



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS FLORIANÓPOLIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

Ian Moreira Cavalcante

**SELEÇÃO DE FORNECEDORES POR MEIO DE SIMULAÇÃO E MACHINE
LEARNING COM AS FERRAMENTAS KNN E LR**

Florianópolis
2019

Ian Moreira Cavalcante

**SELEÇÃO DE FORNECEDORES POR MEIO DE SIMULAÇÃO E MACHINE
LEARNING COM AS FERRAMENTAS KNN E LR**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Santa Catarina para a obtenção do título de mestre em engenharia mecânica.
Orientador: Prof. Fernando Antônio Forcellini, Dr.
Coorientador: Prof. Enzo Morosini Frazzon, Dr.-Ing.

Florianópolis
2019

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Cavalcante, Ian Moreira

Seleção de fornecedores por meio de simulação e machine learning com as ferramentas KNN e LR / Ian Moreira Cavalcante ; orientador, Fernando Antônio Forcellini, coorientador, Enzo Morosini Frazzon, 2019.

98 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica, Florianópolis, 2019.

Inclui referências.

1. Engenharia Mecânica. 2. Indústria 4.0. 3. Machine Learning. 4. Simulação. 5. Seleção de Fornecedores. I. Forcellini, Fernando Antônio. II. Frazzon, Enzo Morosini. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Mecânica. IV. Título.

Ian Moreira Cavalcante

**SELEÇÃO DE FORNECEDORES POR MEIO DE SIMULAÇÃO E MACHINE
LEARNING COM AS FERRAMENTAS KNN E LR**

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Fábio Antônio Xavier, Dr.
Universidade Federal de Santa Catarina

Prof. Helio Aisenberg Ferenhof, Dr.
Universidade Federal de Santa Catarina

Prof. Milton Pereira, Dr.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em engenharia mecânica.

Prof. Jonny Carlos da Silva, Dr.
Coordenador do Programa

Prof. Fernando Antônio Forcellini, Dr.
Orientador

Prof. Enzo Morosini Frazzon, Dr.-Ing.
Coorientador

Florianópolis, 29 de Abril de 2019.

Dedico aos que sonham com o protagonismo brasileiro
frente aos desafios da manufatura do século XXI.

AGRADECIMENTOS

Inicio os agradecimentos creditando o sucesso deste trabalho à minha namorada Rafaella Bisineli. Por deixar de ser arquiteta por um ano e abraçar o desafio de residir na Alemanha durante o ano de 2018. Todos os dias, a sua presença me proporcionou sensação de conforto, carinho e alegria. No fim das contas, a verdade é que nossas economias se foram (*is mir egal*), mas vivemos a melhor experiência das nossas vidas.

Agradeço à minha família, por ter a mentalidade de que estudar nunca é custo e sim investimento. Valor este que, sem dúvida, permanecerá na minha vida e passarei adiante. Agradeço também aos professores Fernando Forcellini e Enzo Frazzon, pela orientação e autonomia para desenvolver o trabalho. Foram dois anos intensos e de muito trabalho duro, os quais me senti motivado cada minuto a executar o melhor possível. Me sinto privilegiado de ter absorvido um pouco da experiência de vocês e, conseqüentemente, ter adquirido um modo único de enxergar a manufatura.

Agradeço ao colega Felipe Martinhuk, pelos valiosos debates acerca da manufatura digital e por sonhar junto por um país tecnologicamente independente. Aos colegas da UFSC, principalmente dos laboratórios GEPPS e ProLogIS, pelo acolhimento e auxílio no início do meu mestrado em Florianópolis, assim como quando iniciei os trabalhos na Alemanha.

Por fim, agradeço a todos os pesquisadores que produzem conhecimento com seriedade, principalmente àqueles que foram a base para a minha pesquisa.

*"All models are wrong, but some are useful."
(George Box, 1979)*

RESUMO

Esta pesquisa discute e aplica conceitos e ferramentas da Indústria 4.0 para resolver um problema clássico da manufatura – a seleção de fornecedores. Para tal, são utilizadas de maneira integrada ferramentas de simulação e *machine learning*, com o objetivo de desenvolver um portfólio de fornecedores com desempenho superior em função da performance de entregas no prazo. São teorizados os potenciais benefícios desta nova abordagem e as consequências que novos serviços orientados a dados podem promover na manufatura. Os algoritmos de *machine learning* utilizados foram a regressão logística e k-vizinhos mais próximos. Além disso, é proposto um método para integração dos modelos de simulação e de *machine learning*. Também foram apresentadas duas abordagens que combinam ambos algoritmos por meio de um sistema baseado em regras, as quais apresentaram performance superior à utilização dos algoritmos de maneira isolada. Portanto, esta pesquisa contribuiu com o conhecimento em manufatura digital pelo fato de investigar quantitativamente ferramentas que têm sido utilizadas no novo paradigma da manufatura, o qual é fortemente influenciado por decisões orientadas a dados.

Palavras-chave: Indústria 4.0. Machine Learning. Simulação. Seleção de Fornecedores.

ABSTRACT

This research applies Industry 4.0 concepts to solve a classic manufacturing problem – the supplier selection. Simulation and machine learning tools are used in an integrated manner, aiming to develop a portfolio of suppliers with superior performance in relation to the on-time delivery key performance indicator. The potential benefits of this novel approach and the consequences that new data-driven services can have on manufacturing are theorized. The machine learning algorithms used were logistic regression and k-nearest neighbors. It is also presented two approaches that combine both algorithms by means of a rule-based system, which presented superior performance to the individual use of the algorithms. Therefore, this research has contributed to digital manufacturing knowledge, which is strongly influenced by data-driven decisions, by quantitatively investigating tools that have been used in the new manufacturing paradigm – simulation and machine learning.

Keywords: Industry 4.0. Machine Learning. Simulation. Supplier Selection.

LISTA DE FIGURAS

Figura 1 – Comparação entre documentos científicos publicados e país de publicação com relação às palavras-chave <i>Industry 4.0</i> e <i>Smart Manufacturing</i>	21
Figura 2 – Número de publicações em função do tempo dos termos <i>Industry 4.0</i> e <i>Smart Manufacturing</i>	22
Figura 3 – Comparação das buscas relativas no website Google com relação às palavras-chave <i>Industry 4.0</i> e <i>Smart Manufacturing</i> utilizando a ferramenta Google Trends.	23
Figura 4 – Agregação valor aos sistemas produtivos no contexto da Indústria 4.0.	23
Figura 5 – Sistema ERP e suas conexões com os diversos departamentos da empresa.	24
Figura 6 – O sistema MRP e suas conexões com outras atividades em uma empresa de manufatura.	24
Figura 7 – <i>Framework</i> simplificado da integração dos sistemas que envolvem a programação e controle da produção em uma empresa de manufatura.	25
Figura 8 – <i>Framework</i> do fluxo de informações em uma empresa de serviços. .	26
Figura 9 – Número de publicações em função do tempo dos termos que designam estratégias de orientação ao serviço em combinação à palavra <i>manufacturing</i> , utilizando a ferramenta de pesquisa Scopus.	27
Figura 10 – Quantidade de publicações das palavras Servitização, Sistema produto-serviço, Manufatura e Dados durante o período de 2010 à 2018.	33
Figura 11 – Quantidade de publicações das palavras Serviço, Manufatura e Dados ao longo do tempo durante o período de 2010 à 2018.	33
Figura 12 – Método para visualização de redes bibliométricas com auxílio do software VOSViewer.	35
Figura 13 – Compatibilização das palavras-chave com o mesmo significado. . .	36
Figura 14 – Quantidade em relação à co-ocorrência entre palavras-chave. . . .	36
Figura 15 – Perda relativa em relação à co-ocorrência entre palavras-chave. . .	37
Figura 16 – Rede bibliográfica com <i>clusters</i> de palavras inter-relacionadas de acordo com o grau de vínculo.	38
Figura 17 – Rede bibliográfica com <i>clusters</i> de palavras inter-relacionadas de acordo com o ano de publicação.	39
Figura 18 – Rede bibliográfica de densidade de palavras de acordo com a quantidade de publicações.	40
Figura 19 – Subconjuntos da mitigação de riscos na cadeia de suprimentos e principais palavras-chave.	45

Figura 20 – Fases do modelo de decisão para seleção de fornecedores.	46
Figura 21 – Métodos para seleção de fornecedores.	47
Figura 22 – Abordagens em modelagem e simulação na escala de nível de abstração.	58
Figura 23 – Diagrama de Venn relacionando sistemas de aprendizado relacionados à IA.	59
Figura 24 – Fluxograma mostrando como as diferentes partes de um sistema de IA se relacionam entre si nas diferentes disciplinas de IA.	60
Figura 25 – Processo de aprendizado supervisionado.	64
Figura 26 – Ilustração da predição de uma instância x utilizando k -vizinhos.	67
Figura 27 – Função logística representando a probabilidade de um sistema estar sob ou fora de controle.	68
Figura 28 – A abordagem de <i>citizen data scientists</i>	71
Figura 29 – Modelo de simulação <i>make-to-order</i>	73
Figura 30 – Modelo de seleção de fornecedores utilizando ML.	74
Figura 31 – Integração entre os modelos de simulação e ML.	76
Figura 32 – Curva ROC de todos os fornecedores com a semente de simulação n_1	79
Figura 33 – Curva ROC de todos os fornecedores com a semente de simulação n_2	80
Figura 34 – Curva ROC de todos os fornecedores com a semente de simulação n_3	80
Figura 35 – Curva ROC de todos os fornecedores com a semente de simulação n_4	81
Figura 36 – Curva ROC do fornecedor S_1 após simulação com quatro diferentes sementes de simulação (n_1 to n_4).	81
Figura 37 – Curva ROC do fornecedor S_2 após simulação com quatro diferentes sementes de simulação (n_1 to n_4).	82
Figura 38 – Comparação da acurácia do modelo que utiliza o algoritmo KNN com cinco sementes de simulação diferentes.	83
Figura 39 – Performance da confiabilidade de entrega após a seleção de fornecedores utilizando as técnicas de ML supervisionado, LR e KNN.	84

LISTA DE TABELAS

Tabela 1 – Definições de sistema produto-serviço.	27
Tabela 2 – Definições das três categorias de um PSS.	29
Tabela 3 – Definições de servitização.	30
Tabela 4 – As 10 palavras-chave mais frequentes utilizadas pelos autores durante o período de 2010 à 2017.	33
Tabela 5 – Principais critérios para seleção de fornecedores.	41
Tabela 6 – Vantagens e desvantagens dos métodos de seleção de fornecedores.	49
Tabela 7 – Principais softwares comerciais de DES.	56
Tabela 8 – Exemplo de uma base de dados com n instâncias, n características e um rótulo com valores binários.	65
Tabela 9 – Predições de probabilidade de entrega no prazo utilizando o algoritmo LR para cada fornecedor e os resultados da classificação. . .	82

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
IBGE	Instituto Brasileiro de Geografia e Estatística
ABM	<i>Agent-based Modeling</i>
ABNT	Associação Brasileira de Normas Técnicas
AHP	<i>Analytic Hierarchy Process</i>
CPS	<i>Cyber-physical systems</i>
DES	<i>Discrete event simulation</i>
DL	<i>Deep Learning</i>
DT	<i>Decision Tree</i>
ERP	<i>Enterprise Resource Planning</i>
FIFO	<i>First In, First Out</i>
GT	<i>Game Theory</i>
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
KNN	<i>k-Nearest Neighbors</i>
LIFO	<i>Last In, First Out</i>
LR	<i>Logistic Regression</i>
MILP	<i>Mixed Integer Linear Programming</i>
ML	<i>Machine Learning</i>
MOP	<i>Multi-Objective Programming</i>
MRP	<i>Materials Requirements Planning</i>
PSS	<i>Product-service systems</i>
ROC	<i>Receiver Operating Characteristic</i>
SD	<i>System dynamics</i>
SPT	<i>Shortest Processing Time</i>
TIC	Tecnologia de Informação e Comunicação

