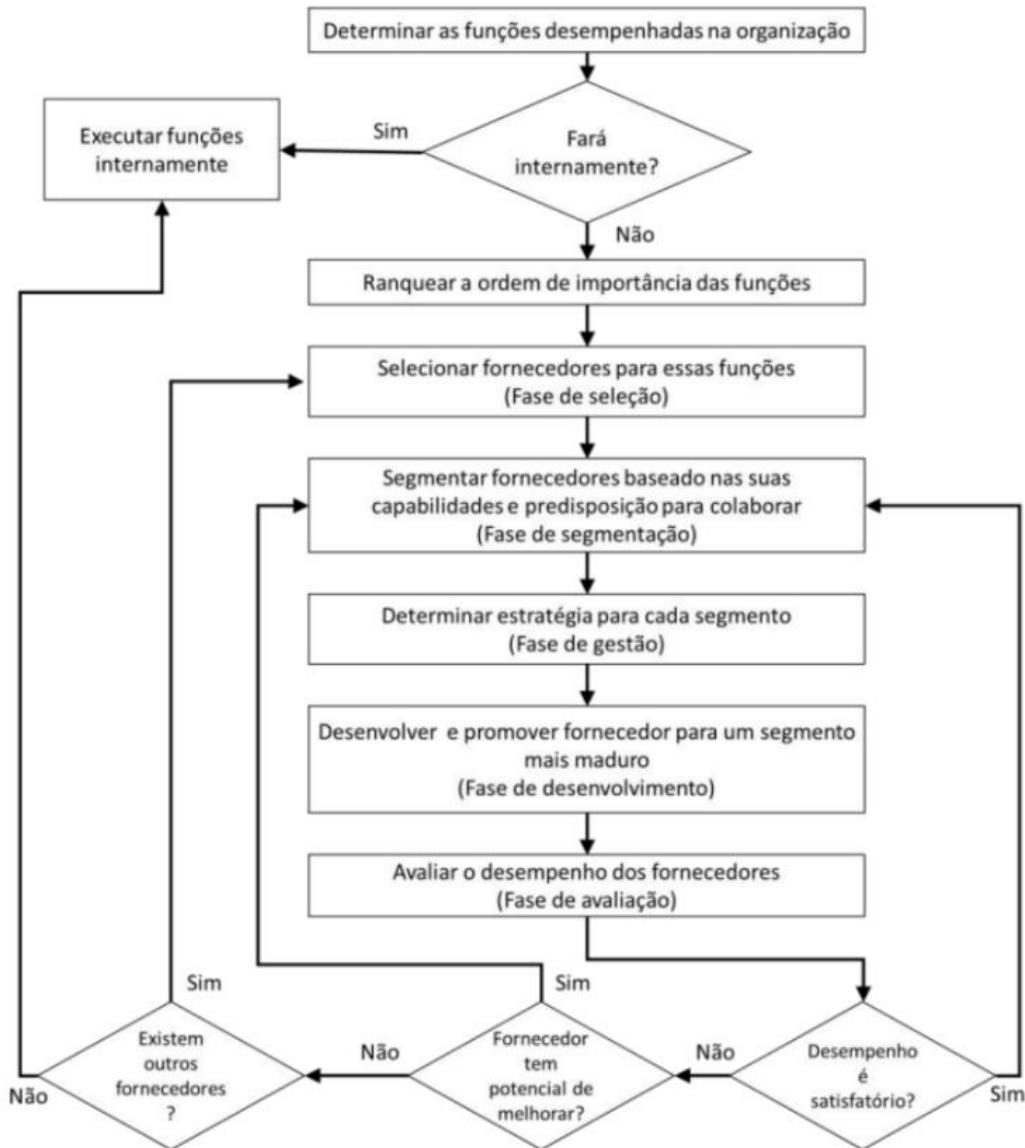
Fonte: Adaptado de Park et al. (2010).

Em uma abordagem menos segmentada, podemos afirmar que a avaliação de fornecedores para apoiar o desenvolvimento da relação com os mesmos se baseia em alocar os fornecedores em diferentes categorias que norteiam quais serão os esforços de melhoria (GALO; CALACHE; CARPINETTI, 2018; REZAEI; ORTT, 2012). Em outras palavras, esse processo de categorização é um processo de segmentação de fornecedores, o qual apoia a escolha das estratégias adequadas para lidar com diferentes segmentos de fornecedores (REZAEI; WANG; TAVASSZY, 2015; REZAEI; ORTT, 2012).

Rezaei e Ortt (2012) também propuseram um *framework* mais conceitual das atividades relacionadas ao SRM com as implicações lógicas do processo, sendo que neste é possível

observar a relação direta entre a etapa de "Avaliação e desenvolvimento de fornecedores" com os processos de segmentação, desenvolvimento e seleção de fornecedores. O framework está representado na Figura 3.

Figura 3 – Framework conceitual das atividades do SRM



Fonte: Adaptado de Rezaei e Ort (2012).

No framework apresentado, quatro processos principais de gestão de fornecedores são identificados: avaliação, seleção, segmentação e desenvolvimento. Esses passos, conforme proposto por Rezaei e Ort (2012), foram fundamentais para a classificação das etapas do processo de SRM ao longo da revisão bibliográfica realizada nesta pesquisa. Para complementar a estrutura das etapas utilizadas na classificação dos artigos em relação aos processos de SRM, De Boer (2001) menciona o processo de pré-qualificação de fornecedores.

Essa etapa ocorre antes da avaliação detalhada e consiste em filtrar um grupo amplo de fornecedores para selecionar apenas aqueles que atendem aos requisitos básicos, de acordo com critérios específicos para o contexto. A pré-qualificação foi incorporada como o quinto processo considerado na revisão sistemática da literatura conduzida na presente pesquisa.

## 2.2 Aprendizado de máquina

O Aprendizado de máquina, também conhecido como *machine learning*, foi definido como o campo derivado da inteligência artificial responsável por estudar o aprendizado automatizado de computadores e máquinas. Segundo Biamonte (2017), os computadores podem ser utilizados para realizar análises de dados clássicas, utilizando métodos de regressão de mínimos quadrados, interpolação polinomial e análise de dados. Os protocolos de aprendizado de máquina podem ser subdivididos nas categorias de supervisionados e não supervisionados.

As diversas técnicas de *machine learning* disponíveis representam uma gama de possibilidades, tais como automatização de processos, reconhecimento de padrões e avaliação de grande volume de dados. Essas possibilidades aplicadas em SRM têm o potencial de impactar positivamente cada uma das etapas dessa área.

### 2.2.1 Aprendizado de máquina supervisionado

Segundo Biamonte (2017), no aprendizado supervisionado, os dados de treinamento são divididos em categorias rotuladas, e o trabalho da máquina é aprender a atribuir rótulos a dados além dos utilizados no processo de treinamento.

Nessas técnicas, é inicialmente feita uma rotulação entre os dados de entrada e as saídas correspondentes esperadas e, com base nesses dados, é realizado o treinamento do modelo. Para testar o modelo, os rótulos são retirados, e o algoritmo trabalha com os dados de acordo com os padrões aprendidos durante o treino, sendo que a saída final tem seu desempenho comparado com os dados de fato rotulados (Fernandes, 2021). Algumas das aplicações mais recorrentes de técnicas supervisionadas estão nos campos de classificação e regressão.

Uma consideração importante para a utilização de técnicas de classificação e regressão em aprendizado de máquina é o potencial que as mesmas possuem de se adaptar excessivamente aos dados, gerando um problema conhecido como *overfitting*. Esse problema ocorre quando a

adaptação aos dados de treinamento é excessiva e o algoritmo não é capaz de realizar generalizações (MAXWELL et al., 2018).

### 2.2.1.1 Classificação

O campo de estudo de classificação em aprendizado de máquina consiste na junção de um conjunto de dados de acordo com rótulos. De acordo com Ngai et al. (2009), esse campo de técnicas atua a partir de categorias pré-definidas, classificando o objeto de estudo a partir de atributos conhecidos. Essas categorias são fornecidas como dados de entrada para o algoritmo, e os padrões de semelhança são identificados durante o treinamento. Durante essa etapa, a maior parte do conjunto de dados é utilizada, enquanto uma porção menor é reservada para testes. O conjunto de teste é fundamental para avaliar o desempenho do algoritmo, o qual é medido por métricas de desempenho.

De acordo com Dos Santos et al. (2022), dentre as métricas mais utilizadas para avaliar o desempenho de modelos de classificação estão a acurácia, precisão, *recall* e *F-score*, as quais são descritas, respectivamente, pelas equações 1 a 4.

$$\text{Acurácia} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} \quad (1)$$

$$\text{Precisão} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{F-score} = \frac{2 \times \text{Recall} \times \text{Precisão}}{\text{Recall} + \text{Precisão}} \quad (4)$$

Onde:

TP = Verdadeiros positivos (True Positive)

TN = Verdadeiros negativos (True Negative)

FP = Falso Positivo (False Positive)

FN = False Negative (False Negative)

A acurácia pode ser interpretada como a proporção de predições corretas feitas pelo modelo em relação ao total de predições feitas, sendo uma medida geral de quão bem o modelo está performando. A precisão, por sua vez, é a proporção de exemplos positivos classificados

corretamente em relação ao total de exemplos classificados como positivos pelo modelo, sendo uma medida da precisão das previsões positivas do modelo. O *recall*, também conhecido como sensibilidade, é a proporção de exemplos positivos classificados corretamente em relação ao total de exemplos reais positivos na base de dados e é uma medida da capacidade do modelo em encontrar todos os exemplos positivos. Por último, o *F-score* é uma medida que combina precisão e *recall* em uma única métrica, calculada como a média harmônica entre essas duas medidas. Essa métrica é útil quando se deseja uma medida que leve em consideração tanto falsos positivos quanto falsos negativos (Miao; Zhu, 2020); Riyanto et al., 2023; (Yacoubi; Axman, 2020).

Algumas das técnicas mais populares para algoritmos de classificação são as *Decision Trees*, *Random Forest* e *Support Vector Machine*(SVM).

### 2.2.1.2 Regressão

A regressão em *machine learning* é aplicada por meio de técnicas de aprendizado supervisionado. Sua base de usabilidade é fazer previsões de dados históricos (*forecasting* e *prediction*) (WU et al., 2019). Sendo assim, podemos definir esse campo de aplicação como destinado a estimar uma variável dependente a partir de uma ou mais variáveis independentes (Roopa & Asha, 2019). Algumas das técnicas mais populares utilizadas para construir algoritmos de regressão são *Decision Tree*, *Random Forest* e regressão linear.

As medidas de sucesso dessa classe de algoritmo se baseiam em medir a discrepância ou distância entre as previsões do modelo e os valores reais dos dados, sendo a medida de sucesso mais comum a soma dos quadrados residuais, também conhecida como *Mean Squared Error* (MSE), descrita pela Equação 5.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

Onde,  $y_i$  corresponde ao valor real e  $\hat{y}_i$  ao valor predito da amostra  $i$  de um conjunto total de amostras de tamanho  $n$ . Vale destacar que o MSE busca sempre ser minimizado em problemas de regressão.

### 2.2.2 Aprendizado de máquina não supervisionado

Segundo Biamonte (2017), no aprendizado não supervisionado não são utilizados conjuntos de treinamentos pré-rotulados, sendo o objetivo das máquinas encontrar categorias naturais nas quais os dados de treinamento se enquadram.

As técnicas não supervisionadas são comumente aplicadas em problemas de agrupamento, também denominados como problemas de clusterização (*clustering*) e em problemas de redução de dimensionalidade (BERTOLINI et al., 2021).

### 2.2.2.1 Clusterização

Segundo Jana et al. (2010), as técnicas de *clusterização* são uma classificação estatística para descobrir se os indivíduos de uma população se enquadram em diferentes grupos, atuando por meio de comparações quantitativas de múltiplas características. Esse campo das técnicas é caracterizado por uma abordagem não supervisionada, que visa identificar dados próximos ou intrinsecamente semelhantes em um mesmo *cluster* (NGAI et al., 2009). A semelhança intrínseca mencionada pode ser interpretada como a distância entre os dados, sendo que os algoritmos atuam buscando minimizar a distância entre os pontos dentro de um *cluster* e maximizar a distância entre os *clusters*, que são localizados através dos centróides. O cálculo dessas distâncias pode ser feito por distância de cossenos (*cosine similarity*) ou distância euclidiana, por exemplo. Além disso, esse campo de aplicações não utiliza rótulos pré-definidos para cada um desses agrupamentos.

As abordagens de *clusterização* podem ainda ser subdivididas em processos de *Hard Clustering*, onde cada dado pertence a um único *cluster*, e em processos de *Soft Clustering*, onde cada dado pode pertencer a mais de um *cluster*, com diferentes níveis de pertencimento que variam entre 0 e 1 (Bora & Gupta, 2014). Um exemplo de algoritmo de *Hard Clustering* é o popular *K-Means* e um exemplo de *Soft Clustering* é o *Fuzzy C-Means*.