LISTA DE SÍMBOLOS

$P(x)$	Função que descreve a probabilidade de um evento estar sob controle
$g(x)$	Combinação linear de variáveis independentes
β_i	Coefficientes de regressão
acc_a	Acurácia do algoritmo que prediz as entregas no prazo
acc_b	Acurácia do algoritmo que prediz as entregas atrasadas
R_{ka}	Resultado do classificador KNN que prediz o fornecedor mais provável a realizar a entrega no prazo
R_{kb}	Resultado do classificador KNN que prediz o fornecedor mais provável a realizar a entrega atrasada
R_{lr1}	Resultado do classificador LR que prediz o fornecedor mais provável a realizar a entrega no prazo
R_{lr2}	Resultado do classificador LR que prediz o segundo fornecedor mais provável a realizar a entrega no prazo
S_i	Fornecedor
τ	Data
γ	Quantidade

SUMÁRIO

1	INTRODUÇÃO	15
1.1	CONTEXTUALIZAÇÃO	15
1.2	OBJETIVOS	18
1.2.1	Objetivo Geral	18
1.2.2	Objetivos Específicos	18
1.3	ESTRUTURA DO TRABALHO	18
2	REFERENCIAL TEÓRICO	20
2.1	<i>INDUSTRY 4.0 E SMART MANUFACTURING</i>	20
2.2	NOÇÕES DE GESTÃO DE OPERAÇÕES	22
2.3	O SERVIÇO E A MANUFATURA	25
2.3.1	Sistemas produto-serviço	26
2.3.2	Servitização	29
2.3.3	Sistema produto-serviço, servitização e dados	31
2.4	SELEÇÃO DE FORNECEDORES	39
2.5	SIMULAÇÃO DE SISTEMAS PRODUTIVOS	51
2.6	<i>MACHINE LEARNING</i>	58
2.6.1	Aprendizado não supervisionado	62
2.6.2	Aprendizado por reforço	62
2.6.3	Aprendizado supervisionado	63
2.6.3.1	<i>K-Nearest Neighbors</i>	66
2.6.3.2	<i>Logistic Regression</i>	66
2.6.4	Ciência de Dados e Manufatura	68
3	EXPERIMENTAÇÃO DIGITAL	72
3.1	MODELO DE SIMULAÇÃO	72
3.2	MODELO DE ML	72
3.3	INTEGRAÇÃO DOS MODELOS	75
3.4	CONSIDERAÇÕES DO MODELO	77
4	RESULTADOS E DISCUSSÕES	79
5	CONCLUSÕES	86
	REFERÊNCIAS	88

1 INTRODUÇÃO

Neste capítulo é apresentado o contexto que esta dissertação de mestrado está inserida e os objetivos geral e específico que pretendem-se alcançar.

1.1 CONTEXTUALIZAÇÃO

O mundo vive em um ambiente com escassez de recursos e a adoção por uma vida sustentável se faz necessária. No contexto de sistemas produtivos, ser sustentável pode significar manter ou aumentar a produtividade com uma contínua tentativa de minimização dos recursos. Partindo desta premissa, a meta por fazer mais com menos se torna algo desejável e/ou necessário para a manutenção da competitividade de qualquer organização. Em direção a esta meta, novas tecnologias de informação e comunicação (TICs) emergem dando suporte à manufatura.

Para Henning Kagermann (2015), é nítido que o impacto da transformação digital será comparável às ondas de inovação passadas, como a mecanização e a eletricidade. Portanto, uma discussão qualitativa e quantitativa é necessária para a caracterização dos novos serviços que visam atender às necessidades de uma indústria que passa por uma transformação digital – Indústria 4.0. Historicamente, o conceito de valor tem sido associado à produtividade por meio da entrega de bens tangíveis (RANDHAWA; SCERRI, 2015) e a transição para uma economia baseada em serviços necessita uma revisão do entendimento de criação de valor nas organizações (PITELIS, 2009). O paradigma da disponibilidade de dados e conectividade dos dispositivos por meio de sistemas ciber-físicos gera novas possibilidades de criação de valor nos sistemas produtivos e logísticos.

Historicamente, as empresas de manufatura têm se apoiado em resultados de processos observáveis por meio de iniciativas no chão de fábrica, como o trabalho padronizado e a melhoria contínua (FLATH; STEIN, 2018). Hoje, as empresas de manufatura têm acesso a diversas fontes de dados que fornecem uma grande quantidade de informações sobre status de processos e a incorporação destes dados permite que possam ser avaliados problemas não observáveis, como por exemplo, a degradação de máquina e defeitos ocultos (FLATH; STEIN, 2018). Consequentemente, sistemas de informação com capacidade de analisar estes dados com objetivo de guiar as organizações na tomada de decisão se fazem necessários (FLATH; STEIN, 2018).

O dicionário Houaiss conceitua serviço como um produto da atividade humana destinado à satisfação de necessidades, mas que não apresenta o aspecto de um bem material (HOUAISS; VILLAR; FRANCO, 2001). No contexto de serviços na manufatura, Kagermann *et al.* (2014) defendem que o atendimento às necessidades do cliente têm se tornado cada vez mais complexo, pois as necessidades são modeladas individualmente e devem ser atendidas a qualquer momento e em qualquer lugar.

Deve-se compreender que esta tarefa não é trivial de ser satisfeita e o estado atual da manufatura – englobando os recursos humanos e materiais – pode não ser capaz de cumpri-la. Para garantir o atendimento aos anseios dos clientes, as TICs são essenciais para entregar os valores requisitados por meio de novos serviços e modelos de negócio fortemente baseados em dados (KAGERMANN *et al.*, 2014).

A manufatura passa por um período de transição no qual a aquisição e análise de dados tem potencial para melhorar o padrão do que se entende por excelência em produtividade. Esta transição pode ser encarada como uma transformação digital e é definida por Henning Kagermann (2015) como o fenômeno de contínua convergência entre o mundo real e virtual, ocorrendo em um ambiente onde há uma grande disponibilidade de dados por meio das TICs. Também chamado de *Digitalization*, este fenômeno é comparável as ondas de inovação ocorridos devido à mecanização e eletricidade (Henning KAGERMANN, 2015). Segundo Lee *et al.* (2013), o modelo de fábrica inteligente que segue os conceitos da Indústria 4.0 utiliza máquinas conectadas que trabalham como uma comunidade colaborativa.

Além disto, por meio da utilização de ferramentas preditivas, dados são sistematicamente coletados, analisados e transformados em informações úteis para uma tomada de decisão apoiada em algoritmos inteligentes (LEE *et al.*, 2013). Portanto, os serviços de aquisição e análise de grandes volumes de dados passam a atuar como protagonistas no planejamento e controle de produção na manufatura digital. Segundo Kusiak (2018), embora as empresas tenham usado cada vez mais tecnologias de comunicação sem fio e sensores para captação de dados, a maioria delas não sabe como interpreta-los adequadamente com o objetivo de melhorar os seus processos e produtos. O autor ainda afirma que um dos fatores que contribuem para as empresas não saberem exatamente quais dados medir e quando medir é o fato da ciência dos dados não estar presente nos cursos de engenharia e de negócios atualmente.

Inicialmente, a inovação era vista como uma disciplina de pesquisa com foco na ciência, tecnologia e desenvolvimento de novos produtos para manufatura. Hoje, ela tem sido vista como fator chave para qualquer organização que almeja um crescimento sustentável para manutenção da sua competitividade (RANDHAWA; SCERRI, 2015). A literatura que aborda inovação em serviços expande na direção da multidisciplinaridade, envolvendo os segmentos de economia, marketing, ciência organizacional e gerenciamento (RANDHAWA; SCERRI, 2015). Ao mesmo tempo que a servitização é a combinação integrada de produtos e serviços para oferecimento de valor (T. S. BAINES; LIGHTFOOT; BENEDETTINI *et al.*, 2009), a inovação em serviços requer um alto grau de interatividade entre o prestador de serviço e o cliente, e portanto, muitas vezes a inovação se faz presente mais nesta interação do que no serviço ou processo em si (RANDHAWA; SCERRI, 2015). Os autores defendem que há diferença entre inovação na manufatura e inovação em serviços, sendo o primeiro orientado ao pro-

duto e tecnologia enquanto que o segundo é orientado a aspectos culturais, humanos, habilidades em comunicação e interface com o consumidor.

Embora os conceitos sejam distintos, devido à disponibilidade de novas tecnologias na era da Indústria 4.0, cada vez mais a manufatura poderá agregar o comportamento humano por meio dos sistemas ciberfísicos (CPS) englobando os aspectos da inovação em serviços. Alinhado a este raciocínio, Frazzon *et al.* (2013) afirmam que, apesar dos CPS serem muito relacionados aos avanços tecnológicos, aspectos como a criatividade, flexibilidade e capacidade de resolver problemas dos colaboradores são extremamente necessários para a operação do sistema.

Entre todas as tecnologias, a tecnologia da informação é dita como a mais significativa para a inovação em serviços e modelos de negócio virtuais, como já acontece no varejo (Amazon), música (Spotify), entretenimento (YouTube) e redes sociais (Facebook) (RANDHAWA; SCERRI, 2015). Neste contexto, a inovação em serviços para a manufatura é fortemente influenciada pela integração dos CPS por meio da tecnologia da informação. Kusiak (2018) afirma que as empresas de manufatura necessitam de uma maior orientação prática, pois o estado atual das pesquisas científicas concentra-se na abordagem da fronteira tecnológica, sem considerar como isto será aplicado. Ele complementa que as necessidades da indústria são outras, como por exemplo saber quais tipos de dados devem ser utilizados na amostragem, quais sensores usar e onde estes sensores devem ser instalados na linha de produção. O autor elenca as cinco principais lacunas de inovação no contexto da manufatura:

- Adoção de estratégias de dados;
- Melhoria na coleta, uso e compartilhamento de dados;
- Projeto de modelos preditivos;
- Estudo de modelos preditivos genéricos;
- Conexão das fábricas e controlar processos.

É possível verificar a forte relação que um grande volume de dados (*Big Data*) tem no novo contexto da Indústria 4.0, no qual a conectividade entre pessoas, máquinas e objetos por meio da internet das coisas será um fator preponderante para manutenção da competitividade das empresas de manufatura.

Inserida no contexto da manufatura digital, esta dissertação de mestrado busca resolver o problema da seleção de fornecedores utilizando tecnologias digitais como simulação de sistemas produtivos e técnicas de *machine learning*. O papel da simulação é a de representar um ambiente de manufatura real e o modelo de *machine learning* atua como uma ferramenta de tomada de decisões que, neste caso, refere-se

à seleção de fornecedores. O sistema que envolve a integração destas duas ferramentas pode ser abstraído como um serviço digital de seleção de fornecedores que utiliza dados passados para realizar a tomada de decisão.

Apesar dos tópicos de simulação e *machine learning* serem consolidados na literatura em geral, não há uma abordagem integrada e quantitativa dos possíveis benefícios da utilização destas ferramentas no contexto da seleção de fornecedores. Portanto, a motivação para realização deste trabalho se dá no ineditismo do assunto.

1.2 OBJETIVOS

Nesta seção são apresentados os objetivos geral e específicos que esta dissertação de mestrado pretende alcançar.

1.2.1 Objetivo Geral

O objetivo geral desta dissertação é o desenvolvimento de um modelo de *machine learning* que atua como serviço digital de seleção de fornecedores, a qual é avaliada de maneira integrada por meio de um modelo de simulação.

1.2.2 Objetivos Específicos

- Apresentar os principais conceitos da Indústria 4.0 e a sua relação ao tópico de serviços aplicados na manufatura;
- Apresentar o estado atual da literatura com relação à seleção de fornecedores;
- Discutir o impacto da cultura de dados na manufatura;
- Desenvolver uma base de dados sintética por meio de um modelo de simulação;
- Propor um método para integração dos modelos de simulação e *machine learning*;
- Apresentar uma abordagem orientada a dados para o problema de seleção de fornecedores por meio de um modelo de *machine learning*;
- Avaliar o desempenho da abordagem, discutir melhorias e novas oportunidades de pesquisa.

1.3 ESTRUTURA DO TRABALHO

O trabalho está estruturado em cinco capítulos. O Capítulo 1 contextualiza o cenário de transformação digital da manufatura e a inovação em serviços com orientação a dados. Além disso, neste capítulo são apresentados os objetivos geral e

específicos que pretendem-se alcançar e a estruturação do trabalho. No Capítulo 2, é apresentado um referencial teórico detalhado para o entendimento pleno desta pesquisa. Os conceitos apresentados englobam tópicos como Indústria 4.0, noções de gestão de operações, serviços na manufatura, seleção de fornecedores, simulação de sistemas produtivos e técnicas de *machine learning*. No Capítulo 3, são apresentados os modelos de simulação e *machine learning*, assim como o método de integração e as considerações da modelagem para realizar a experimentação digital do problema. No Capítulo 4, são apresentados os resultados e discussões da experimentação. Finalmente, no Capítulo 5, são apresentadas as conclusões desta dissertação e novas oportunidades de pesquisa.