### 2.2.2.2 Redução de dimensionalidade

Algoritmos de redução de dimensionalidade no contexto de aprendizado de máquina são técnicas que buscam reduzir o número de variáveis em um conjunto de dados, preservando ao máximo as informações relevantes para a análise (NGAI et al., 2009). Esse tipo de técnica representa os dados de um conjunto em um espaço de menor dimensão, mantendo as características essenciais dos dados originais, permitindo uma simplificação do modelo e facilitando a visualização dos dados. Na redução de dimensionalidade, costuma-se utilizar a

técnica denominada análise de componentes principais (*PCA - Principal Component Analysis*), que se baseia na criação de variáveis e na captura da maior variabilidade da combinação linear dos preditores, que são as variáveis ou características utilizadas para análise nos modelos não supervisionados (Fernandes, 2021).

### 2.2.3 Algoritmos de *Machine Learning*

#### 2.2.3.1 *Decision Tree*

As *Decision Trees* (DTs) são estruturas que, a partir de um nó inicial (nó raiz), se ramificam em outros nós (filhos) até que cheguem às folhas (XU et al., 2022). Nessa estrutura, cada nó denota um cenário de decisão, que pode se ramificar em diversas alternativas, levando a novos nós. Assim, as DTs são capazes de mapear os possíveis resultados finais decorrentes de uma série de decisões, sendo esses resultados as folhas, as quais podem ser associadas a classes de saída do algoritmo.

Dessa forma, as DTs são métodos de aprendizado de máquinas supervisionados e preditivos, capazes de prever valores numéricos reais esperados, atuando como uma abordagem de regressão, ou classes esperadas, atuando como uma abordagem de classificação, sendo ambas as abordagens baseadas em um conjunto de amostras e atributos previamente definidos. Existem diversos algoritmos de treinamento para as DTs, entre eles: *Classification e Regression Trees* (CART) (BREIMAN et al., 1986); *Iterative Dichotomiser 3* (ID3) (QUINLAN, 1986), C4.5 (QUINLAN, 1993), *SLIQ* (MEHTA; AGRAWAL; RISSANEN, 1996) e *SPRINT* (SHAFER, 1996).

O algoritmo CART é o mais conhecido e difundido para processos de regressão e classificação (YAN et al., 2016). O CART divide o espaço em conjuntos retangulares, de modo que classifica qualquer dado de acordo com os espaços definidos. Além disso, esse algoritmo utiliza o conceito de impureza e ganho de informação como critério para a construção de sua árvore, de modo que a divisão de nó pai para nó filho maximize o ganho de informação. A fórmula da Impureza de Gini utilizada como critério para a construção da árvore é descrita pela Equação 6:

$$\text{Gini}(S) = 1 - \sum_{j=1}^n (pj^2) \quad (6)$$

Onde  $S$  representa o conjunto de dados a ser calculado a impureza,  $p_j$  indica a frequência relativa de  $j$  em  $S$  e  $j$  se refere a cada classe ou rótulo presente no conjunto  $S$ . Para o cálculo do ganho de informação, que deve ser maximizado, deve-se calcular a impureza do nó pai subtraída da impureza ponderada dos nós filhos, sendo essa ponderação feita proporcionalmente ao número de amostras de cada nó filho.

### 2.2.3.2 Random Forest

*Random Forest* (RF) é um algoritmo de aprendizado supervisionado amplamente utilizado em diversas áreas de aplicação, tanto para regressão quanto para classificação (Yunxia, 2017). Nesse contexto, as RFs atuam por meio da geração de múltiplas Árvores de Decisão (DTs), cada uma utilizando um subconjunto aleatório dos dados de treinamento (BELGIU; DRĂGUȚ, 2016). Proposto por Breiman (2001), o RF é uma abordagem popular devido à sua robusta capacidade de generalização e à sua habilidade em mitigar problemas de overfitting quando comparado a uma única Decision Tree (ZHANG; QUOST; MASSON, 2023).

Como podemos observar, a *Random Forest* tem sua base na técnica de *Decision Tree*, sendo uma abordagem semelhante. Os detalhes que garantem uma maior robustez da primeira técnica em relação à segunda estão centrados nas etapas de obtenção dos conjuntos de treinamento a serem utilizados na construção de cada DT, na seleção dos critérios e variáveis durante a construção das DTs e na agregação desses múltiplos resultados. As fórmulas utilizadas para a construção das DTs contidas em uma *Random Forest* são as mesmas descritas anteriormente pela Equação 6.

### 2.2.3.3 Regressão linear

A regressão linear convencional é um modelo com uma única variável independente. De acordo com Acharya et al. (2019), essa técnica define a dependência da variável dependente e é descrita pela Equação 7:

$$y = \beta_0 + \beta_1 + \varepsilon \quad (7)$$

Onde  $y$  é a variável dependente,  $x$  é a variável independente, também conhecida como variável preditora,  $\alpha$  é o coeficiente linear, intercepto da reta onde a variável independente vale 0,  $\beta$  é o coeficiente da variável independente, ou coeficiente angular, e  $\epsilon$  é o erro ou resíduo.

Como para todas as técnicas de regressão, a base para o cálculo é a Soma dos Quadrados Residuais (SQR), de forma que essa soma seja minimizada ao fim da previsão. Essa métrica quantifica a diferença entre os valores observados e os valores preditos pelo modelo, sendo essencialmente uma medida de quão bem o modelo representa os dados observados. A fórmula do SQR está descrita pela Equação 8:

$$Variância(SQR) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

Onde  $y_i$  corresponde ao valor real e  $\hat{y}_i$  ao valor predito da amostra  $i$  de um conjunto total de amostras de tamanho  $n$ .

#### 2.2.3.4 *K-Means*

Segundo Bora e Gupta (2014), o algoritmo *K-Means* é um dos mais populares algoritmos de *Hard Clustering*. Ele recebe o parâmetro de entrada  $k$ , que representa o número de *clusters*, e divide um conjunto de  $n$  objetos em  $k$  *clusters* de modo que a similaridade intra-cluster resultante seja alta, mas a similaridade inter-cluster seja baixa. Essa similaridade pode ser interpretada como a distância entre cada ponto do conjunto de dados em um espaço multidimensional, sendo medida pela tradicional distância euclidiana. Esse algoritmo é aplicado de forma iterativa, definindo  $k$  centróides (um para cada *cluster*) e seguindo uma sequência de passos que passa pela escolha inicial do número de *clusters*  $k$ , pela seleção dos centróides iniciais, atribuição dos pontos aos *clusters* e atualização dos centróides, sendo que as duas últimas etapas são repetidas até que o resultado obtido seja satisfatório. Para encontrar o número ideal de *clusters*, é comumente utilizado o *Elbow Method*, ou gráfico do cotovelo, onde o valor ideal de  $k$  é aquele em que  $k+1$  não fornecerá uma redução significativa da Soma dos Quadrados Residuais (SQR), métrica utilizada para medir a variância.

#### 2.2.3.5 *Fuzzy C-Means*

O *Fuzzy C-Means* (FCM) é uma técnica de *Soft Clustering* na qual cada ponto do conjunto de dados pertence a um cluster em algum grau especificado, definido por um grau de

pertencimento. Essa técnica apresenta um método de agrupamento dos pontos que povoam algum espaço multidimensional, sendo esse agrupamento feito em um número específico de *clusters* diferentes (Bora; Gupta, 2014).

Assim como no algoritmo *K-Means*, a maneira mais utilizada para encontrar o número ideal de *clusters* é realizando um processo iterativo chamado de *Elbow Method*, ou gráfico do cotovelo. No FCM, a métrica para priorização dos agrupamentos também é a soma dos quadrados residuais, sendo que o ponto onde é observado que o acréscimo de mais um *cluster* não gera uma redução significativa do RSS é assumido como o número ideal de *clusters*. O diferencial do FCM em relação ao *K-Means* é a possibilidade de representação do grau de incerteza na atribuição de um ponto a um *cluster*.

### 3 METODOLOGIA

A metodologia a ser utilizada no desenvolvimento da revisão sistemática da literatura, representada na Figura 4, é baseada nas metodologias propostas por Randolph (2009) e Tranfield, Denyer e Smart (2003).

Em linhas gerais, para a realização da pesquisa foi necessário definir as *strings* de busca relacionadas ao tema de forma adequada, com o auxílio do professor orientador e de seu aluno de doutorado, Rafael Arantes. Logo em seguida, foi necessário buscar nas bases de dados selecionadas as publicações com proposições de modelos de seleção, segmentação, avaliação e desenvolvimento de fornecedores apoiadas por técnicas de *machine learning*. A segunda grande parte da pesquisa consiste em analisar esses artigos selecionados e classificá-los em relação aos métodos utilizados, aplicação no contexto de SRM, objetivo de aplicação das técnicas, tipo de aprendizado e outros critérios.

Figura 4 – Esquemático das etapas de realização da pesquisa



Fonte: Elaborado pela autora.

As etapas estão melhor detalhadas a seguir:

1. **Etapa 1:** A primeira etapa consistiu na definição exata das *strings* de busca nas bases de dados com auxílio do professor orientador responsável e especialistas no tema. As *strings* deveriam buscar por palavras e expressões relacionadas a *machine learning* e à gestão de fornecedores. As mesmas foram inseridas nas bases de dados escolhidas, e foi feita a seleção de todos os artigos que contêm uma das combinações definidas.
2. **Etapa 2:** A etapa 2 consistiu na eliminação de artigos duplicados, para a qual foi feita uma comparação dos títulos dos artigos obtidos de cada uma das bases de dados.
3. **Etapa 3:** Essa etapa consistiu na pré-classificação dos artigos restantes da etapa anterior, que foram divididos em três categorias: A, B e C, melhor definidas abaixo:
  - o **Categoria A:** Engloba os artigos que claramente abordam o foco da pesquisa. Para identificá-los, analisamos o título e o resumo, verificando se é explicitamente afirmado que o estudo trata de métodos de *machine learning* para apoiar a gestão de relacionamento com fornecedores.
  - o **Categoria B:** Engloba os artigos que possivelmente tratam do nosso tema de pesquisa. Ao examiná-los, procuramos menções separadas aos temas de *machine learning* e gestão de fornecedores no título e no resumo, mesmo que não estejam diretamente relacionadas. No entanto, ainda há a possibilidade de esses artigos estarem ligados ao escopo do estudo.

- **Categoria C:** Engloba os artigos que não se encaixam no escopo da pesquisa. Os temas são mencionados de forma completamente desconexa, sem qualquer conexão perceptível com a temática de *machine learning* para gestão de fornecedores.
4. **Etapa 4:** Após a classificação ABC, os artigos classificados como “C” foram descartados, e foram definidos os critérios para classificação aprofundada. Os campos para classificação foram decididos em conjunto com o professor orientador, levando em consideração as análises foco da pesquisa.
  5. **Etapa 5:** Os artigos previamente classificados como A e B foram analisados com base nos critérios definidos na etapa anterior. Os artigos foram baixados e acessados integralmente, assim, além da classificação detalhada, a etapa 5 verificou quais artigos realmente abordavam o tema central da revisão e quais eram acessíveis. Após essa categorização, os resultados foram analisados e a conclusão do relatório foi desenvolvida.

## 4 DESENVOLVIMENTO

As atividades desenvolvidas e apresentadas a seguir foram obtidos a partir das cinco etapas descritas na metodologia em conjunto com o professor orientador. Abaixo, é fornecida uma descrição de como esses resultados foram alcançados em sequência cronológica:

### **I. Definição do protocolo de revisão:** 2 meses

Durante esse período, discutiu-se quais *strings* utilizar e como aprofundar teoricamente nas técnicas de *machine learning*. Decidiu-se que a autora faria apresentações periódicas sobre diferentes técnicas de *machine learning* para o grupo de pesquisa, em paralelo ao andamento da pesquisa. A técnica a ser apresentada era sempre escolhida pelo professor orientador, priorizando as mais relevantes no contexto estudado.

### **II. Busca sistemática de estudos:** 2 meses

Durante esse período, ocorreram as etapas 1 e 2 descritas anteriormente (Figura 4), respectivamente as etapas de definição das *strings* de busca e extração dos artigos com eliminação dos duplicados.

### **III. Seleção de estudos:** 2 meses

A seleção dos estudos foi feita por meio da etapa 3 descrita anteriormente (Figura 4), que retirou artigos classificados como fora do escopo de pesquisa das análises.

#### IV. Extração de dados: 4 meses

A etapa de extração de dados compreende as etapas 4 e 5 do processo (Figura 4). Durante essas etapas, foram definidos os campos de análise detalhada e realizada a leitura dos artigos previamente selecionados, visando extrair os dados e informações conforme os critérios estabelecidos. Além da extração dos dados, foi feito um fichamento simples de cada um dos artigos classificados como dentro do escopo da revisão.