2 REFERENCIAL TEÓRICO

Neste capítulo são apresentados os conceitos fundamentais para o entendimento deste trabalho. Ao final deste capítulo é esperado que o leitor compreenda tópicos relacionados à manufatura – desde tópicos clássicos de gestão de operações às possibilidades de utilização da ciência de dados na manufatura – base da transformação digital na manufatura (Indústria 4.0).

2.1 *INDUSTRY 4.0 E SMART MANUFACTURING*

Dois termos comumente utilizados para descrever a modernização do setor industrial com a utilização de novas tecnologias são: Indústria 4.0 (*Industry 4.0*) e Manufatura Inteligente (*Smart Manufacturing*). Segundo Kang *et al.* (2016), os dois termos se equivalem, sendo a Indústria 4.0 uma realização da Manufatura Inteligente por meio da integração entre setor privado, governo e academia.

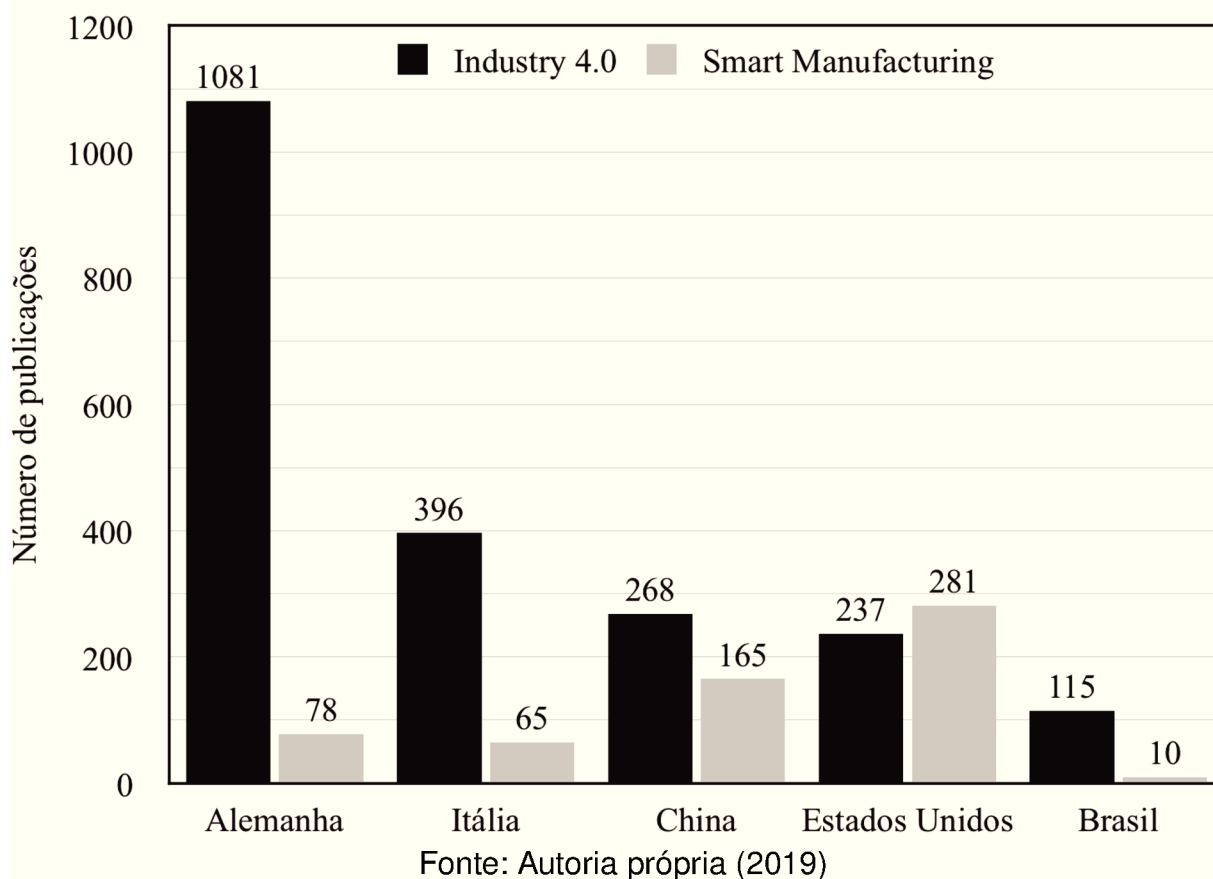
Na Alemanha, o termo Indústria 4.0 foi apresentado na feira de Hannover em 2011 como forma de descrever a revolução digital nas cadeias de valor causada por tecnologias avançadas e integradas que podem transformar a economia global e a sociedade como um todo (KLAUS, 2016). As ditas fábricas inteligentes que derivam da quarta revolução industrial criam um ambiente no qual recursos físicos e virtuais cooperam por meio da internet utilizando sensores cada vez mais baratos em combinação a técnicas de inteligência artificial (KLAUS, 2016). Segundo Henning Kagermann *et al.* (2013), a Indústria 4.0 não deve ser vista como uma fábrica inteligente isolada e sim como um elemento entre vários outros que, juntos, implicam em criação de valor, novos modelos de negócios, serviços e organização do trabalho. Portanto, é necessária uma implementação interdisciplinar de forma cooperativa entre as áreas chave – mobilidade, logística, construção, produto e energia inteligentes, que fazem parte da conexão entre a internet das coisas e a internet dos serviços (Henning KAGERMANN, 2015).

Kusiak (2017) afirma que, apesar de não existir uma única definição aceita do que seria a Manufatura Inteligente, pode-se entender como uma forma de produção integrada às tecnologias de comunicação, controle, simulação, plataformas computacionais e sensores que se utiliza do uso intensivo de dados e engenharia preditiva. A manufatura inteligente utiliza os conceitos dos sistemas ciber-físicos como a internet das coisas, computação em nuvem, computação orientada ao serviço, inteligência artificial e ciência dos dados. Segundo Davis *et al.* (2012), as fábricas inteligentes possibilitadas pelo uso intensivo da tecnologia da informação e redes de suprimento, podem responder melhor às decisões estratégicas e tem potencial para revitalizar o setor industrial, pois resulta na inovação na manufatura, competitividade e performance produtiva. Para Kusiak (2017), a essência da manufatura inteligente é baseada em seis

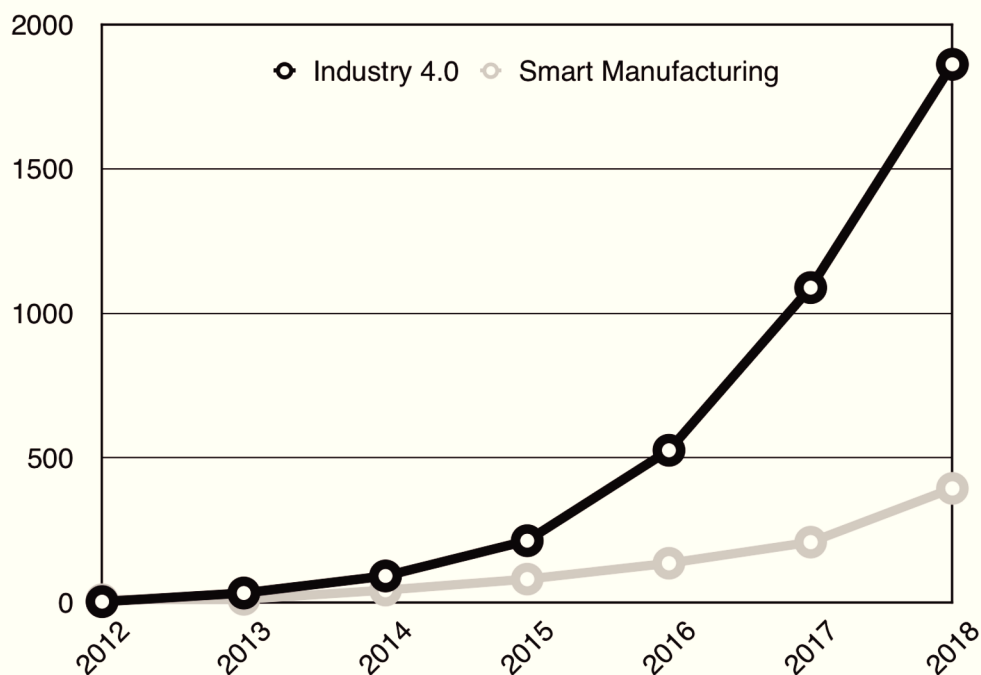
pilares: tecnologia de manufatura e processos, materiais, dados, engenharia preditiva, sustentabilidade e compartilhamento dos recursos em rede.

Por fim, nota-se que as definições são relativamente equivalentes, pois buscam propor um novo paradigma para os sistemas produtivos e logísticos. Após uma coleta de dados de trabalhos científicos na plataforma Scopus, percebe-se que há uma forte relação do termo Indústria 4.0 à Alemanha ao mesmo tempo que a Manufatura Inteligente está correlacionada aos Estados Unidos. A Figura 1 ilustra a coleta de dados realizada na plataforma Scopus. Pode-se perceber que apesar dos termos serem equivalentes, o termo Indústria 4.0 têm se tornado cada vez mais popular devido à sua preferência na literatura científica, como ilustrado na Figura 2, e também nas mais variadas empresas de consultoria empresarial e serviços como McKinsey (D. WEE *et al.*, 2015), Deloitte (SCHLÄPFER; M. KOCH; MERKHOFFER, 2015), Boston Consulting Group (RÜSSMANN *et al.*, 2015), PwC (V. KOCH *et al.*, 2014) e Fórum Econômico Mundial (LEOPOLD; ZAHIDI; RATCHEVA, 2016).

Figura 1 – Comparação entre documentos científicos publicados e país de publicação com relação às palavras-chave *Industry 4.0* e *Smart Manufacturing*.



Além disso, a Figura 3 mostra que o termo Industry 4.0 tem sido o termo mais procurado no maior website de buscas do mundo – Google – e portanto, o conjunto destes fatores pode fazer com que o termo Indústria 4.0 se consolide como mais popular em relação à Manufatura Inteligente no longo prazo.

Figura 2 – Número de publicações em função do tempo dos termos *Industry 4.0* e *Smart Manufacturing*.

Fonte: Autoria própria (2019)

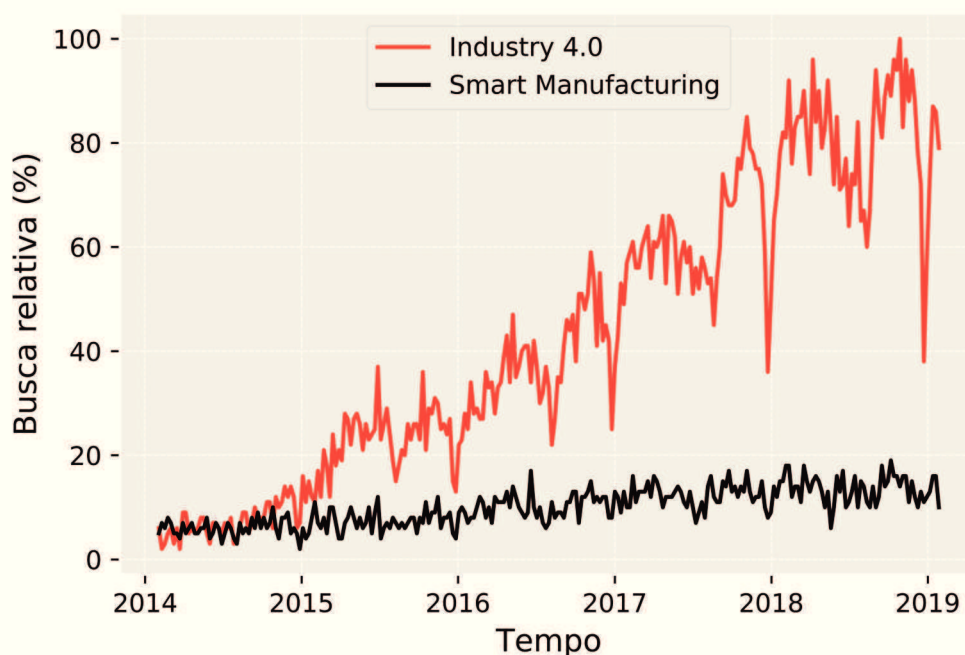
Segundo Monostori *et al.* (2016), o objetivo da Indústria 4.0 é garantir que a necessidade do cliente seja atendida ao preço da produção em massa, e para alcançá-lo, novas maneiras de agregar valor à empresa e ao controle do sistema devem ser implementadas. Por fim, as novas oportunidades devido à transformação digital da produção em combinação aos sistemas ciber-físicos envolvem a participação de recursos humanos e materiais com o intuito de otimizar as suas performances por meio de uma integração inteligente. A Figura 4 resume as maneiras de agregar valor aos sistemas produtivos no contexto da Indústria 4.0.