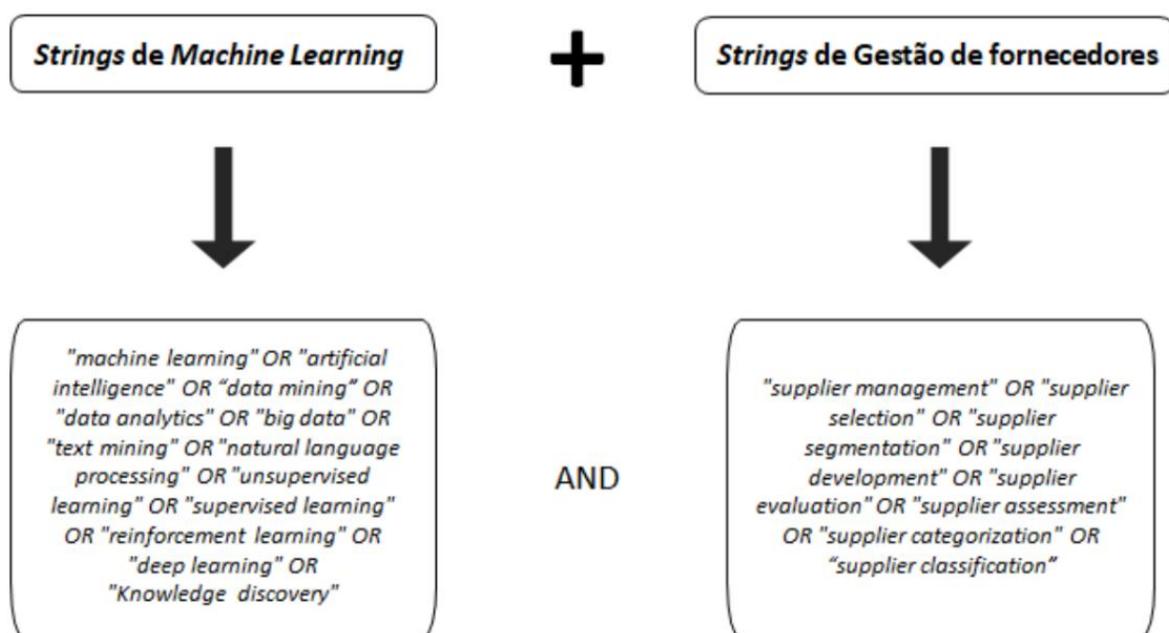
##### 4.1 Definição do protocolo de revisão e aprofundamento teórico em *machine learning*

As *strings* selecionadas após as discussões foram as descritas abaixo:

- Para a temática de Inteligência Artificial: "*machine learning*" OR "*artificial intelligence*" OR "*data mining*" OR "*data analytics*" OR "*big data*" OR "*text mining*" OR "*natural language processing*" OR "*unsupervised learning*" OR "*supervised learning*" OR "*reinforcement learning*" OR "*deep learning*" OR "*knowledge discovery*".
- Para a temática de gestão de fornecedores: "*supplier management*" OR "*supplier selection*" OR "*supplier segmentation*" OR "*supplier development*" OR "*supplier evaluation*" OR "*supplier assessment*" OR "*supplier categorization*" OR "*supplier classification*".

A Figura 5 abaixo representa como as *strings* foram inseridas nas bases de busca.

Figura 5 – *Strings* de busca dos artigos



Fonte: Elaborado pela autora.

A seleção de termos de pesquisa específicos foi fundamental para garantir a abrangência e relevância da revisão sistemática da literatura sobre métodos de aprendizado de máquina aplicados para apoiar a gestão de fornecedores.

Na temática de Inteligência Artificial, foram selecionadas palavras-chave que abrangem uma variedade de técnicas e conceitos amplamente utilizados em estudos relacionados a *machine learning* e análise de dados, tais como "*machine learning*", "*artificial intelligence*", "*data mining*", "*big data*", entre outros. Essa abordagem permitiu a inclusão de estudos que exploram diversas abordagens e tecnologias neste campo em constante evolução.

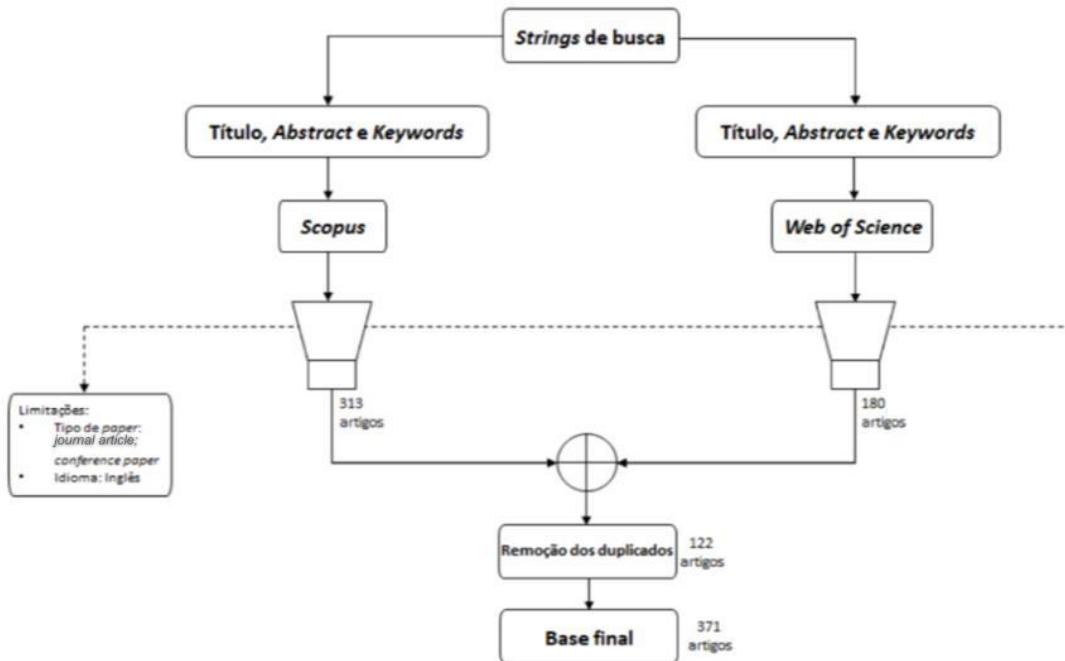
Já na temática de gestão de fornecedores, foram escolhidas palavras-chave que refletem os principais aspectos e processos envolvidos na gestão eficaz dos fornecedores, como "*supplier management*", "*supplier selection*", "*supplier development*", entre outros. Esses termos foram selecionados com o intuito de abranger estudos que discutem estratégias, modelos e ferramentas para melhorar a eficiência e eficácia da gestão de fornecedores.

Além da definição do protocolo, em paralelo às etapas diretamente relacionadas à pesquisa, foram feitos estudos de técnicas de *machine learning* para um aprofundamento teórico no tema. A ideia desses estudos era que fosse feita uma apresentação para o grupo de pesquisa de uma técnica a ser escolhida pelo professor orientador. As apresentações seguiram uma estrutura de introdução, contextualização da técnica, aplicação do algoritmo em um exemplo simples e, por último, uma aplicação em forma de código. A ideia de fazer uma aplicação simples era que os algoritmos estudados pudessem ser aplicados em um conjunto de dados simples e que os passos fossem apresentados também de forma visual, priorizando a didática. Já a aplicação em forma de código, teve como objetivo a capacitação para usos convencionais dessas técnicas. As técnicas estudadas foram *Decision tree*, aplicada para classificação e regressão, *Random Forest*, também aplicada para classificação e regressão, *K-means* e *Fuzzy C-Means*.

## 4.2 Busca Sistemática por Estudos

A busca sistemática por estudos consiste na inserção das *strings* definidas nas bases de dados e na remoção de artigos duplicados. Essa etapa extraiu inicialmente um total de 493 artigos de ambas as bases de dados e, após a remoção dos duplicados, esse número caiu para 371. A Figura 6 abaixo faz uma representação visual desse processo.

Figura 6 – Esquematização da extração dos artigos



Fonte: Elaborado pela autora.

#### 4.3 Seleção dos estudos

A seleção dos estudos ocorreu durante a etapa 3 apresentada pela Figura 4, que consistiu na pré-classificação dos artigos entre as categorias A, B e C. Essa atribuição foi baseada na análise dos resumos, os quais são acessíveis a todos, não sendo necessário abrir cada um dos *papers*.

#### 4.4 Definição dos campos para análise detalhada

A definição dos campos de análise dos artigos foi realizada em conjunto com o professor orientador e seu aluno de doutorado Rafael Arantes. Foi levada em conta a relevância de cada campo de classificação para as análises quantitativas e qualitativas que deveriam ser feitas ao fim do processo de classificação dos artigos.

Os campos para análise detalhada definidos foram os seguintes:

- **Autores:** Autores responsáveis pela publicação da pesquisa
- **Ano:** Ano de publicação
- **Países:** Países que contribuíram na publicação da pesquisa
- **Instituições:** Instituições responsáveis pela publicação do estudo

- **Técnicas de machine learning:** Técnicas de *machine learning* que foram utilizadas em etapas do método proposto pelos autores nos artigos.
- **Tarefas de machine learning:** A tarefa das técnicas de machine learning, classificadas em algoritmos de classificação, regressão, clusterização ou associação.
- **Abordagem de machine learning:** Esse campo se refere ao tipo de aprendizado dos algoritmos de machine learning, que se dividem entre supervisionados e não supervisionados.
- **Objetivo de aplicação da técnica:** Esse campo exige uma análise mais ampla do intuito de utilização da técnica de machine learning dentro do método proposto pelo autor, relacionando esse objetivo de aplicação da técnica com o processo de SRM envolvido.
- **Tipo de data set:** Análise da origem do *data set* utilizado, entre real ou simulado.
- **Processo de Supplier Relationship Management (SRM):** Essa classe refere-se à identificação de qual etapa do SRM o método proposto na publicação mais se alinha. As opções para essa classificação incluem: avaliação, seleção, segmentação, desenvolvimento e pré-qualificação. A avaliação de fornecedores envolve métodos que analisam um conjunto de fornecedores com base em critérios pré-definidos ou processos que estabelecem esses critérios. A seleção, embora relacionada à avaliação, distingue-se pela criação de um ranking dos fornecedores avaliados. A segmentação divide um conjunto maior de fornecedores, classificando-os conforme o objetivo do estudo. A pré-qualificação refere-se à validação preliminar de um grupo de fornecedores. Por fim, o desenvolvimento está relacionado a métodos que visam a criação de planos para melhorar o desempenho dos fornecedores.
- **Representação linguística:** Refere-se à classificação entre linguagem convencional e linguagem *fuzzy*.
- **Método Multi-Criteria Decision Making (MCDM):** Relaciona-se à utilização de alguma técnica de suporte MCDM em uma das etapas do método proposto na publicação.

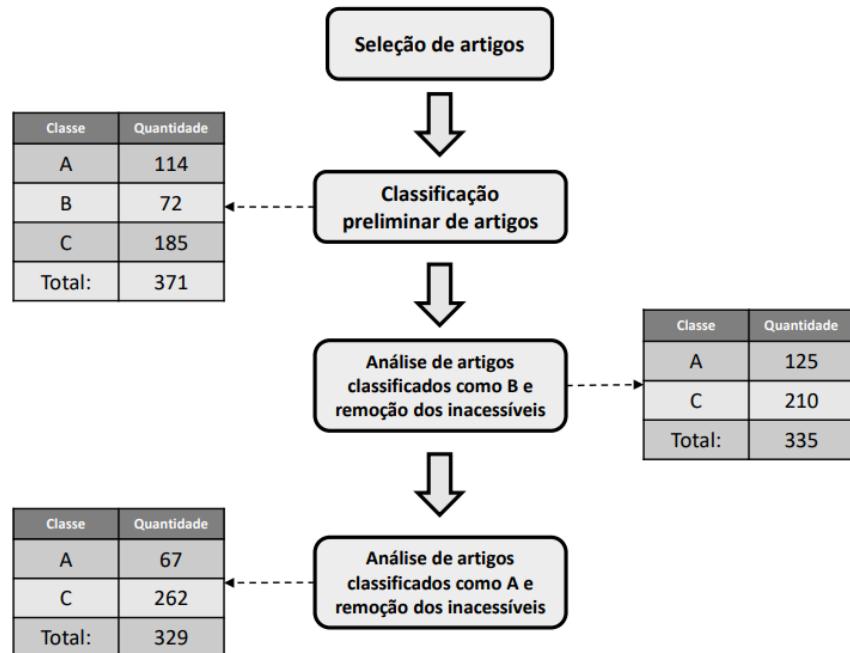
**Aplicação/ Tipo de indústria:** Diz respeito à área ou indústria onde o método foi aplicado, caso tenha sido testado em um cenário real.

#### 4.5 Extração dos dados

Neste período da pesquisa, foi realizada a leitura dos artigos previamente selecionados, visando extrair os dados e informações conforme os critérios estabelecidos para sua análise

detalhada. A leitura de cada artigo foi feita acessando cada um individualmente, portanto, a conferência da disponibilidade do artigo para o domínio de acesso utilizado foi feita em paralelo a essa classificação. No total, 67 publicações foram classificadas como A e acessíveis para o estudo. Além da categorização dos artigos nos campos de classificação definidos em conjunto ao grupo de pesquisa, um pequeno fichamento foi feito também para cada um dos 67 artigos, com o objetivo de facilitar eventuais análises qualitativas dos conteúdos abordados. Os fichamentos seguiram um formato simples, com uma secção dedicada a resumir o propósito da pesquisa e motivação do estudo, uma para sumarizar os pontos mais relevantes do referencial teórico, outra para resumir o método desenvolvido pelos autores, destacando seu passo a passo e uma última secção para destacar os principais pontos apresentados na conclusão do estudo. A figura 7 abaixo sintetiza o volume de publicações obtidas ao longo da pesquisa.

Figura 7 - Volume de publicações obtidos ao longo da pesquisa



Fonte: Elaborado pela autora

## 5 RESULTADOS E DISCUSSÃO

### 5.1 Análises bibliométricas

No total, foram encontrados 67 artigos que abordavam o tema em análise e que estavam disponíveis para acesso. O primeiro resultado obtido foi acerca do volume de publicações nos

últimos anos. A mais antiga publicação analisada foi feita no ano de 2001 e no ano de 2023 foram consideradas apenas publicações feitas até o mês de abril. Na figura 8 é apresentado um gráfico com essa distribuição ao longo dos anos.

Figura 8 - Distribuição do número de publicações ao longo dos anos



Fonte: Elaborado pela autora.

Não foi observada uma tendência constante de crescimento de publicações sobre o tema desde 2001, mas é possível observar um aumento considerável na média de publicações desde o ano de 2018 até o período de formação da base de dados para análises (Abril de 2023).

Em relação aos tipos de publicação, 45 das publicações foram obtidas no formato de artigo e 22 foram obtidas no formato de *conference paper*. Foi observada também uma clara predominância de alguns países nas publicações relacionadas ao tema. Para essa análise, vale ressaltar que um único artigo pode ter tido colaboração de mais de um país para seu desenvolvimento. A tabela 1 abaixo apresenta os países que publicaram sobre o tema e seu respectivo número de colaboração em publicações.

Tabela 1 - Países que publicaram sobre o tema e seus números de publicações

País	Nº de publicações
China	24
Estados Unidos	8

Irã	7
Índia	5
Canadá	4
Taiwan	4
Coreia do sul	3
Bangladesh	2
Indonésia	2
Malásia	2
Turquia	2
Alemanha	1
Austrália	1
Bélgica	1
Brasil	1
Islândia	1
Itália	1
Japão	1
Malásia	1
México	1
Palestina	1

Portugal	1
Romênia	1
Rússia	1
Reino Unido	1

Fonte: Elaborado pela autora.

É possível observar uma clara predominância de colaborações da China sobre o tema analisado, que participou de aproximadamente 36% do total de publicações. Estados Unidos, Irã e Índia foram os demais países que apresentaram um número considerável de participação em publicações, com aproximadamente 12%, 10% e 7% de participação nas publicações, respectivamente. Não foi observada uma correlação forte de colaboração entre esses países em suas publicações.

Em relação ao Brasil, um único estudo foi observado sobre o tema. A universidade responsável por sua publicação foi a Universidade Federal de Santa Catarina, em parceria com a *Berlin School of Economics and Law*. A publicação em questão propôs um método de segmentação de fornecedores, analisando resiliência e risco nesse processo, onde a técnica de machine learning utilizada foi o K-Nearest Neighbours (KNN), uma técnica de classificação que teve como objetivo segmentar os fornecedores entre três classes, sendo elas fornecedores resilientes, não resilientes e moderadamente resilientes.

Foi analisado também o número de publicações por autores, mas não foi observado nenhum autor que tenha contribuído em mais de 2 publicações sobre o tema, o que não demonstrou uma predominância individual de ninguém no contexto acadêmico. Vale ressaltar que foram considerados para análise apenas os 3 primeiros autores de cada publicação.

Além da análise por autores, foi feita também uma análise das publicações por instituições de pesquisa. O resultado das instituições que participaram de pelo menos 2 pesquisas sobre o tema é descrito na Tabela 2.

Tabela 2 - Instituições que mais publicaram sobre o tema

Universidade	Número de publicações participante

Islamic Azad University	5
National Taipei University of Technology	3
Shandong University	2
Hong Kong Polytechnic University	2

Fonte: Elaborado pela autora.

Em relação aos autores dos artigos, também não foi observado um número elevado de publicações de um mesmo autor sobre o tema ou uma correlação recorrente de colaboração entre autores.

## 5.2 Análises quantitativas

Durante a fase de leitura e análise, uma das informações mais relevantes coletadas foi a identificação das técnicas de *machine learning* mais utilizadas no contexto de *Supplier Relationship Management* (SRM). A Tabela 3 mostra a frequência com que cada técnica foi aplicada nos métodos propostos nas publicações. Vale destacar que uma mesma publicação pode empregar mais de uma técnica de *machine learning*.

Tabela 3 - Recorrência de uso de cada técnica de *Machine Learning*

Técnica de Machine Learning	Número de vezes que apareceu
<i>Neural Networks</i>	23
<i>Decision Tree</i>	10
<i>Support Vector Machine (SVM)</i>	9
<i>K-Means</i>	7
<i>Association Rule Mining</i>	5

<i>Logistic Regression</i>	4
<i>Natural Language Processing (NLP)</i>	3
<i>Fuzzy Neural Network</i>	3
<i>K-Nearest Neighbors (KNN)</i>	3
<i>Random Forest</i>	3
<i>Recurrent Neural Network (RNN)</i>	3
<i>Apriori</i>	2
<i>Autoregressive Integrated Moving Average (ARIMA)</i>	2
<i>Bayesian Networks</i>	2
<i>Fuzzy C-Means</i>	2
<i>Naive bayes</i>	2
<i>Adaboost</i>	1
<i>Clustering K-Prototyping</i>	1
<i>Gradient Boosting</i>	1
<i>Hierarchical Clustering</i>	1
<i>Holt's Linear Trend (HLT)</i>	1
<i>i-PM</i>	1
<i>Least Squares Support Vector Machine (LS-SVM)</i>	1
<i>Prediction Generation Probability Model (PGPM)</i>	1

<i>Projected Clustering</i>	1
Relational Regressor Chain (RRC)	1
Repeated Incremental Pruning to Produce Error Reduction (RIPPER)	1
Self-Organizing Maps (SOM)	2
Semi-Fuzzy Support Vector Domain Description	1

Fonte: Elaborado pela autora.

As técnicas mais frequentes foram as Redes Neurais (*Neural Networks*), presentes em cerca de 34% das publicações analisadas. Em seguida, aparecem *Decision Tree* (15%), *Support Vector Machine* (13%), *K-means* (10%) e *Association Rule Mining* (7%)

Também foi investigado o tipo de abordagem de *machine learning* utilizada nos artigos, considerando se foram supervisionadas, não supervisionadas ou uma combinação de ambas. Os resultados dessa análise estão detalhados na Tabela 4.

Tabela 4 - Segmentação por abordagens supervisionadas ou não supervisionadas

<b>Tipo de abordagem</b>	<b>Número de publicações</b>
Não supervisionado	18
Supervisionado	43
Mesclou ambas as abordagens	6

Fonte: Elaborado pela autora.

É possível observar que aproximadamente 64% das publicações adotaram apenas técnicas supervisionadas em seus métodos propostos, enquanto 27% utilizaram apenas abordagens não supervisionadas. As demais publicações (9%) mesclaram ambas as abordagens.

Foram analisados os conjuntos de dados empregados nas publicações, classificando-os entre reais e simulados. A Tabela 5 apresenta esses dados.

Tabela 5 - Tipos de data set utilizados

Tipo de data set	Quantidade
Real	51
Simulado	16

Fonte: Elaborado pela autora.

A análise revelou que 76% das publicações utilizaram conjuntos de dados reais, enquanto 24% optaram por conjuntos simulados.

As tarefas de *machine learning* mais recorrentes nos estudos analisados também foram mapeadas. Vale ressaltar que alguns artigos empregaram mais de uma tarefa ao longo de seus métodos propostos. A distribuição dessas tarefas está na Tabela 6.

Tabela 6 - Recorrência de tarefas de *Machine Learning*

Task de Machine Learning	Recorrência
Associação	10
Classificação	26
Clusterização	17
Regressão	23

Fonte: Elaborado pela autora.

A maioria das publicações focou em abordagens de classificação e regressão, enquanto técnicas de associação e clusterização foram menos abordadas.

Foi analisada também a distribuição dos métodos propostos nos artigos em relação às etapas do processo de Gerenciamento de Relacionamento com Fornecedores (SRM). A Tabela 7 apresenta essa distribuição. É importante observar que, para a categorização, considerou-se

apenas a etapa mais relevante para o objetivo do método proposto, ou seja, cada publicação foi associada a apenas uma etapa do processo SRM.

Tabela 7 - Distribuição das aplicações em relação à etapa do SRM

Etapas do processo SRM	Recorrência
Avaliação	17
Desenvolvimento	2
Pré-qualificação	2
Segmentação	13
Seleção	33

Fonte: Elaborado pela autora.

Observa-se que a etapa do processo SRM mais frequente nos estudos foi a de seleção (49%), enquanto as menos frequentes foram desenvolvimento e pré-qualificação, representando a etapa principal em apenas 3% das publicações, cada uma.

Também foi analisado o tipo de linguagem utilizado nos artigos, distinguindo entre linguagem *fuzzy* e lógica clássica. Dos artigos analisados, cerca de 67% (45 artigos) utilizaram apenas lógica clássica, enquanto aproximadamente 33% (22 artigos) aplicaram linguagem *fuzzy* em algum momento da proposição do método.

A Tabela 8 detalha a distribuição dessas linguagens em relação às técnicas de machine learning utilizadas nos artigos. É importante destacar que alguns artigos empregaram mais de uma técnica de machine learning ao longo da proposta do método.

Tabela 8 - Distribuição dos quantidade de referente ao uso do tipo de linguagem e abordagens de  
*Machine Learning*

Representação Linguística	Associação	Classificação	Clusterização	Régressão	Total

<i>Fuzzy</i>	4	9	5	5	23
Clássica	6	19	8	18	51

Fonte: Elaborado pela autora.

A distribuição do uso das linguagens *fuzzy* e convencional nas tarefas de *Machine Learning* foi proporcional à frequência geral dessas abordagens: aproximadamente 31% para a linguagem *fuzzy* e 69% para a lógica clássica.

Em relação à distribuição dessas linguagens nas etapas do processo de SRM, os resultados estão detalhados na Tabela 9.

Tabela 9 - Distribuição dos artigos sobre as linguagens utilizadas e as etapas do processo de SRM

<b>Representação Linguística</b>	<b>Avaliação</b>	<b>Desenvolvimento</b>	<b>Pré-qualificação</b>	<b>Segmentação</b>	<b>Seleção</b>	<b>Total</b>
<i>Fuzzy</i>	3	0	1	5	13	22
Clássica	14	2	1	8	20	45

Fonte: Elaborado pela autora.

Nessa análise, cada publicação foi associada apenas à etapa mais relevante dentro do método proposto. Observou-se que, especialmente na etapa de avaliação, a lógica clássica foi mais utilizada do que a *fuzzy*. Nas demais etapas, a distribuição se mostrou proporcional à frequência geral de uso de cada abordagem.

Além das abordagens de *machine learning*, das etapas do processo SRM e do tipo de linguagem utilizada, foram também analisadas as técnicas de *Multi-criteria Decision Making* (MCDM) presentes nas publicações. Dos estudos analisados, 44 publicações (cerca de 66%) não utilizaram abordagens MCDM, enquanto 23 publicações (aproximadamente 34%) empregaram essas técnicas. As técnicas MCDM mais utilizadas estão listadas na Tabela 10.

Tabela 10 - Recorrência de uso de cada técnica MCDM

Técnica MCDM	Número de vezes que foi utilizada
<i>AHP (Analytical Hierarchy Process)</i>	9
<i>DEA (Data Envelopment Analysis)</i>	7
<i>Fuzzy TOPSIS (Technique for Order Preference by Similarity to Ideal Solution)</i>	2
<i>Best/Worst Method</i>	1
<i>ELECTRE (ELimination Et Choix Traduisant la REALité)</i>	1
<i>Fuzzy Best Worst Method</i>	1
<i>Grey Correlation Analysis</i>	1
<i>MULTIMOORA (Multi-Objective Optimization by Ratio Analysis)</i>	1
<i>SWOT (Strengths Weaknesses Opportunities and Threats) Fuzzy</i>	1
<i>TOPPRA (Technique for Order of Preference by Similarity to Ideal Solution)</i>	1
<i>Weighted Sum Model (WSM)</i>	1

Fonte: Elaborado pela autora.

Observa-se que apenas três abordagens MCDM foram aplicadas pelo menos duas vezes: AHP, DEA e *Fuzzy Topsis*. A distribuição do uso dessas técnicas nas diferentes etapas do processo SRM está apresentada na Tabela 11. Vale ressaltar que alguns artigos utilizaram mais de uma técnica MCDM ao longo da proposição de seus métodos.