2.2 NOÇÕES DE GESTÃO DE OPERAÇÕES

O software que faz o controle e gerenciamento das informações do sistema é geralmente chamado de *Enterprise Resource Planning*, ou sistema ERP, e é responsável por integrar os dados e processos da organização de forma que a informação esteja disponível para todos os departamentos, que por sua vez, o abastecem com informações atualizadas dos status das tarefas e máquinas (PINEDO, 2016). A Figura 5, proposta por Slack, Chambers e Johnston (2013), ilustra o sistema ERP e suas conexões com os diversos departamentos da empresa.

Segundo Pinedo (2016), a programação da produção afeta a organização também no médio e longo prazo devido à conexão com o *Materials Requirements Planning* (MRP). O MRP é o sistema que calcula quando e quantas peças e materiais são requi-

Figura 3 – Comparação das buscas relativas no website Google com relação às palavras-chave *Industry 4.0* e *Smart Manufacturing* utilizando a ferramenta Google Trends.



Fonte: Autoria própria (2019)

Figura 4 – Agregação valor aos sistemas produtivos no contexto da Indústria 4.0.

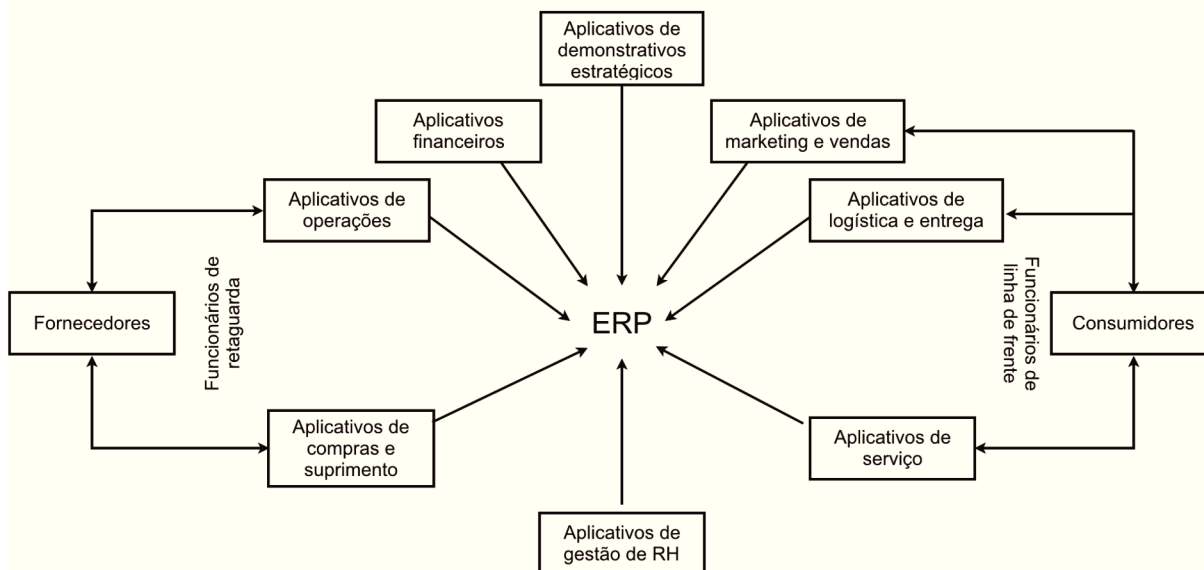


Fonte: Adaptado de Monostori *et al.* (2016)

sitadas (SLACK; CHAMBERS; JOHNSTON, 2013). A Figura 6 mostra o esquemático do MRP e a sua relação com outras atividades em uma empresa de manufatura.

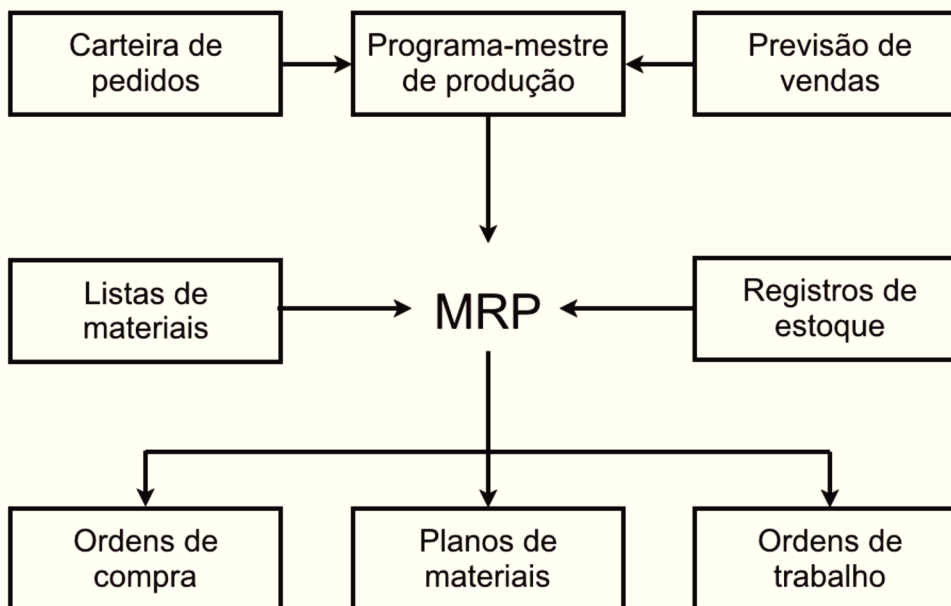
Pinedo (2016) propõe um diagrama de fluxo de informação em um sistema de manufatura apresentado pela Figura 7, a qual ilustra, de maneira abstraída a integração dos sistemas que envolvem a programação e controle da produção em uma

Figura 5 – Sistema ERP e suas conexões com os diversos departamentos da empresa.



Fonte: Adaptado de Slack, Chambers e Johnston (2013)

Figura 6 – O sistema MRP e suas conexões com outras atividades em uma empresa de manufatura.

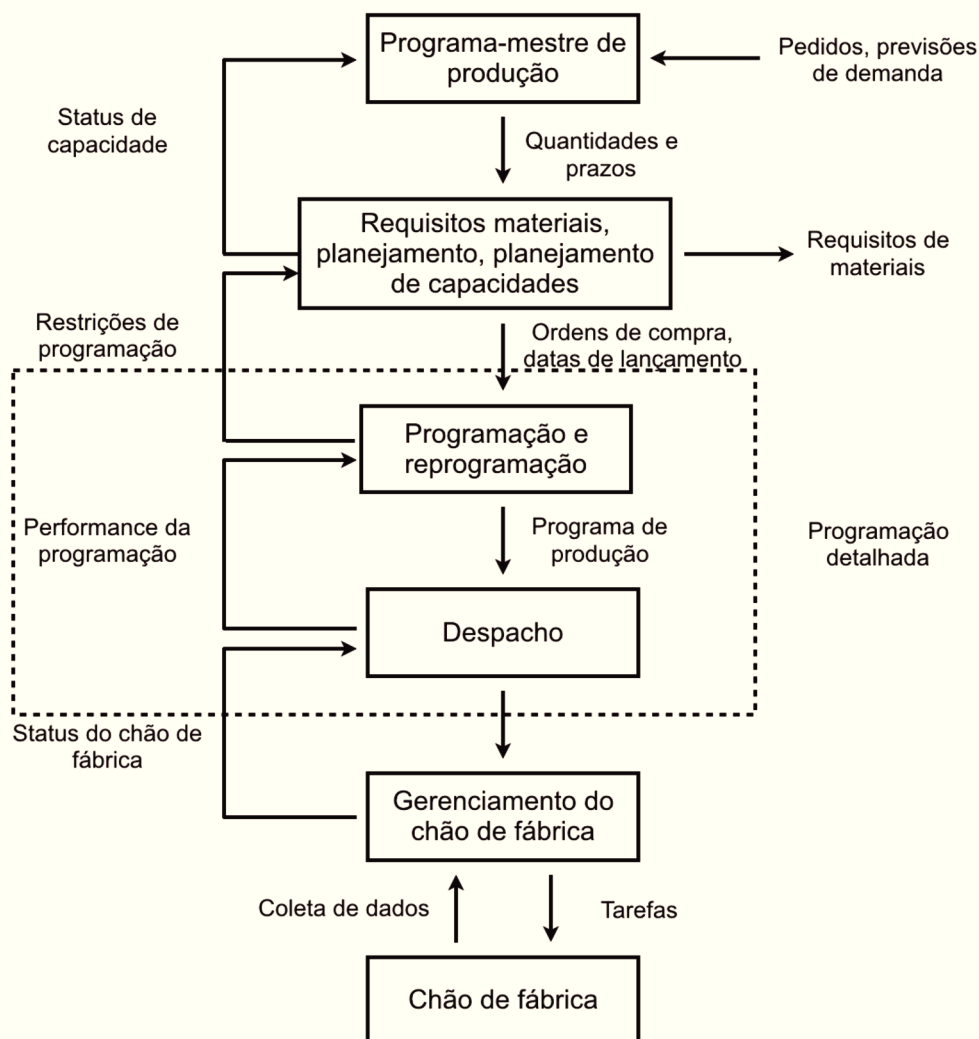


Fonte: Adaptado de Slack, Chambers e Johnston (2013)

empresa de manufatura.

Para as empresas de serviços, o fluxo de informação é a entidade principal e o seu sistema não é fácil de ser generalizado como as empresas de manufatura, pois deve ser coordenado por outras funções de tomada de decisão (PINEDO, 2016). Estas funções de tomada de decisão são baseadas em grandes bancos de dados que contém informações relativas, geralmente, à disponibilidade de recursos e potenciais clientes, e interagem frequentemente com os módulos de previsão de demanda e

Figura 7 – *Framework* simplificado da integração dos sistemas que envolvem a programação e controle da produção em uma empresa de manufatura.



Fonte: Adaptado de Pinedo (2016)

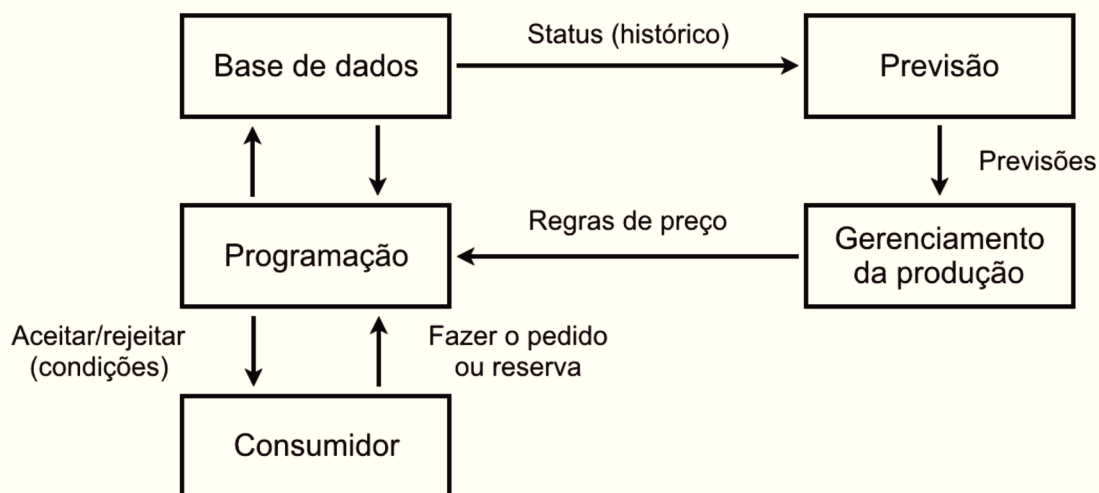
gerenciamento de produção (PINEDO, 2016). A Figura 8 ilustra o fluxo de informações em uma empresa de serviços proposta por Pinedo (2016).

Vale observar que, diferentemente das empresas de manufatura, o MRP não está presente no ambiente de serviços, pois não existem recursos tangíveis.

2.3 O SERVIÇO E A MANUFATURA

Empresas com modelos de negócio orientados ao produto buscam o aumento do número de produtos vendidos. Este é o principal método aplicado para aumentar o faturamento, a fatia de mercado e a geração de lucros (TUKKER, 2015). No entanto, em modelos de negócio orientados ao serviço, as empresas faturam por meio dos serviços oferecidos e, portanto, os produtos físicos e consumíveis passam a ser encarados como custos (TUKKER, 2015).

A terminologia utilizada para designar estratégias de orientação ao serviço é

Figura 8 – *Framework* do fluxo de informações em uma empresa de serviços.

Fonte: Adaptado de Pinedo (2016)

diversificada e abrange vários termos na literatura científica como: servitização, sistema produto-serviço, vendas funcionais e produtos funcionais (OPRESNIK; TAISCH, 2015; KARLSSON; LARSSON; ÖHRWALL RÖNNBÄCK, 2018). No entanto, em uma pesquisa realizada na plataforma Scopus com estas palavras-chave em combinação à palavra “manufacturing”, os termos com maior número de publicações e relevância são os termos “servitização” e “sistemas produto-serviço”, apresentada na Figura 9. Portanto, as pesquisas nesta dissertação de mestrado estão fortemente relacionadas a estes dois termos – sistema produto-serviço e servitização – quando o objetivo for abordar estratégias de manufatura orientadas ao serviço.

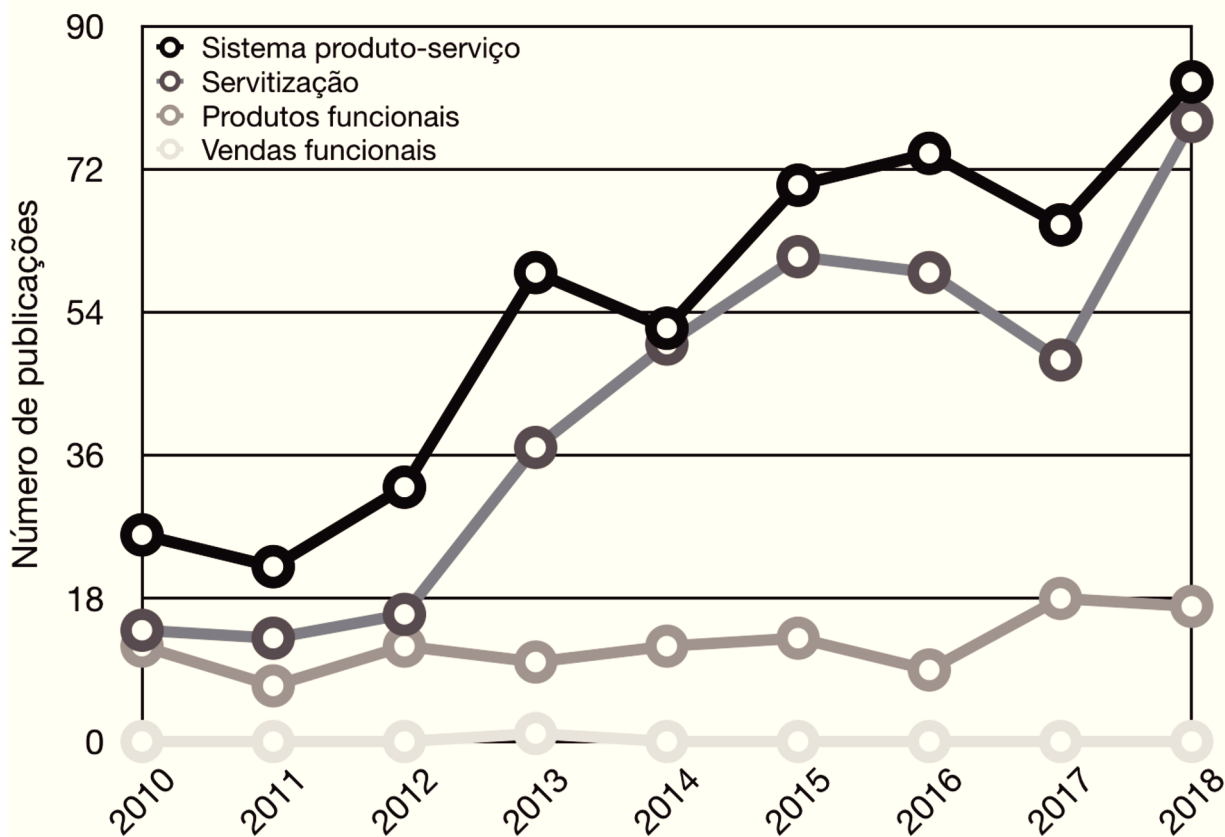
Ressalta-se que a plataforma Scopus foi escolhida, pois é a melhor ferramenta disponível para pesquisa eletrônica de literatura científica em comparação a alternativas como Web of Science e Google Scholar, particularmente para artigos publicados após o ano de 1995 (FALAGAS *et al.*, 2008).

2.3.1 Sistemas produto-serviço

Uma definição consolidada de sistema produto-serviço (PSS) pela literatura científica é proposta por T. S. Baines, Lightfoot, Evans *et al.* (2007), o qual apresenta o PSS como uma forma de entregar valor por meio do oferecimento de um produto integrado a um serviço. Para Boehm e Thomas (2013), o PSS é um pacote integrado de produtos e serviços que objetiva a criação de valor e utilidade para o cliente. Goedkoop (1999) define as palavras-chave que compõem o PSS:

- Produto: uma *commodity* tangível manufaturada para ser vendida. É capaz de “cair nos seus pés” e atender as necessidades dos clientes;
- Serviço: uma atividade realizada com valor econômico e geralmente feita com

Figura 9 – Número de publicações em função do tempo dos termos que designam estratégias de orientação ao serviço em combinação à palavra *manufacturing*, utilizando a ferramenta de pesquisa Scopus.



Fonte: Autoria própria (2019)

fins comerciais;

- Sistema: uma coleção de elementos incluindo suas relações.

Em um PSS, o foco não está na venda do produto e sim na sua utilização, o que pode fazer com que haja uma reestruturação dos riscos, responsabilidades e custos que são atrelados à posse (T. S. BAINES; LIGHTFOOT; EVANS *et al.*, 2007). A Tabela 1 compila as diversas definições de PSS segundo as revisões de literatura apresentadas por Tukker (2015) e T. S. Baines, Lightfoot, Evans *et al.* (2007):

Tabela 1 – Definições de sistema produto-serviço.

Autor (ano)	Definição de sistema produto-serviço
Goedkoop (1999)	Um conjunto comercializável de produtos e serviços capazes de cumprir, juntos, as necessidades do cliente.

Continua na próxima página

Tabela 1 – Definições de sistema produto-serviço.

Autor (ano)	Definição de sistema produto-serviço
Manzini e Vezzoli (2003)	Uma estratégia de inovação que muda o foco comercial de projetar e vender apenas produtos físicos para projetar e vender um sistema de produtos e serviços que, de maneira conjunta, atendem às demandas específicas dos clientes.
Mont (2004)	Um sistema de produtos, serviços e redes com infraestrutura de suporte que é projetado para ser competitivo, satisfazendo as necessidades dos clientes com baixo impacto ambiental em relação aos modelos de negócio tradicionais.
Tukker e Tischner (2017)	Consiste na combinação de produtos tangíveis e serviços intangíveis que, juntos, são capazes de suprir as necessidades dos clientes.
Berkovich, Leimeister e Krcmar (2011)	Um pacote integrado de hardware, software e elementos de serviço que objetiva a solução de problemas do cliente.
F. Zhang <i>et al.</i> (2012)	Um pacote sistemático no qual serviços intangíveis são anexados a produtos tangíveis para suprir atividades industriais em todo o ciclo de vida do produto.
Boehm e Thomas (2013)	Um pacote integrado de produtos e serviços que objetiva a criação de valor e utilidade para o cliente.

Fonte: Adaptado de Tukker (2015) e T. S. Baines, Lightfoot, Evans *et al.* (2007)

Os PSS podem ser subdivididos em três categorias ou tipos: os orientados ao produto, orientados ao uso e orientados ao resultado. Esta classificação foi primeiramente cunhada por Tukker (2004) e é amplamente aceita na literatura. A Tabela 2 apresenta as definições das três categorias.

Tabela 2 – Definições das três categorias de um PSS.

Categoria do PSS	Características
Orientado ao produto	Modelo de negócio ainda focado na venda de produtos, embora hajam serviços adicionados. A motivação da empresa reside na minimização dos recursos para manter o produto em funcionamento adequado e no maior tempo possível, aumentando o seu ciclo de vida. Exemplos: serviços de pós-venda como contratos de manutenção, seguro e consultoria.
Orientado ao uso	Embora o produto tradicional ainda possua papel central, o modelo de negócio não é direcionado à venda. O produto pertence ao fornecedor e está disponível aos clientes de diferentes formas, inclusive podendo ser compartilhado entre usuários. Neste caso, o objetivo da empresa é maximizar o uso do produto e o seu ciclo de vida e minimizar os materiais utilizados para produzi-lo. O produto pode ser utilizado por aluguel (uso por um único usuário).
Orientado ao resultado	Neste caso o fornecedor e cliente acordam um resultado desejado e não há produto envolvido. Por exemplo, terceirização de alimentos, pagamento por cópias em uma copiadora ou serviços de gerenciamento de pneus baseado na decolagem e pouso de aviões.

Fonte: Adaptado de Tukker (2004) e T. S. Baines, Lightfoot, Evans *et al.* (2007)

Diante das três categorias apresentadas, Tukker (2015) afirma que o tipo de PSS orientado ao resultado é o mais promissor, pois facilita a mudança para uma economia circular e eficiente em recursos, visto que o lucro é dependente do resultado e não somente da venda do produto. Neste tipo de PSS, a minimização dos produtos e consumíveis é incentivada, pois estes agora são vistos como custos para o fornecedor.

2.3.2 Servitização

A primeira definição do termo servitização é apresentado como o processo de mudar a visão de produtos e serviços separadamente para um pacote integrado de sistemas, com serviços atuando no papel principal (VANDERMERWE; RADA, 1988). Outra definição apresentada por T. S. Baines, Lightfoot, Benedettini *et al.* (2009) é a servitização como inovação das capacidades e processos de uma organização, substituindo a venda de produtos para a venda de produtos e serviços integrados, i.e., PSS, que oferecem valor na sua utilização. Portanto, constata-se que ambos autores

concordam que servitizar é fazer com que o produto não seja mais o elemento central da oferta de valor para o cliente. Outras definições de servitização são apresentadas na Tabela 3, adaptada de T. S. Baines, Lightfoot, Benedettini *et al.* (2009):

Tabela 3 – Definições de servitização.

Autor (ano)	Definição de servitização
Vandermerwe e Rada (1988)	Pacotes integrados de produtos, serviços, suporte, self-service e conhecimento focados no cliente.
White, Stoughton e Feng (1999)	Serviços baseados em produtos que desfazem a distinção entre atividades industriais e setoriais de serviços tradicionais.
Ren e Gregory (2007)	Um processo de mudança em que as empresas de manufatura adotam a orientação do serviço e/ou desenvolvem mais e melhores serviços, com o objetivo de satisfazer às necessidades do cliente, obter vantagens competitivas e melhorar o desempenho da empresa.
Desmet <i>et al.</i> (2013)	Uma tendência em que empresas de manufatura adotam mais e mais componentes de serviço nas suas ofertas.
T. Baines e W. Lightfoot (2013)	Refere-se à transformação em que empresas de manufatura estão oferecendo cada vez mais serviços diretamente relacionados aos seus produtos.