Tabela 11 - Distribuição do uso das técnicas MCDM

Técnicas MCDM	Etapa do processo SRM					
	Avaliação	Desenvolvimento	Pré- qualificação	Segmentação	Seleção	Total
<i>AHP</i>	2	0	0	1	6	9
<i>DEA</i>	2	0	0	4	1	7
<i>Best Worst Method</i>	0	0	0	0	1	1
<i>ELECTRE</i>	0	0	0	0	1	1
<i>Fuzzy Best Worst Method</i>	1	0	0	0	0	1
<i>Fuzzy Topsis</i>	1	0	0	0	1	2
<i>Grey Correlation Analysis</i>	0	0	0	0	1	1
<i>MULTIMOO RA</i>	0	0	0	0	1	1
<i>SWOTfuzzy</i>	0	0	0	0	1	1
<i>TOPPRA</i>	1	0	0	0	0	1
<i>Weighted Sum Model</i>	0	0	0	0	1	1
<b>Total:</b>	7	0	0	5	14	26

Fonte: Elaborado pela autora.

A Tabela 12 apresenta um resumo detalhado da aplicação das técnicas de *Machine Learning*, abordagens MCDM e as etapas do processo SRM em que foram empregadas.

Tabela 12 - Distribuição dos artigos entre as abordagens de *machine learning* e as etapas do processo SRM

Processo de SRM	Abordagem de <i>Machine Learning</i>	Técnica <i>Machine learning</i>	Nome do artigo	Autores	Ano
Avaliação	Associação	<i>Apriori</i>	<i>Mining Pattern of Supplier with the Methodology of Domain-Driven DataMining</i>	Xu, Xu and Lin, Jie and Xu, Dongming	2009
		<i>Association Rule Mining</i>	<i>Optimal control in dynamic food supply chains using big data</i>	Kappelman, Ashton Conrad and Sinha, Ashesh Kumar	2020
			<i>Application of association rule mining in supplier selection criteria</i>	Haery, A. and Salmasi, N. and Modarres Yazdi, M. and Iranmanesh, H.	2008
		<i>Natural Language Processing (NLP)</i>	<i>A novel Multiple Attribute Decision Making approach based on interval data using U2P-Miner algorithm</i>	Golpira, Heris	2018
	Classificação	<i>Natural Language Processing (NLP)</i>	<i>A Patent-Based Tool to Support Component Suppliers Assessment in the Smartphone Supply Chain</i>	Cammarano, A. and Varriale, V. and Michelino, F. and Caputo, M.	2021

		<i>Neural Networks; Clustering K-Prototyping; Decision Trees</i>	<i>A case based reasoning approach on supplier selection in petroleum enterprises</i>	<i>Zhao, K. and Yu, X.</i>	2010
		<i>Neural Networks</i>	<i>Evaluation Model of Power Operation and Maintenance Based on TextEmotion Analysis</i>	<i>Wan, Wei and Liu, Yuanlong and Han, Xingwang and Wang, Huijian</i>	2021
		<i>Prediction Generation Probability Model (PGPM); Neural Networks</i>	<i>Evaluation Model of Power Operation and Maintenance Based on Text Emotion Analysis</i>	<i>Wang, H. and Wang, H. and Hu, J. and Xu, J. and Li, J. and He, B.</i>	2021
		<i>Neural Network; Natural Language Processing (NLP)</i>	<i>Risk assessment for global supplier selection using text mining</i>	<i>Su, Chuan-Jun and Chen, Yin-An</i>	2018
		<i>Support Vector Machine (SVM)</i>	<i>A Study of Supplier Selection Method Based on SVM for Weighting Expert Evaluation</i>	<i>Zhao, L. and Qi, W. and Zhu, M.</i>	2021
			<i>Application of an MCDM model with data mining techniques for green</i>	<i>Liou, James J. H. and Chang, Mu-Hsin and Lo, Huai-Wei</i>	2021

			<i>supplier evaluation and selection</i>	<i>and Hsu, Min-Hsi</i>	
Clusterizaçã o	<i>K-Means</i>	<i>Dashboard application model in supplier evaluation by using artificial immune system and data mining methods</i>	<i>Yurtay, Yuksel and Ayanoglu, Murat</i>	2020	
		<i>Estimation of the features influence on cluster partition</i>	<i>Kolesnikova, D. and Andreev, Y. and Iureva, R.</i>	2021	
	<i>Neural Networks; Self-Organizing Maps (SOM)</i>	<i>A hybrid approach to supplier selection for the maintenance of a competitive supply chain</i>	<i>Ha, S.H. and Krishnan, R.</i>	2007	
	<i>Neural Networks; Clustering K-Prototyping; Decision Trees</i>	<i>A case based reasoning approach on supplier selection in petroleum enterprises</i>	<i>Zhao, K. and Yu, X.</i>	2010	
Regressão	<i>Least Squares Support Vector Machine (LS-SVM); Neural Networks</i>	<i>A new enhanced support vector model based on general variable neighborhood search algorithm for supplier performance evaluation: A case study</i>	<i>Vahdani, B. and Mousavi, S.M. and Tavakkoli-Moghaddam, R. and Hashemi, H.</i>	2016	

		Neural Networks	<i>Supply Capability Evaluation of Intelligent Manufacturing Enterprises Based on Improved BP Neural Network</i>	Quan, Quan and Zhang, Zhongqiang	2022
			<i>An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information</i>	Ã‡elebi, D. and Bayraktar, D.	2007
		Neural Networks; Self-Organizing Maps (SOM)	<i>A hybrid approach to supplier selection for the maintenance of a competitive supply chain</i>	Ha, S.H. and Krishnan, R.	2007
Desenvolvimento	Associação	Apriori	<i>Data mining for the optimization in manufacturing process</i>	He, T.	2010
	Clusterização	K-Means	<i>Data analytics for transforming towards smart supplier relationship management (A case study in manufacturing company)</i>	Anggrahini, D. and Kurniati, N. and Sukma, A.P.	2021
Pré-qualificação	Associação	i-PM	<i>Using data mining synergies for</i>	Jain, Rajeev and Singh, A.	2012

			<i>evaluating criteria at pre-qualification stage of supplier selection</i>	<i>R. and Yadav, H. C. and Mishra, P. K.</i>	
	Classificaçã o	<i>Support Vector Machine (SVM)</i>	<i>Machine learning approach for finding business partners and building reciprocal relationships</i>	<i>Mori, Junichiro and Kajikawa, Yuya and Kashima, Hisashi and Sakata, Ichiro</i>	2012
Segmentação	Classificaçã o	<i>Adaboost</i>	<i>A Hybrid DEA-Adaboost Model in Supplier Selection for Fuzzy Variable and Multiple Objectives</i>	<i>Cheng, Yijun and Peng, Jun and Zhou, Zhuofu and Gu, Xin and Liu, Weirong</i>	2017
		<i>Decision Trees; Neural Networks</i>	<i>Supplier selection: A hybrid model using DEA, decision tree and neural network</i>	<i>Wu, Desheng</i>	2008
			<i>Designing a supplier evaluation model in the cheese industry using hybrid method</i>	<i>Wahyudi, S. and Asrol, M.</i>	2022
	<i>K-Nearest Neighbors (KNN); Logistic Regression; Naïve Bayes; Decision Tree; support Vector Machine (SVM)</i>		<i>Supervised machine learning approach for effective supplier classification</i>	<i>Harikrishnaku mar, R. and Dand, A. and Nannapaneni, S. and Krishnan, K.</i>	2019

		<i>K-Nearest Neighbors (KNN)</i>	<i>A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing</i>	<i>Cavalcante, I.M. and Frazzon, E.M. and Forcellini, F.A. and Ivanov, D.</i>	2019
		<i>Neural Networks</i>	<i>Using artificial neural networks and analytic hierarchy process for the supplier selection problem</i>	<i>Kar, A.K.</i>	2013
		<i>Random Forest</i>	<i>Identification of Core Suppliers Based on E-Invoice Data Using Supervised Machine Learning</i>	<i>Hong, Jung-sik and Yeo, Hyeongyu and Cho, Nam-Wook and Ahn, Taeuk</i>	2018
		<i>Support Vector Machine (SVM)</i>	<i>Study on the application of SVM in supplier primary election</i>	<i>Cai, L. and Song, F. and Yuan, D.</i>	2008
	Clusterizaçã o	<i>Fuzzy C-Means</i>	<i>A clustering algorithm FCM-ACO for supplier base management</i>	<i>Liu, W. and Jiang, L.</i>	2010
		<i>K-Means; Association Rule Mining</i>	<i>A multi-agent infrastructure for enhancing ERP system intelligence</i>	<i>Symeonidis, A.L. and Chatzidimitriou, K.C. and Kehagias, D.</i>	2001

				<i>and Mitkas, P.A.</i>	
		<i>K-Means</i>	<i>Study and realization of supplier business intelligence system for chain supermarket</i>	<i>Xue, H. and Guo, P. and Zhang, H. and Kang, B.</i>	2009
		<i>Projected Clustering</i>	<i>Hybrid-LWM: A linear-model based hybrid clustering algorithm for supplier categorisation</i>	<i>Irfan, D. and Shengchun, D. and Xiaofei, X.</i>	2011
	Regressão	<i>Support Vector Machine (SVM)</i>	<i>An Intelligence- Based Model for Supplier Selection Integrating DataEnvelopment Analysis and Support Vector Machine</i>	<i>Fallahpour, Alireza and Kazemi, Nima and Molani, Mohammad and Nayeri, Sina and Ehsani, Mojtaba</i>	2018
	Associação	<i>K-Means; Association Rule Mining</i>	<i>A multi-agent infrastructure for enhancing ERP system intelligence</i>	<i>Symeonidis, A.L. and Chatzidimitrio w, K.C. and Kehagias, D. and Mitkas, P.A.</i>	2001
Seleção	Associação	<i>Association Rule Mining</i>	<i>An integrated method for finding key suppliers in SCM</i>	<i>Lin, Rong-Ho and Chuang, Chun-Ling and Liou, James J.</i>	2008

				<i>H. and Wu, Guo-Dong</i>	
		<i>A Novel Approach to Find Patterns of Supplier</i>	<i>Xu, Xu and Lin, Jie</i>	2009	
	<i>Self-Organizing Maps (SOM); Decision Tree; Bayesian Networks</i>	<i>Supplier selection in webinar supply chain using self-organizing maps and data mining</i>	<i>Hosseini, R. and Saryazdi, N.G. and Golpayegani, S.A.H.</i>	2011	
Classificaçã o	<i>Decision Tree; Support Vector Machine (SVM)</i>	<i>Integration of machine learning and mathematical programming methods into the biomass feedstock supplier selection process</i>	<i>Mirkouei, A. and Haapala, K.R.</i>	2014	
	<i>Decision Trees</i>	<i>Supplier Evaluation Model in Freight Activity Microsimulation Estimator</i>	<i>Mahmoudifard , Seyed Mehdi and Shabanpour, Ramin and Golshani, Nima and Mohamma dian, Kiana and Mohammadian , Abolfazi</i>	2018	
	<i>Fuzzy Neural Network</i>	<i>Intelligent decision making for service</i>	<i>Ren, Weibo and Wu,</i>	2019	

			<i>providers selection in maintenance service network: An adaptive fuzzy-neuro approach</i>	<i>Kezhong and Gu, Qiusheng and Hu, Yaoguang</i>	
			<i>Supply and demand matching model for third party logistics integrated platform</i>	<i>Ju, C. and Sun, B.</i>	2009
			<i>Integration of particle swarm optimization-based fuzzy neural network and artificial neural network for supplier selection</i>	<i>Kuo, R.J. and Hong, S.Y. and Huang, Y.C.</i>	2010
	<i>Neural Networks</i>		<i>A hybrid group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network</i>	<i>Kar, A.K.</i>	2014
	<i>Random Forest; Bayesian Networks; Repeated Incremental Pruning to Produce Error Reduction (RIPPER); Naïve Bayes</i>		<i>Supplier classification by applying AutoML</i>	<i>Hernandez-Sanchez, D.-L. and Sanchez-Delacruz, E. and Loeza-Mejia, C.-I. and Hernandez-Gonzalez, A.-G.</i>	2022

		<i>Random Forest</i>	<i>A decision support system for classifying supplier selection criteria using machine learning and random forest approach</i>	<i>Ali, M.R. and Nipu, S.M.A. and Khan, S.A.</i>	2023
		<i>Semi-Fuzzy Support Vector Domain Description</i>	<i>Integration of semi-fuzzy SVDD and CC- Rule method for supplier selection</i>	<i>Guo, Xuesong and Zhu, Zhengwei and Shi, Jia</i>	2013
		<i>Neural Networks; Decision Trees; K-Nearest Neighbors (KNN)</i>	<i>A knowledge-based supplier intelligence retrieval system for outsource manufacturing</i>	<i>Choy, K.L. and Lee, W.B. and Lau, H.C.W. and Choy, L.C.</i>	2005
		<i>Decision Trees; Neural Networks; Logistic Regression</i>	<i>A hybrid ensemble and AHP approach for resilient supplier selection</i>	<i>Hosseini, Seyedmohsen and Al Khaled, Abdullah</i>	2016
		<i>Self-Organizing Maps (SOM); Decision Trees; Bayesian Networks</i>	<i>Supplier selection in webinar supply chain using self-organizing maps and data mining</i>	<i>Hosseini, R. and Saryazdi, N.G. and Golpayegani, S.A.H.</i>	2011
	Clusterizaçāo	<i>Fuzzy C-Means</i>	<i>Supplier Selection: A Hybrid Approach Using ELECTRE and Fuzzy Clustering</i>	<i>Azadnia, Amir Hossein and Ghadimi, Pezhman and Saman, Muhamad</i>	2011