Fonte: Adaptado de T. S. Baines, Lightfoot, Benedettini *et al.* (2009)

Assim como no caso dos PSS, as estratégias de servitização na manufatura podem ser tipificadas e são compilados de acordo com os estudos de Wise e Baumgartner (2000), T. S. Baines, Lightfoot, Benedettini *et al.* (2009), Raddats e Easingwood (2010) e Rymaszewska, Helo e Gunasekaran (2017):

- Serviços acoplados a produtos em seus próprios produtos e produtos terceirizados (e.g. instalação, treinamento e suporte);
- Serviços gerais de mercado (e.g. atividades financeiras);
- Serviços operacionais de diagnóstico de fornecedores (e.g. consultoria técnica e integração de sistemas);
- Serviços operacionais nos próprios produtos (e.g. serviços de gerenciamento e disponibilidade de ativos);
- Serviços de controle e distribuição (e.g. roteamento logístico);

- Soluções integradas nas quais a empresa vai além da performance dos seus produtos-base tradicionais, com o intuito de atender a necessidade do cliente (e.g. preocupação com a infraestrutura de estações de recarregamento de carros elétricos pela Tesla).

Observa-se que as estratégias para servitização na manufatura englobam serviços que estão fortemente relacionados à melhoria do sistema e não apenas do produto em si, tornando assim, não apenas a empresa mais competitiva, mas a cadeia como um todo.

Por fim, conclui-se que o conceito de sistemas produto-serviço é um caso especial de servitização, pois valoriza o desempenho ou utilização dos ativos ao invés objetivar a propriedade, provoca diferenciação e fornece valor na integração do produto e serviço (T. S. BAINES; LIGHTFOOT; EVANS *et al.*, 2007).

2.3.3 Sistema produto-serviço, servitização e dados

A servitização pode ser vista como uma boa estratégia econômica de diferenciação de produtos em comparação às *commodities* (OPRESNIK; TAISCH, 2015), as quais são economicamente bem sucedidas quando oferecidas aos menores preços. Mas, ao mesmo tempo que a servitização na manufatura tem sido vista como uma estratégia de negócio para alavancar a competitividade das empresas no passado, hoje, ela tem se tornado cada vez mais comum e menos aceita como um diferencial para as empresas e clientes, fenômeno que pode ser chamado de “comoditização” (OPRESNIK; TAISCH, 2015).

De acordo com Neely (2008), a China possuía 1% das suas empresas oferecendo serviços em 2007, já em um levantamento de 2011, o mesmo país alcançou a marca dos 20% (OPRESNIK; TAISCH, 2015). Ou seja, as economias que historicamente sempre foram fortemente engajadas na venda de commodities têm passado a criar valor por meio de serviços e competindo com economias de países desenvolvidos que já possuíam uma cultura de oferecimento de serviços, como o caso dos Estados Unidos, o qual 58,6% das empresas já ofereciam serviços em 2007 (NEELY, 2008). Portanto, o questionamento é: como criar valor e se diferenciar num mundo competitivo onde todos, aparentemente, caminham para o oferecimento de produtos e serviços integrados? Uma possível tentativa de resposta inicia ao perceber a intersecção entre os interesses das pesquisas científicas no âmbito da *Industry 4.0* e das grandes empresas de manufatura e tecnologia: a ciência dos dados.

O entusiasmo com relação à utilização de dados é notório no mundo e nas mais diversas organizações, principalmente as grandes empresas de tecnologia, que cada vez mais investem em equipes para analisar e criar valor do fluxo de informações dentro e fora das suas empresas. Novas profissões surgem no meio corporativo, como por

exemplo, o cientista de dados, profissão que pode ter uma ampla atuação e geração valor em diversos ambientes, desde o campo (agricultura) até a indústria de transformação. A frase *“Data is the new oil”*, ou seja, “o dado é o novo petróleo” – ou, com a variação, “o dado é o novo recurso natural” – foi primeiramente cunhada por Clive Humby em 2006 (ARTHUR, 2013) e representa este entusiasmo com a utilização dos dados no mundo capitalista. Esta frase tem sido repetida por vários executivos nos últimos anos – como Ginni Rometty, atual CEO da IBM (ROMETTY, 2013) e Qi Lu, chefe do grupo de Aplicações e Serviços da Microsoft (LU, 2017), além de ter sido assunto de capa da revista *The Economist*, na primeira edição de maio de 2017 (ECONOMIST, 2017).

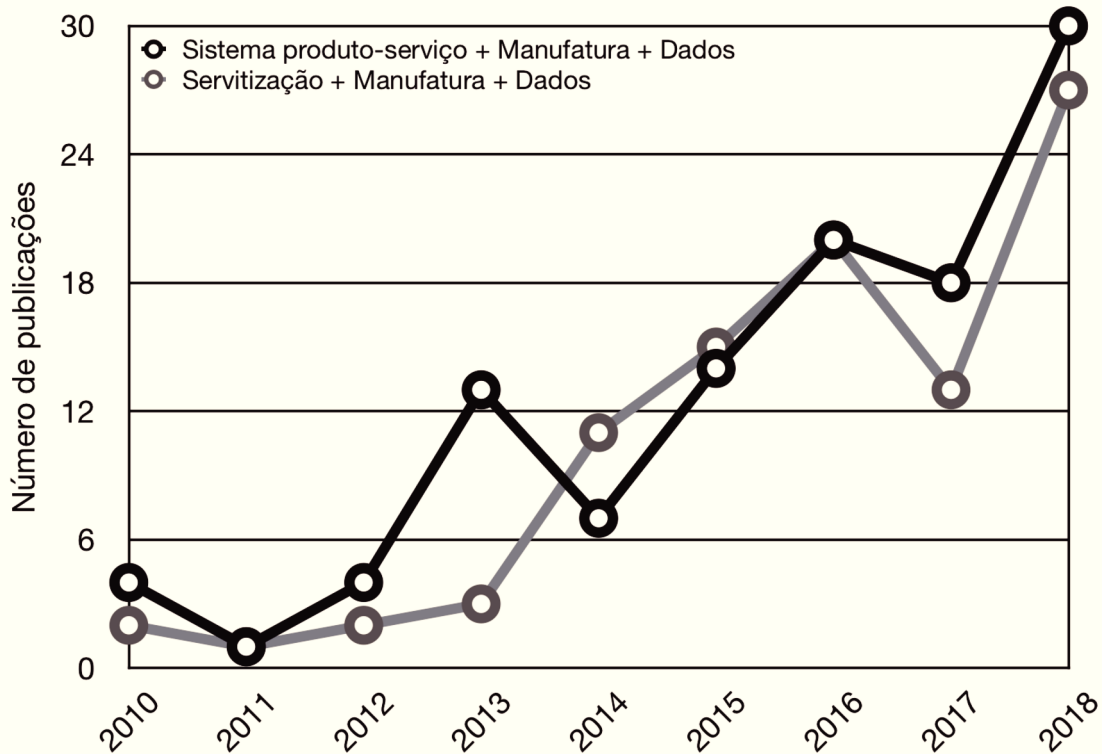
Para investigar a relação entre servitização, sistemas produto-serviço e a ciência dos dados, foi conduzida uma pesquisa na plataforma Scopus na qual verificou-se o crescente número de publicações que abordam o serviço na manufatura e a exploração dos dados como palavras-chave providas pelos autores. A Figura 10 ilustra a quantidade de publicações em relação ao tempo com relação às palavras-chave TITLE-ABS-KEY (“Product-service system” AND “Manufacturing” AND “Data”). De maneira análoga, a mesma busca foi realizada substituindo a palavra “Product-Service System” pela palavra “Servitization”. Vale ressaltar que a plataforma Scopus desconsidera o uso do hífen e automaticamente busca por palavras escritas no singular e plural.

Na tentativa de obter um resultado mais abrangente, foi observado um comportamento semelhante ao combinar as palavras serviço, manufatura e dados com a busca: TITLE-ABS-KEY (“Service” AND “Manufacturing” AND “Data”). A Figura 11 apresenta o resultado da quantidade de publicações em função do tempo.

O fato da quantidade de publicações da palavra “Serviço(s)” ser maior é justificada pelo fato de englobar tanto pesquisas relativas a PSS e Servitização como também outros grandes campos de pesquisa – ciência dos serviços (service science), engenharia de serviços (service engineering), projeto de serviços (service design) e inovação em serviços (service innovation), que possuem grupos de estudo diversificados, incluindo, além da engenharia, também o design, marketing, economia e administração.

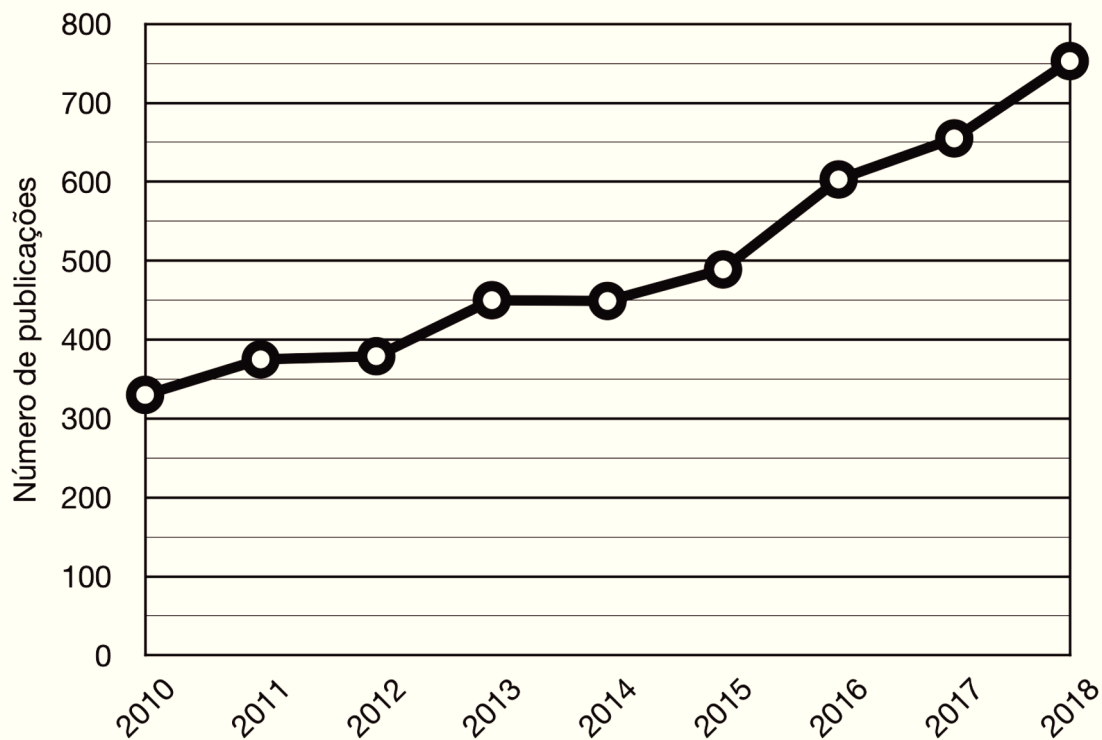
Com o intuito de ilustrar as palavras-chave mais frequentes utilizadas pelos autores, dados os parâmetros de busca mencionados anteriormente, a Tabela 4 compila os resultados das 10 palavras-chave mais frequentes.

Figura 10 – Quantidade de publicações das palavras Servitização, Sistema produto-serviço, Manufatura e Dados durante o período de 2010 à 2018.



Fonte: Autoria própria (2019)

Figura 11 – Quantidade de publicações das palavras Serviço, Manufatura e Dados ao longo do tempo durante o período de 2010 à 2018.



Fonte: Autoria própria (2019)

Tabela 4 – As 10 palavras-chave mais frequentes utilizadas pelos autores durante o período de 2010 à 2017.

PSS		Servitização		Serviços	
<i>Manufacture</i>	33	<i>Big Data</i>	23	<i>Data Envelopment Analysis</i>	687
<i>Big Data</i>	29	<i>Manufacture</i>	13	<i>Data Mining</i>	655
<i>Information Management</i>	18	<i>Industry 4.0</i>	7	<i>Big Data</i>	562
<i>Life Cycle</i>	18	<i>Servitization</i>	7	<i>Manufacture</i>	493
<i>Product-service systems</i>	18	<i>Data Analytics</i>	6	<i>Information Management</i>	427
<i>Decision Making</i>	17	<i>Digital Storage</i>	5	<i>Decision Making</i>	367
<i>Product Design</i>	15	<i>Manufact. Industries</i>	5	<i>Data Handling</i>	286
<i>Data Mining</i>	11	<i>Big Data Analytics</i>	4	<i>Efficiency</i>	216
<i>Maintenance</i>	11	<i>Data Mining</i>	4	<i>Life Cycle</i>	190
<i>Data Envelopment Analysis</i>	10	<i>Embedded Systems</i>	4	<i>Optimization</i>	186

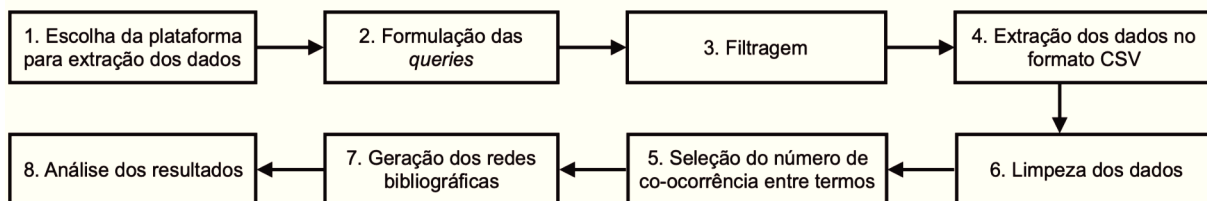
Fonte: Autoria própria (2018)

Percebe-se que a palavra *Big Data* apresenta-se com forte tendência para pesquisas futuras, pois figura entre as três mais frequentes palavras-chave utilizadas pelos autores com publicação nos temas: manufatura, serviços, servitização, sistemas produto-serviço e dados entre os anos de 2010 e 2017.

Uma última análise foi conduzida com relação às palavras-chave relacionadas a serviços e manufatura utilizando o software VOSviewer, o qual é uma ferramenta para construir e visualizar redes bibliométricas. O software também oferece funcionalidades de mineração de texto que podem ser usadas para construir e visualizar co-ocorrências de termos extraídos da literatura científica (ECK; WALTMAN, 2009). A Figura 12 ilustra o método utilizado para visualização das redes bibliométricas nesta dissertação de mestrado.

A plataforma Scopus foi selecionada para a extração dos dados científicos. Para esta análise foi realizada uma *query*, englobando ambos os termos Servitização e Sis-

Figura 12 – Método para visualização de redes bibliométricas com auxílio do software VOSViewer.



Fonte: Autoria própria (2018)

temas produto-serviço: TITLE-ABS-KEY(("Servitization" OR "Product-service system") AND "Manufacturing"), retornando 671 documentos no total. A filtragem realizada foi em função do tempo, ou seja, somente publicações dos últimos 6 anos foram compiladas, no caso, de 2012 à 2017, resultando no total de 547 documentos extraídos para próxima etapa, a limpeza dos dados.

A fase de limpeza dos dados é fundamental para que os resultados gráficos estejam coerentes sob o ponto de vista terminológico. Por exemplo, a palavra-chave "PSS" possui o mesmo significado do termo "Product-Service System", e caso não haja uma compatibilização destes termos, estes serão contabilizados como palavras-chave diferentes, o que não é verdade. Portanto, os termos que foram identificados e compatibilizados nesta pesquisa são apresentados na Figura 13.

Ressalta-se que, apesar da plataforma Scopus significar igualmente as palavras no plural e/ou com hífen de maneira automática, o software VOSViewer não possui tal funcionalidade, o que faz com que a fase de limpeza de dados seja fundamental para a geração de uma rede bibliométrica adequada.

Após a fase de limpeza de dados é necessário selecionar o número de co-ocorrência entre termos para finalmente gerar a rede bibliográfica. Este número pode ser aleatoriamente escolhido a critério do autor ou possuir uma metodologia (MONOSTORI *et al.*, 2016), pois isto definirá a densidade da rede bibliométrica. Nesta pesquisa, o método proposto consiste na avaliação da redução do número de palavras-chave em função do número de termos em co-ocorrência. Ou seja, selecionar uma condição na qual haja uma significativa redução no número de palavras-chave visando obter um mapa legível, e ao mesmo, sem perder palavras-chave fundamentais. Observa-se, a partir da Figura 14, que a escolha de 8 co-ocorrências entre termos é uma boa opção para o equilíbrio entre densidade do mapa e quantidade de palavras chave.

Optou-se por selecionar 8 co-ocorrências pois caso sejam escolhidas 9 co-ocorrências, haverá uma perda de 23 para 16 palavras-chave neste passo, o que representa percentualmente, uma perda relativa de 30,4% dos termos como observado na Figura 15, logo, isto acarretaria em uma rede bibliográfica pouco detalhada.

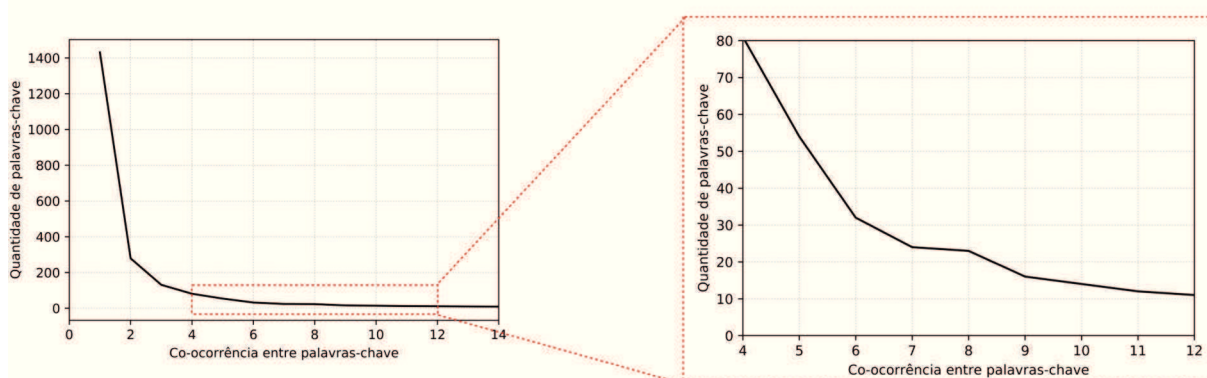
Por fim, as últimas etapas consistem na geração das redes bibliográficas e análise dos resultados. Três redes bibliográficas são possíveis. A primeira forma clusters

Figura 13 – Compatibilização das palavras-chave com o mesmo significado.

Product-service system	Product-service system Product service system Product-service systems Product service systems Industrial product-service system Industrial product service system Industrial product-service systems Industrial product service systems PSS Product-service system (PSS) Product-service system design
Servitization	Servitization Servitisation Manufacturing servitization Manufacturing servitisation
Manufacturing	Manufacturing Manufacturing industry Manufacturing industries
Business model	Business model Business models

Fonte: Autoria própria (2018)

Figura 14 – Quantidade em relação à co-ocorrência entre palavras-chave.



Fonte: Autoria própria (2018)

de palavras que são inter-relacionadas com linhas mais finas ou grossas, de acordo com o grau de vínculo que possuem, sendo as mais finas com menor vínculo que as mais grossas. A Figura 16 ilustra esta rede para a *query* mencionada na etapa 2. A

Figura 15 – Perda relativa em relação à co-ocorrência entre palavras-chave.



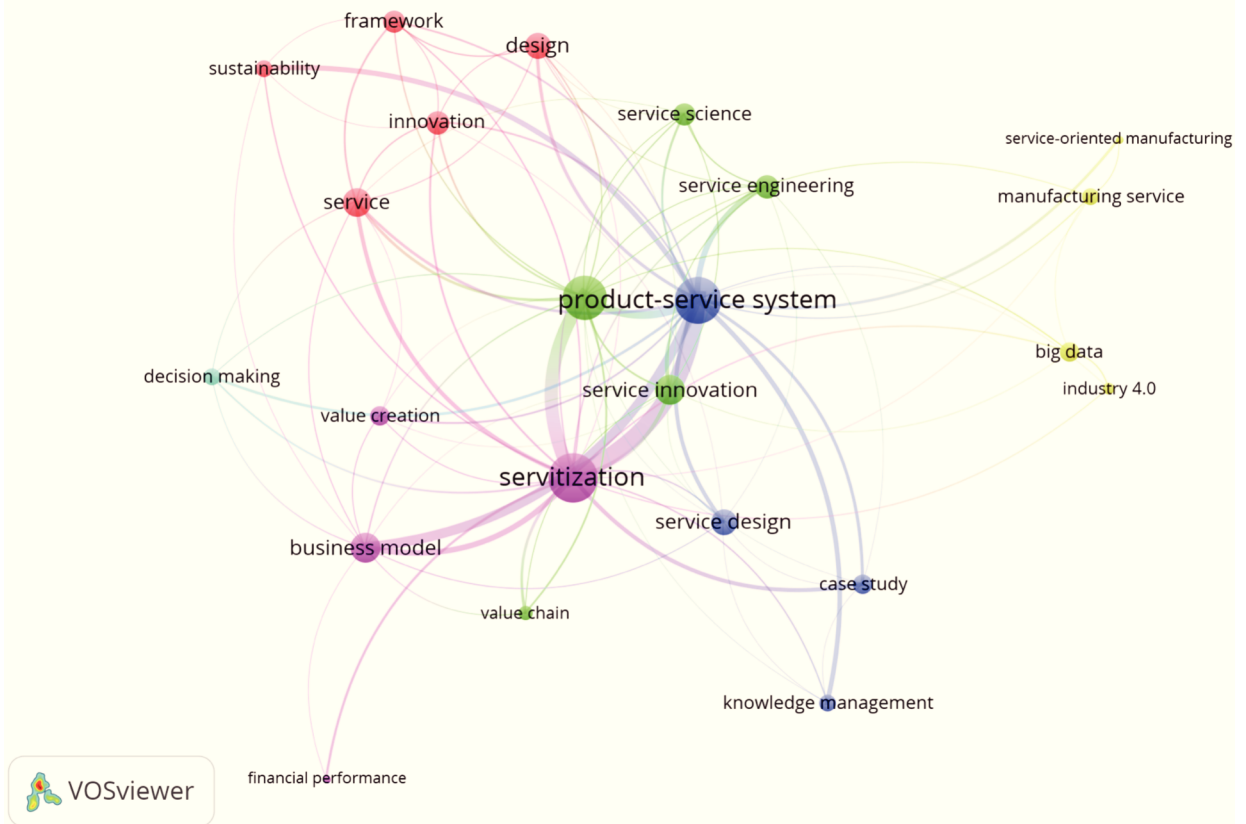
Fonte: Autoria própria (2018)

segunda possibilidade é a formação de clusters que relacionam o aparecimento destas palavras em função do tempo por meio de um gráfico que relaciona as cores e anos. Este segundo tipo de rede é mostrado na Figura 17. Por fim, é possível visualizar a densidade do uso das palavras-chave por meio de um gráfico em cores, no qual as cores vermelha e verde significam um alto e baixo número de publicações, respectivamente. Esta última rede bibliográfica é apresentada por meio da Figura 18.

A Figura 16 apresenta a formação de seis clusters: (1.) *design, framework, innovation, service e sustainability*, (2.) *manufacturing, service engineering, service innovation, service science e value chain*, (3.) *product-service system, service design, knowledge management e case study*, (4.) *servitization, value creation, business model e financial performance*, (5.) *big data, industry 4.0, manufacturing service e service-oriented manufacturing* e (6.) *decision making*. Ou seja, agrupamentos de palavras-chave que possuem forte vínculo e aparecem de forma conjunta nas publicações.

Analisando a Figura 17, visualiza-se que há uma tendência de oportunidades de pesquisa científica relacionada aos serviços e manufatura direcionar-se para a exploração de dados, traduzida nas palavras-chave *Big Data* e *Industry 4.0*. Por exemplo, no múltiplo estudo de caso apresentado por Cenamor, Sjödin e Parida (2017), as empresas de manufatura entrevistadas já enxergam o uso de dados (informação) como elemento central para a proposição de valor, pois tem potencial para geração de receita e diminuição dos custos associados ao oferecimento de serviços. Desta forma, evita-se o paradoxo dos serviços – o fato do oferecimento de serviços gerar mais receita e ao mesmo tempo diminuir os lucros devido aos seus altos custos de

Figura 16 – Rede bibliográfica com *clusters* de palavras inter-relacionadas de acordo com o grau de vínculo.