				Zameri Mat and Wong, Kuan Yew and Sharif, Safian	
	<i>Hierarchical Clustering</i>	<i>Partner selection in virtual enterprises: a multi-criteria decision support approach</i>		Crispim, Jose Antonio and de Sousa, Jorge Pinho	2009
	<i>K-Means</i>	<i>An improved K- means algorithm for supplier evaluation and recommendation of purchase and supply platform</i>		Li, J. and Qiu, W. and Li, W.	2020
		<i>CLUS-MCDA: A novel framework based on cluster analysis and multiple criteria decision theory in a supplier selection problem</i>		Maghsoodi, Abteen Ijadi and Kavian, Azad and Khalilzadeh, Mohammad and Brauers, Willem K. M.	2018
	<i>Self-Organizing Maps (SOM); Decision Trees; Bayesian Networks</i>	<i>Supplier selection in webinar supply chain using self-organizing maps and data mining</i>		Hosseini, R. and Saryazdi, N.G. and Golpayegani, S.A.H.	2011
	Regressão	<i>Artificial Neural Networks; Decision Trees; K-Nearest</i>	<i>A knowledge-based supplier intelligence retrieval system for outsource manufacturing</i>	Choy, K.L. and Lee, W.B. and Lau, H.C.W. and Choy, L.C.	2005

		<i>Neighbors (KNN)</i>			
		<i>Autoregressive Integrated Moving Average (ARIMA); Neural Networks</i>	<i>Supplier selection and order allocation planning using predictive analytics and multi-objective programming</i>	<i>Islam, Samiul and Amin, Saman Hassanzadeh and Wardley, Leslie J.</i>	2022
		<i>Neural Networks; Support Vector Machine (SVM)</i>	<i>Comparison on neural networks and support vector machines in suppliers' selection</i>	<i>Guosheng, H. and Guohong, Z.</i>	2006
		<i>Decision Trees; Neural Networks; Logistic Regression</i>	<i>A hybrid ensemble and AHP approach for resilient supplier selection</i>	<i>Hosseini, Seyedmohsen and Al Khaled, Abdullah</i>	2016
		<i>Gradient Boosting</i>	<i>Resilient and sustainable supplier selection: an integration of SCOR 4.0 and machine learning approach</i>	<i>Khan, Md Muzahid and Bashar, Imranul and Minhaj, Golam Morshed and Wasi, Absar Ishraq and Hossain, Niamat Ullah Ibne</i>	2023
		<i>Logistic Regression</i>	<i>Optimal Strategy for Supplier Selection in a Global Supply</i>	<i>Gaida, Itoua Wanck Eyika and Mittal,</i>	2022

			<i>Chain Using Machine Learning Technique</i>	<i>Mandeep and Yadav, Ajay Singh</i>	
	<i>Neural Networks</i>		<i>The application of data mining and BP neural network in supplier selection</i>	<i>Shi, C.-D. and Chen, J.-H. and Sun, Q.-X.</i>	2008
			<i>Study on supplier selection based on AHP and BP neural network</i>	<i>Bin, X. and Bin, P.</i>	2008
			<i>Supplier selection based on a neural network model using genetic algorithm</i>	<i>Golmohammad i, D. and Creese, R.C. and Valian, H. and Kolassa, J.</i>	2009
			<i>Design of an intelligent supplier relationship management system: a hybrid case based neural network approach</i>	<i>Choy, KL and Lee, WB and Lo, V</i>	2003
			<i>Evaluation of material suppliers based on BP neural network under the background of big data</i>	<i>Hou, H. and Meng, H.</i>	2020
			<i>Business transaction recommendation for discovering potential</i>	<i>Lee, Donghun and Kim, Kwanho</i>	2022

			<i>business partners using deep learning</i>		
			<i>A locally linear neuro-fuzzy model for supplier selection in cosmetics industry</i>	<i>Vahdani, B. and Iranmanesh, S.H. and Mousavi, S.M. and Abdollahzade, M.</i>	2011
			<i>Selection of suppliers based on BP neutral networks and grey correlation analysis</i>	<i>Chen, K. and Xuan, Z. and Shang, X.</i>	2009
			<i>Forecasting contractor performance using a neural network and genetic algorithm in a pre-qualification model</i>	<i>El-Sawalhi, N. and Eaton, D. and Rustom, R.</i>	2008
			<i>Relational Regressor Chain (RRC); Holt's Linear Trend (HLT); Autoregressive Integrated Moving Average (ARIMA)</i>	<i>Machine learning and optimization models for supplier selection and order allocation planning</i>	<i>Islam, Samiul and Amin, Saman Hassanzadeh and Wardley, Leslie J.</i>
			<i>Support Vector Machine</i>	<i>A Hybrid Supplier Selection Approach</i>	<i>Aggarwal, I. and Gunreddy,</i>

		(SVM); <i>Decision Trees;</i> <i>Logistic</i> <i>Regression</i>	<i>Using Machine</i> <i>Learning and Data</i> <i>Envelopment</i> <i>Analysis</i>	<i>N. and Rajan,</i> <i>A.J.</i>	
--	--	--	---	-------------------------------------	--

Fonte: Elaborado pela autora.

A Tabela 13 apresenta a quantidade de publicações para cada tipo de abordagem SRM e abordagem de *Machine Learning*, facilitando a visualização das distribuições.

Tabela 13 - Divisão dos artigos entre etapas do processo SRM e tarefas de machine learning com quantidade

Etapa do processo SRM	Tarefa de Machine Learning	Quantidade	Total
Avaliação	Associação	4	21
	Classificação	7	
	Clusterização	6	
	Regressão	4	
Desenvolvimento	Associação	1	2
	Classificação	0	
	Regressão	0	
	Clusterização	1	
Pré-qualificação	Associação	1	2
	Clusterização	0	
	Regressão	0	
	Classificação	1	

Segmentação	Associação	1	14
	Classificação	8	
	Clusterização	4	
	Regressão	1	
Seleção	Associação	3	37
	Classificação	11	
	Clusterização	5	
	Regressão	18	

Fonte: Elaborado pela autora.

### 5.3 Discussão dos resultados

Foram analisadas 67 publicações com o objetivo de investigar a aplicação de técnicas de aprendizado de máquina no contexto dos processos de SRM para identificar os principais objetivos de utilização dessas técnicas.

No contexto de avaliação de fornecedores, seria esperado que abordagens de aprendizado de máquina voltadas à regressão fossem predominantes, especialmente devido a capacidade desse tipo de técnica em prever a performance e desempenho futuro de fornecedores com base em critérios como, como qualidade, custo e eficiência.

De fato, foram utilizados modelos preditivos baseados em redes neurais para atribuição de pontuações de desempenho e a estimativa da eficiência dos fornecedores, por meio de notas, reforçando a eficácia da regressão nesse contexto. No entanto, as técnicas de classificação foram as mais recorrentes, principalmente em contextos que apresentaram a necessidade de categorizar fornecedores em grupos de risco, avaliar a credibilidade de dados subjetivos e lidar com informações textuais e emocionais, como as extraídas de diálogos e opiniões de especialistas.

Muitas das avaliações envolveram classificar fornecedores em categorias de risco ou credibilidade, onde técnicas de classificação, como *decision trees* e redes neurais, se destacam por sua capacidade de lidar com múltiplos atributos qualitativos e subjetivos. Além disso, a análise emocional de textos e diálogos, uma tarefa complexa e de natureza subjetiva, tem demandado a aplicação de técnicas que classifiquem os fornecedores com base em sentimentos e opiniões, um domínio tipicamente reservado às abordagens de classificação.

De maneira geral, as aplicações convergiram para uma previsão mais precisa e abrangente do desempenho dos fornecedores, otimizando o processo de avaliação como um todo.

Na etapa de desenvolvimento do relacionamento com fornecedores, o uso de *machine learning*, especificamente através da mineração de regras de associação e clusterização, teve como objetivo otimizar esse processo, levando em consideração o seu desempenho. A mineração de regras de associação permitiu identificar padrões ocultos entre *feedback* dos clientes e o desempenho dos fornecedores, auxiliando na retenção de clientes ao mesmo tempo que ajusta e melhora a rede de fornecedores existente. Já o *K-Means*, técnica de clusterização utilizada, tem como objetivo segmentar fornecedores com base em critérios de desempenho, agrupando-os em clusters que possibilitam à empresa tomar decisões mais informadas e estratégicas, como a aplicação de estratégias de gerenciamento específicas para cada grupo com o objetivo de desenvolver os mesmos.

No contexto da pré-qualificação, técnicas de mineração de processos, como o i-PM, e algoritmos de classificação, como o *Support Vector Machine* (SVM), foram os utilizados pelos autores para lidar com a complexidade e incerteza associadas ao processo. O i-PM identifica associações ocultas nos dados de pré-qualificação e define os critérios essenciais para a classificação de fornecedores. O SVM por sua vez é utilizado para prever se as relações entre fornecedores e compradores serão mutuamente benéficas, classificando-as em pares que têm potencial de sucesso ou não. Ambas as técnicas são aplicadas com o objetivo de permitir uma construção de relações estratégicas mais eficazes, seja na tentativa de identificar regras de relacionamentos de sucesso ou na tentativa de prever pares de cliente e fornecedor que possuem potencial.

Em relação à segmentação de fornecedores, as técnicas de aprendizado de máquina, especialmente classificação e clusterização, têm sido amplamente empregadas com objetivos distintos, dependendo da abordagem utilizada. No caso das técnicas de clusterização, como o *k-means*, o objetivo é segregar fornecedores em grupos distintos com base em características específicas, permitindo uma análise detalhada dos perfis de cada cluster. Essa segmentação

facilita o gerenciamento e a identificação de fornecedores com características semelhantes, possibilitando uma análise mais precisa e escalável, com o objetivo de melhorar a eficiência e a precisão da segmentação.

Por outro lado, as abordagens de segmentação que utilizaram técnicas de classificação têm como foco a categorização direta dos fornecedores baseada em critérios, como eficiência, qualidade ou resiliência. Em alguns casos, a classificação entre fornecedores eficientes e ineficientes pode ser feita a partir de saídas do DEA (*Data Envelopment Analysis*), uma abordagem MCDM. A utilização de classificações também tem o objetivo de superar a subjetividade inerente ao julgamento humano. Dessa forma, a segmentação com auxílio de técnicas de *machine learning* permitem uma tomada de decisão mais objetiva e menos sujeita a vieses.

Por fim, no processo de seleção de fornecedores, as técnicas de regressão tiveram como principais objetivos em suas aplicações prever o desempenho dos fornecedores baseado em dados históricos. Modelos como redes neurais e metodologias híbridas são amplamente utilizados para capturar padrões complexos e atribuir pesos aos critérios de seleção, otimizando esse processo e reduzindo a subjetividade. Além disso, essas técnicas também foram utilizadas com o objetivo de simular o julgamento dos gestores e fornecer uma avaliação menos subjetiva.

As técnicas de classificação, por sua vez, são empregadas para categorizar os fornecedores com base em critérios quantitativos e qualitativos. Abordagens como SVM e *Random Forest* são as mais frequentemente usadas para identificar os fornecedores mais adequados. Essas técnicas contribuem para um processo também mais automatizado e transparente, reduzindo a subjetividade e otimizando a tomada de decisão.

Além dos contextos específicos do Gerenciamento do Relacionamento com os fornecedores (SRM), algumas outras discussões podem ser feitas. Aproximadamente 49% das publicações analisadas abordam especificamente essa etapa. Isso reflete a centralidade da seleção eficiente de fornecedores dentro do ciclo do processo de SRM, sendo um dos principais objetivos nesse contexto.

Além da seleção, as etapas de segmentação e avaliação também foram amplamente exploradas, correspondendo a 21% e 25% das publicações, respectivamente. Essas etapas envolvem processos quantitativos diretamente alinhados com as técnicas de aprendizado de máquina, como atribuição de notas, agrupamento e categorização. Em contrapartida, as etapas de desenvolvimento do relacionamento com fornecedores e pré-qualificação foram menos abordadas, representando apenas 3% das publicações cada uma. A baixa recorrência de estudos sobre a pré-qualificação pode ser explicada pelo fato de essa etapa ser menos explorada no

contexto geral de SRM. Já o desenvolvimento de relacionamento, por sua natureza mais subjetiva, apresenta desafios adicionais para a aplicação de técnicas de aprendizado de máquina, o que justifica a menor quantidade de estudos focados nessa etapa.

No que tange às abordagens de aprendizado de máquina, as técnicas de classificação e regressão se destacaram como as mais frequentes, ambas associadas a modelos supervisionados. A preferência por essas abordagens se justifica pela compatibilidade com as demandas do processo de SRM, onde é comum a necessidade de atribuir pontuações e rótulos aos fornecedores. Técnicas de clusterização foram utilizadas apenas para agrupar fornecedores sem a necessidade de rótulos, enquanto as técnicas de associação concentraram-se na extração de regras, com o objetivo principal de reduzir a subjetividade dos processos e permitir decisões mais informadas.

Por fim, no contexto das técnicas de *Multi-Criteria Decision Making* (MCDM), observou-se que a maioria dos métodos analisados não recorreu a essas abordagens como suporte. Entretanto, entre os 35% das publicações que empregaram técnicas MCDM, destacam-se principalmente o AHP e o DEA. O AHP é frequentemente mencionado como a técnica MCDM mais utilizada em processos de SRM, por sua adequação à atribuição de notas e pesos para os critérios de avaliação de fornecedores. Já o DEA se mostra relevante por sua flexibilidade no tratamento de diferentes tipos de dados e sua capacidade de calcular eficiência, características que também se alinham bem ao contexto de SRM.

## 6 CONCLUSÃO

Este estudo revisou o estado da arte da aplicação de técnicas de aprendizado de máquina na gestão de fornecedores, abordando seleção, segmentação, avaliação e desenvolvimento. A revisão da literatura identificou as principais técnicas, como Redes Neurais e Árvores de Decisão, com foco na sua aplicabilidade utilizando conjuntos de dados reais.

No que tange à comparação das abordagens, verificou-se que técnicas supervisionadas predominam, e que os métodos não supervisionados foram consideravelmente menos explorados no contexto de gestão de fornecedores, com ressalva para o processo de segmentação de fornecedores. Desafios como a disponibilidade de dados em larga escala e a interpretabilidade de modelos complexos foram destacados pelos autores, tendo sido consideradas barreiras para uma aplicação mais ampla, visto que implicam em uma dificuldade para a capacidade de treinamento e generalização dos modelos. Além disso, a eficácia das

técnicas de *machine learning* empregadas está frequentemente condicionada à qualidade desses dados.

O estudo aponta que há oportunidades para preencher as lacunas identificadas, como a necessidade de mais pesquisas em técnicas não supervisionadas e *fuzzy logic*, além de uma maior integração com métodos de decisão multicritério (MCDM). Além disso, os processos de pré-qualificação e desenvolvimento de fornecedores foram pouco abordados, representando áreas promissoras para estudos futuros.

Em relação às limitações do estudo, identificou-se ao final da pesquisa que não foram incluídas *strings* de busca diretamente relacionadas à técnica PCA (Análise de Componentes Principais), o que representa uma lacuna na revisão da literatura realizada e uma oportunidade de melhoria para trabalhos futuros sobre o tema. Além disso, a expressão “*supplier clustering*”, comumente utilizada na área, também não foi incluída, o que indica que alguns artigos envolvendo a clusterização de fornecedores podem não ter sido contemplados na revisão devido à ausência dessa *string* específica nos critérios de busca.

Por fim, a síntese das descobertas oferece uma base sólida para que futuras pesquisas possam explorar o tema de forma mais direcionada e eficaz. Em suma, o campo de estudo da aplicação de técnicas de *machine learning* em SRM está em expansão, mas ainda enfrenta limitações a serem superadas e lacunas a serem preenchidas para que seu potencial seja plenamente alcançado. Vale também ressaltar que não foram encontradas pesquisas anteriores que apresentassem uma revisão sistemática da literatura sobre a aplicação de técnicas de *machine learning* no contexto de gestão de fornecedores.

## 7 REFERÊNCIAS

- ACHARYA, M. S.; ARMAAN, A.; ANTONY, A. S. **A comparison of regression models for prediction of graduate admissions.** International Conference on Computational Intelligence in Data Science (ICCIDDS), 2019. p. 1-5.
- BELGIU, Mariana; DRĂGUȚ, Lucian. **Random forest in remote sensing: A review of applications and future directions.** ISPRS journal of photogrammetry and remote sensing, v. 114, p. 24-31, 2016.
- BIAMONTE, Jacob et al. **Quantum machine learning.** Nature, v. 549, n. 7671, p. 195-202, 2017.

- BORA, D. J.; GUPTA, A. K. **A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm.** International Journal of Computer Trends and Technology (IJCTT), vol. 10, no. 2, 2014.
- ARANTES, R. F. M. **Avaliação e segmentação de fornecedores apoiadas por técnicas de aprendizado de máquina e interpretadas por XAI - eXplainable Artificial Intelligence.** 2024. Tese (Doutorado) – Departamento de Engenharia de Produção - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2024.
- BREIMAN, Leo. **Random forests Machine learning**, v. 45, n. 1, p. 5-32, 2001.
- DE BOER, Luitzen; LABRO, Eva; MORLACCHI, Pierangela. **A review of methods supporting supplier selection.** European journal of purchasing & supply management, v. 7, n. 2, p. 75-89, 2001.
- DOS SANTOS, Bruno Samways; STEINER, Maria Teresinha Arns; LIMA, Rafael Henrique Palma. **Proposal of a method to classify female smokers based on data mining techniques.** Computers & Industrial Engineering, v. 170, p. 108363, 2022
- DUBEY, Rameshwar; GUNASEKARAN, Angappa; CHILDE, Stephen J.; PAPADOPOULOS, Thanos; HELO, Petri. **Supplier relationship management for circular economy: influence of external pressures and top management commitment.** Management Decision, v. 57, n. 4, p. 767-790, 2018.
- FERNANDES, Fernando Timoteo. **Machine learning em saúde e segurança do trabalhador: perspectivas, desafios e aplicações.** 2021. Tese de Doutorado. Universidade de São Paulo.
- GALO, Nadya Regina; CALACHE, Lucas Daniel Del Rosso; CARPINETTI, Luiz Cesar Ribeiro. **A group decision approach for supplier categorization based on hesitant fuzzy and ELECTRE TRI.** International Journal of Production Economics, v. 202, p. 182-196, 2018.
- GUO, Yunxia. **Application research of supplier evaluation based on random forest.** Proceedings of the 6th International Conference on Software and Computer Applications. 2017. p. 316-323.
- HÜLLERMEIER, Eyke. **Prescriptive Machine Learning for Automated Decision Making: Challenges and Opportunities.** 2021. JAIN, A. K. Data clustering: 50 years beyond K-means. Pattern Recognition Letters, vol. 31, no. 8, 2010.
- KESHAVARZ GHORABAE, Mehdi; ARIMI, Maghsoud; ZAVADSKAS, Edmundas Kazimieras; ANTUCHEVICIENE, Jurgita. **Supplier evaluation and selection in fuzzy**

- environments: a review of MADM approaches.** Economic research-Ekonomska istraživanja, v. 30, n. 1, p. 1073-1118, 2017.
- KRAUSE, Daniel R.; HANDFIELD, Robert B.; SCANNELL, Thomas V. An empirical investigation of supplier development: reactive and strategic processes.** Journal of operations management, v. 17, n. 1, p. 39-58, 1998.
- KRAUSE, Daniel R.; HANDFIELD, Robert B.; TYLER, Beverly B. The relationships between supplier development, commitment, social capital accumulation and performance improvement.** Journal of operations management, v. 25, n. 2, p. 528-545, 2007.
- KRALJIC, Peter. Purchasing must become supply management.** Harvard business review, v. 61, n. 5, p. 109-117, 1983.
- LIMA-JUNIOR, Francisco Rodrigues; CARPINETTI, Luiz Cesar Ribeiro. Predicting supply chain performance based on SCOR® metrics and multilayer perceptron neural networks.** International Journal of Production Economics, v. 212, p. 19-38, 2019.
- MAXWELL, A. E., WARNER, T. A., & FANG, F. (2018). Implementation of machine-learning classification in remote sensing: an applied review.** Remote Sensing Letters, NGAI, Eric WT; XIU, Li; CHAU, Dorothy CK. **Application of data mining techniques in customer relationship management: A literature review and classification.** Expert systems with applications, v. 36, n. 2, p. 2592-2602, 2009.
- OLSEN, Rasmus Friis; ELLRAM, Lisa M. A portfolio approach to supplier relationships. Industrial marketing management, v. 26, n. 2, p. 101-113, 1997.**
- PARK, Jongkyung; SHIN, Kitae; CHANG, Tai-Woo; PARK, Jinwoo. An integrative framework for supplier relationship management.** Industrial Management & Data Systems, 2010.
- PRAJOGO, Daniel; OLHAGER, Jan. Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration.** International Journal of Production Economics, v. 135, n. 1, p. 514-522, 2012.
- RANDOLPH, Justus. A guide to writing the dissertation literature review.** Practical Assessment, Research, and Evaluation, v. 14, n. 1, p. 13, 2009.
- REZAEI, Jafar; ORTT, Roland. A multi-variable approach to supplier segmentation.** International Journal of Production Research, v. 50, n. 16, p. 4593-4611, 2012 REZAEI, Jafar; WANG, Jing; TAVASSZY, Lori. **Linking supplier development to supplier segmentation using Best Worst Method.** Expert Systems with Applications, v. 42, n. 23, p. 9152-9164, 2015.

- RIYANTO, Slamet et al. **Comparative Analysis using Various Performance Metrics in Imbalanced Data for Multi-class Text Classification.** International Journal of Advanced Computer Science and Applications, v. 14, n. 6, 2023.
- ROOPA, H.; ASHA, T. **A linear model based on principal component analysis for disease prediction.** IEEE Access, vol. 7, pp. 105314-105318, 2019
- RUDBERG, M.; OLHAGER, J. **Manufacturing networks and supply chains: an operations strategy perspective.** OMEGA – The International Journal of Management Science, 31, p.29-39, 2003.
- SANTOS, Luiz Felipe de Oliveira Moura; OSIRO, Lauro; LIMA, Rafael Henrique Palma. **A model based on 2-tuple fuzzy linguistic representation and Analytic Hierarchy Process for supplier segmentation using qualitative and quantitative criteria.** Expert Systems with Applications, v. 79, p. 53-64, 2017
- SEGURA, Marina; MAROTO, Concepción. **A multiple criteria supplier segmentation using outranking and value function methods.** Expert Systems with Applications, v. 69, p. 87-100, 2017.
- SUNDTOFT HALD, K.; ELLEGAARD, C. **Supplier evaluation processes: the shaping and reshaping of supplier performance.** International Journal of Operations & Production Management, v. 31, n. 8, p. 888-910, 2011.
- TRANFIELD, David; DENYER, David; SMART, Palminder. **Towards a methodology for developing evidence-informed management knowledge by means of systematic review.** British journal of management, v. 14, n. 3, p. 207-222, 2003.
- WU, J.; LIU, C.; CUI, W.; ZHANG, Y. **Personalized Collaborative Filtering Recommendation Algorithm based on Linear Regression.** IEEE International Conference on Power Data Science (ICPDS), 2019. p. 139-142.
- YUNXIA, Guo. **Application research of supplier evaluation based on random forest.** Proceedings of the 6th International Conference on Software and Computer Applications. 2017. p. 316-323.
- ZHANG, Haifei; QUOST, Benjamin; MASSON, Marie-Hélène. **Cautious weighted random forests.** Expert Systems with Applications, v. 213, p. 118883, 2023.
- CHENG, Yijun; PENG, Jun; ZHOU, Zhuofu; GU, Xin; LIU, Weirong. **A hybrid DEA-Adaboost model in supplier selection for fuzzy variable and multiple objectives.** IFAC PAPERSONLINE.
- HE, T. **Data mining for the optimization in manufacturing process.** Proceedings - 2010 2nd International Workshop on Intelligent Systems and Applications, ISA 2010.

- XU, Xu; LIN, Jie; XU, Dongming. **Mining pattern of supplier with the methodology of domain-driven data mining.** 2009 IEEE INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS, VOLS 1-3.
- CHOY, K.L.; LEE, W.B.; LAU, H.C.W.; CHOY, L.C. **A knowledge-based supplier intelligence retrieval system for outsource manufacturing.** Knowledge-Based Systems.
- LIN, Rong-Ho; CHUANG, Chun-Ling; LIU, James J. H.; WU, Guo-Dong. **An integrated method for finding key suppliers in SCM.** EXPERT SYSTEMS WITH APPLICATIONS.
- KAPPELMAN, Ashton Conrad; SINHA, Ashesh Kumar. **Optimal control in dynamic food supply chains using big data.** COMPUTERS & OPERATIONS RESEARCH.
- HAERY, A.; SALMASI, N.; MODARRES YAZDI, M.; IRANMESH, H. **Application of association rule mining in supplier selection criteria.** World Academy of Science, Engineering and Technology.
- XU, Xu; LIN, Jie. **A novel approach to find patterns of supplier.** SNPD 2009: 10TH ACIS INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING, ARTIFICIAL INTELLIGENCES, NETWORKING AND PARALLEL DISTRIBUTED COMPUTING, PROCEEDINGS.
- ISLAM, Samiul; AMIN, Saman Hassanzadeh; WARDLEY, Leslie J. **Supplier selection and order allocation planning using predictive analytics and multi-objective programming.** COMPUTERS & INDUSTRIAL ENGINEERING.
- GUOSHENG, H.; GUOHONG, Z. **Comparison on neural networks and support vector machines in suppliers' selection.** Journal of Systems Engineering and Electronics.
- SHI, C.-D.; CHEN, J.-H.; SUN, Q.-X. **The application of data mining and BP neural network in supplier selection.** Proceedings of the International Conference on Computer Science and Information Technology, ICCSIT 2008.
- YURTAY, Yuksel; AYANOGLU, Murat. **Dashboard application model in supplier evaluation by using artificial immune system and data mining methods.** Proceedings of the International Conference on Computer Science and Information Technology, ICCSIT 2008.
- KOLESNIKOVA, D.; ANDREEV, Y.; IUREVA, R. **Estimation of the features influence on cluster partition.** Proceedings of the 18th International Conference on Informatics in Control, Automation and Robotics, ICINCO 2021.
- WU, Desheng. **Supplier selection: A hybrid model using DEA, decision tree and neural network.** EXPERT SYSTEMS WITH APPLICATIONS.
- WAHYUDI, S.; ASROL, M. **Designing a supplier evaluation model in the cheese industry using hybrid method.** Academic Journal of Manufacturing Engineering.

- HOSSEINI, Seyedmohsen; AL KHALED, Abdullah. **A hybrid ensemble and AHP approach for resilient supplier selection.** JOURNAL OF INTELLIGENT MANUFACTURING.
- MIRKOUEI, A.; HAAPALA, K.R. **Integration of machine learning and mathematical programming methods into the biomass feedstock supplier selection process.** FAIM 2014 - Proceedings of the 24th International Conference on Flexible Automation and Intelligent Manufacturing: Capturing Competitive Advantage via Advanced Manufacturing and Enterprise Transformation.
- MAHMOUDIFARD, Seyed Mehdi; SHABANPOUR, Ramin; GOLSHANI, Nima; MOHAMMADIAN, Kiana; MOHAMMADIAN, Abolfazi. **Supplier Evaluation Model in Freight Activity Microsimulation Estimator.** TRANSPORTATION RESEARCH RECORD.
- LIU, W.; JIANG, L. **A clustering algorithm FCM-ACO for supplier base management.** Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).
- AZADNIA, Amir Hossein; GHADIMI, Pezhman; SAMAN, Muhamad Zameri Matand; WONG, Kuan Yew; SHARIF, Safian. **Supplier Selection: A Hybrid Approach Using ELECTRE and Fuzzy Clustering.** INFORMATICS ENGINEERING AND INFORMATION SCIENCE, PT II.
- REN, Weibo; WU, Kezhong; GU, Qiusheng; HU, Yaoguang. **Intelligent decision making for service providers selection in maintenance service network: An adaptive fuzzy-neuro approach.** KNOWLEDGE-BASED SYSTEMS.
- JU, C.; SUN, B. **Supply and demand matching model for third party logistics integrated platform.** IJCAI International Joint Conference on Artificial Intelligence.
- KUO, R.J.; HONG, S.Y.; HUANG, Y.C. **Integration of particle swarm optimization-based fuzzy neural network and artificial neural network for supplier selection.** Applied Mathematical Modelling.
- KHAN, Md Muzahid; BASHAR, Imranul; MINHAJ, Golam Morshed; WASI, Absar Ishraq; HOSSAIN, Niamat Ullah Ibne. **Resilient and sustainable supplier selection: an integration of SCOR 4.0 and machine learning approach.** SUSTAINABLE AND RESILIENT INFRASTRUCTURE.
- CRISPIM, Jose Antonio; DE SOUSA, Jorge Pinho. **Partner selection in virtual enterprises: a multi-criteria decision support approach.** INTERNATIONAL JOURNAL OF PRODUCTION RESEARCH.

- JAIN, Rajeev; SINGH, A. R.; YADAV, H. C.; MISHRA, P. K. **Using data mining synergies for evaluating criteria at pre-qualification stage of supplier selection.** JOURNAL OF INTELLIGENT MANUFACTURING.
- SYMEONIDIS, A.L.; CHATZIDIMITRIOU, K.C.; KEHAGIAS, D.; MITKAS, P.A. **A multi-agent infrastructure for enhancing ERP system intelligence.** Scalable Computing.
- LI, J.; QIU, W.; LI, W. **An improved K-means algorithm for supplier evaluation and recommendation of purchase and supply platform.** Journal of Physics: Conference Series.
- VAHDANI, B.; MOUSAVI, S.M.; TAVAKKOLI-MOGHADDAM, R.; HASHEMI, H. **A new enhanced support vector model based on general variable neighborhood search algorithm for supplier performance evaluation: A case study.** International Journal of Computational Intelligence Systems.
- MAGHSOODI, Abteen Ijadi; KAVIAN, Azad; KHALILZADEH, Mohammad; BRAUERS, Willem K. M. **CLUS-MCDA: A novel framework based on cluster analysis and multiple criteria decision theory in a supplier selection problem.** COMPUTERS & INDUSTRIAL ENGINEERING.
- ANGGRIHINI, D.; KURNIATI, N.; SUKMA, A.P. **Data analytics for transforming towards smart supplier relationship management (A case study in manufacturing company).** Proceedings of the International Conference on Industrial Engineering and Operations Management.
- CAMMARANO, A.; VARRIALE, V.; MICHELINO, F.; CAPUTO, M. **A patent-based tool to support component suppliers assessment in the smartphone supply chain.** IEEE Transactions on Engineering Management.
- XUE, H.; GUO, P.; ZHANG, H.; KANG, B. **Study and realization of supplier business intelligence system for chain supermarket.** Proceedings - 2009 International Conference on Computational Intelligence and Software Engineering, CiSE 2009.
- HARIKRISHNAKUMAR, R.; DAND, A.; NANNAPANENI, S.; KRISHNAN, K. **Supervised machine learning approach for effective supplier classification.** Proceedings - 18th IEEE International Conference on Machine Learning and Applications, ICMLA 2019.
- CAVALCANTE, I.M.; FRAZZON, E.M.; FORCELLINI, F.A.; IVANOV, D. **A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing.** International Journal of Information Management.
- GOLPIRA, Heris. **A novel Multiple Attribute Decision Making approach based on interval data using U2P-Miner algorithm.** DATA & KNOWLEDGE ENGINEERING.

- GAIDA, Itoua Wanck Eyika; MITTAL, Mandeep; YADAV, Ajay Singh. **Optimal Strategy for Supplier Selection in a Global Supply Chain Using Machine Learning Technique.** INTERNATIONAL JOURNAL OF DECISION SUPPORT SYSTEM TECHNOLOGY.
- BODENDORF, Frank; WYTOPIL, Benedict; FRANKE, Joerg. **Business Analytics in Strategic Purchasing: Identifying and Evaluating Similarities in Supplier Documents.** APPLIED ARTIFICIAL INTELLIGENCE.
- QUAN, Quan; ZHANG, Zhongqiang. **Supply Capability Evaluation of Intelligent Manufacturing Enterprises Based on Improved BP Neural Network.** JOURNAL OF MATHEMATICS.
- HA, S.H.; KRISHNAN, R. **A hybrid approach to supplier selection for the maintenance of a competitive supply chain.** Expert Systems with Applications.
- BIN, X.; BIN, P. **Study on supplier selection based on AHP and BP neural network.** Proceedings - 1st International Workshop on Knowledge Discovery and Data Mining, WKDD.
- GOLMOHAMMADI, D.; CREESE, R.C.; VALIAN, H.; KOLASSA, J. **Supplier selection based on a neural network model using genetic algorithm.** IEEE Transactions on Neural Networks.
- CHOY, K.L.; LEE, W.B.; LO, V. **Design of an intelligent supplier relationship management system: a hybrid case based neural network approach.** EXPERT SYSTEMS WITH APPLICATIONS.
- HOU, H.; MENG, H. **Evaluation of material suppliers based on BP neural network under the background of big data.** Proceedings - 2020 Management Science Informatization and Economic Innovation Development Conference, MSIED 2020.
- KAR, A.K. **Using artificial neural networks and analytic hierarchy process for the supplier selection problem.** 2013 IEEE International Conference on Signal Processing, Computing and Control, ISPCC 2013.
- LEE, Donghun; KIM, Kwanho. **Business transaction recommendation for discovering potential business partners using deep learning.** EXPERT SYSTEMS WITH APPLICATIONS.
- ZHAO, K.; YU, X. **A case based reasoning approach on supplier selection in petroleum enterprises.** Expert Systems with Applications.
- VAHDANI, B.; IRANMESH, S.H.; MOUSAVI, S.M.; ABDOLLAHZADE, M. **A locally linear neuro-fuzzy model for supplier selection in cosmetics industry.** Applied Mathematical Modelling.

- CHEN, K.; XUAN, Z.; SHANG, X. **Selection of suppliers based on BP neural networks and grey correlation analysis.** IJCAI International Joint Conference on Artificial Intelligence.
- EL-SAWALHI, N.; EATON, D.; RUSTOM, R. **Forecasting contractor performance using a neural network and genetic algorithm in a pre-qualification model.** Construction Innovation.
- KAR, A.K. **A hybrid group decision support system for supplier selection using analytic hierarchy process, fuzzy set theory and neural network.** Journal of Computational Science.
- WAN, Wei; LIU, Yuanlong; HAN, Xingwang; WANG, Huijian. **Evaluation Model of Power Operation and Maintenance Based on Text Emotion Analysis.** MATHEMATICAL PROBLEMS IN ENGINEERING.
- IRFAN, D.; SHENGCHUN, D.; XIAOFEI, X. **Hybrid-LWM: A linear-model based hybrid clustering algorithm for supplier categorisation.** International Journal of Systems, Control and Communications.
- HERNANDEZ-SANCHEZ, D.-L.; SANCHEZ-DELACRUZ, E.; LOEZA-MEJIA, C.-I.; HERNANDEZ-GONZALEZ, A.-G. **Supplier classification by applying AutoML.** CCE 2022 - 2022 19th International Conference on Electrical Engineering, Computing Science and Automatic Control.
- HONG, Jung-sik; YEO, Hyeongyu; CHO, Nam-Wook; AHN, Taeuk. **Identification of Core Suppliers Based on E-Invoice Data Using Supervised Machine Learning.** JOURNAL OF RISK AND FINANCIAL MANAGEMENT.
- ALI, M.R.; NIPU, S.M.A.; KHAN, S.A. **A decision support system for classifying supplier selection criteria using machine learning and random forest approach.** Decision Analytics Journal.
- ÇELEBI, D.; BAYRAKTAR, D. **An integrated neural network and data envelopment analysis for supplier evaluation under incomplete information.** Expert Systems with Applications.
- ISLAM, Samiul; AMIN, Saman Hassanzadeh; WARDLEY, Leslie J. **Machine learning and optimization models for supplier selection and order allocation planning.** INTERNATIONAL JOURNAL OF PRODUCTION ECONOMICS.
- HOSSEINI, R.; SARYAZDI, N.G.; GOLPAYEGANI, S.A.H. **Supplier selection in webinar supply chain using self-organizing maps and data mining.** Proceedings of the 1st International Technology Management Conference, ITMC 2011.
- GUO, Xuesong; ZHU, Zhengwei; SHI, Jia. **Integration of semi-fuzzy SVDD and CC-Rule method for supplier selection.** EXPERT SYSTEMS WITH APPLICATIONS.

- AGGARWAL, I.; GUNREDDY, N.; RAJAN, A.J. **A Hybrid Supplier Selection Approach Using Machine Learning and Data Envelopment Analysis**. 3rd IEEE International Virtual Conference on Innovations in Power and Advanced Computing Technologies, i-PACT 2021.
- MORI, Junichiro; KAJIKAWA, Yuya; KASHIMA, Hisashi; SAKATA, Ichiro. **Machine learning approach for finding business partners and building reciprocal relationships**. EXPERT SYSTEMS WITH APPLICATIONS.
- WANG, H.; WANG, H.; HU, J.; XU, J.; LI, J.; HE, B. **Evaluation Model of Power Operation and Maintenance Based on Text Emotion Analysis**. Dianli Zidonghua Shebei/Electric Power Automation Equipment.
- CAI, L.; SONG, F.; YUAN, D. **Study on the application of SVM in supplier primary election**. Proceedings - 1st International Workshop on Knowledge Discovery and Data Mining, WKDD.
- SU, Chuan-Jun; CHEN, Yin-An. **Risk assessment for global supplier selection using text mining**. COMPUTERS & ELECTRICAL ENGINEERING.
- FALLAHPOUR, Alireza; KAZEMI, Nima; MOLANI, Mohammad; NAYERI, Sina; EHSANI, Mojtaba. **An Intelligence-Based Model for Supplier Selection Integrating Data Envelopment Analysis and Support Vector Machine**. IRANIAN JOURNAL OF MANAGEMENT STUDIES.
- ZHAO, L.; QI, W.; ZHU, M. **A Study of Supplier Selection Method Based on SVM for Weighting Expert Evaluation**. Discrete Dynamics in Nature and Society.
- LIOU, James J. H.; CHANG, Mu-Hsin; LO, Huai-Wei; HSU, Min-Hsi. **Application of an MCDM model with data mining techniques for green supplier evaluation and selection**. APPLIED SOFT COMPUTING.