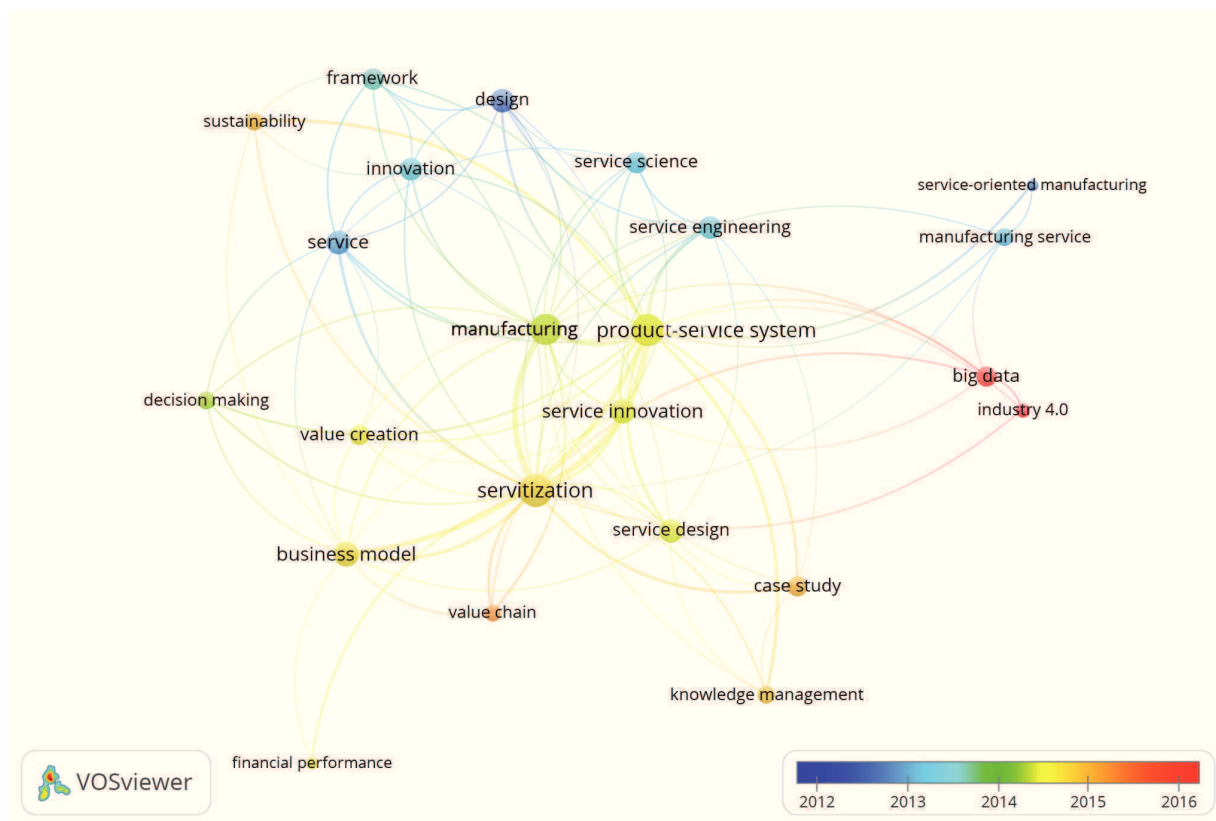
Fonte: Autoria própria (2018)

implementação (GEBAUER; FLEISCH; FRIEDLI, 2005) – pois estimula a eficiência operacional.

Na Figura 18, é observado que a inovação em serviços é um elemento altamente correlacionado à servitização e aos PSS. A inovação em serviços é descrita como um dos fatores mais importantes para as empresas de manufatura alcançarem vantagem competitiva (M. ZHANG *et al.*, 2016). No entanto, poucos estudos acadêmicos têm mostrado o que é necessário para ser uma empresa de manufatura eficiente na inovação em serviços (CHENG; KRUMWIEDE, 2017). Este tipo de pensamento é compartilhado por outros pesquisadores, que afirmam que há uma falta de métodos de pesquisa quantitativos por meio de entrevistas e/ou análises estatísticas de dados endereçando implementações reais de serviços inovadores, sendo estas implementações, um grande desafio neste campo de pesquisa (TUKKER, 2015; CENAMOR; SJÖDIN; PARIDA, 2017; RABETINO; KOHTAMÄKI; GEBAUER, 2017).

Por fim, nesta seção observou-se que há uma corrente transformação digital do serviço na manufatura por meio do uso intensivo de dados que pode ser observado nos ambientes acadêmico e organizacional. Segundo a literatura científica, o desafio de implementar novas soluções e obter resultados quantitativos é sinalizado como indispensável para a desenvolvimento do tema.

Figura 17 – Rede bibliográfica com *clusters* de palavras inter-relacionadas de acordo com o ano de publicação.



Fonte: Autoria própria (2018)

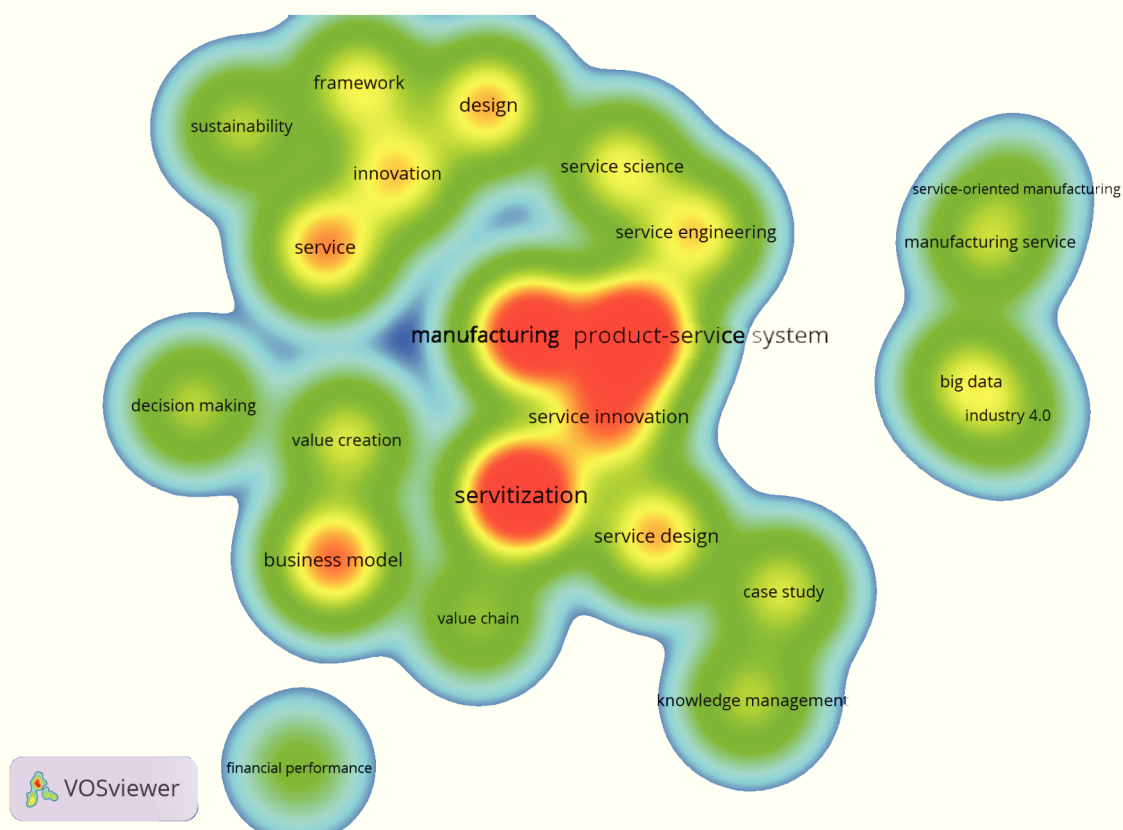
2.4 SELEÇÃO DE FORNECEDORES

A manufatura é um ambiente com incertezas associadas à demanda, fornecimento, custo, lead time e sujeita a desastres catastróficos (RAJAGOPAL; VENKATESAN; GOH, 2017). Os autores afirmam que estas incertezas podem resultar em perdas econômicas, insatisfatório nível de serviço ao consumidor e perda da reputação da empresa no mercado. Portanto, o controle e atenuação destas incertezas resulta em um melhor desempenho da atividade de manufatura.

De acordo com um estudo conduzido pela PwC e Business Continuity Institute em 2013, 75% das empresas enfrentam pelo menos uma grande interrupção na cadeia de suprimentos por ano, sendo a maioria destas interrupções causadas por problemas de fornecimento (YOON *et al.*, 2018). Ademais, à medida que a estrutura da cadeia de suprimentos se torna maior, mais complexa e globalizada, as empresas de manufatura se tornam cada vez mais dependentes dos seus fornecedores (YOON *et al.*, 2018). Neste contexto, a seleção de fornecedores se torna um elemento chave no ambiente de manufatura, pois pode aumentar a previsibilidade do fornecimento com potencial de resultar em benefícios como uma produção mais eficiente, sincronizada e resiliente.

Tradicionalmente, a seleção de fornecedores está principalmente associada

Figura 18 – Rede bibliográfica de densidade de palavras de acordo com a quantidade de publicações.



Fonte: Autoria própria (2018)

aos custos. No entanto, outros fatores como qualidade, entrega, flexibilidade e sustentabilidade também têm sido considerados (YOON *et al.*, 2018). Hamdi *et al.* (2018) afirmam que a escolha do melhor fornecedor geralmente está associada àquele que consegue entregar o produto certo, na hora certa, no local certo, na quantidade certa a um preço competitivo.

Portanto, para a etapa de seleção de fornecedores, a empresa deve avaliar os aspectos – preço, qualidade do produto, capacidade de oferecer quantidade e serviço, como também levar em conta outros fatores de risco como as incertezas de fornecimento, vulnerabilidade do fornecedor e possibilidades de interrupção do fornecimento (HAMDI *et al.*, 2018). A Tabela 5, adaptada de A. Chen, Hsieh e H. Wee (2016), apresenta os principais critérios e suas definições para a seleção de fornecedores. Observa-se que vários critérios podem inter-relacionar-se, ou seja, apresentar algumas definições e variáveis similares, e de fato, Hamdi *et al.* (2018) afirmam que considerar os critérios como interdependentes representa melhor a realidade do que considera-los independentes.

Tabela 5 – Principais critérios para seleção de fornecedores.

Item	Critério	Definição
1	Financeiro	Rentabilidade, custo (do produto, logística, taxas e impostos), lucro bruto, desconto por quantidade, marcação (porcentagem contratual de redução nas despesas indiretas unitárias relacionados aos custos diretos unitários de fabricação), custos de material ecologicamente sustentável, custos para compra de novos equipamentos, custos de projeto de produtos, treinamento de colaboradores, custos de reciclagem, retorno sobre os ativos, preço (de material, montagem, transporte, gerenciamento e negociação), prestação de contas, incentivos, índice recíproco de preço, custo de encomenda de peças, preço unitário das peças necessárias para os clientes, desconto por volume de negócios, preço dos fatores, custo fixo e custo variável.
2	Qualidade	Conformidade com a especificação, confiabilidade do produto, capacidade técnica de avaliação de qualidade, qualidade do produto, sistema de qualidade (presença ou ausência de certificações), instalações de produção, informação de gerenciamento de qualidade, resultado do sistema de qualidade, reclamações, qualidade de melhorias, taxa de rendimento, confiabilidade, habilidade de reparo, qualidade taxa de rejeição, taxa de defeito aceitável/esperado, nível técnico e defeitos.

Continua na próxima página

Tabela 5 – Principais critérios para seleção de fornecedores.

Item	Critério	Definição
3	Entrega	Pontualidade de entrega, desempenho de entrega, lead time, taxa de entrega no prazo, flexibilidade de entrega, programação da entrega, tempo de processamento (tempo necessário para desenvolver o projeto estrutural do produto), tempo de prototipagem (velocidade na construção de protótipos), tempo de revisão do projeto (flexibilidade para aceitar e executar revisões de projeto), serviço (facilidade de comunicação, resposta a mudanças e flexibilidade do processo), resposta a reclamações, precisão, localização, distância, número de entregas padrão e taxa de entrega atrasada aceitável.
4	Relacionamento	Cooperação técnica, cooperação de mercado, tempo de cooperação, porcentagem de fornecedores de longo prazo aceitáveis, crescimento conjunto, construção de relacionamento, confiança, credibilidade, consistência, confiança mútua, facilidade de comunicação, atrasos nos pedidos, incapacidade de atender requisitos adicionais e relação de longo prazo.
5	Serviço	Suporte ao cliente, desempenho do serviço do fornecedor (confiabilidade de entrega, compartilhamento de informações, flexibilidade e capacidade de resposta e resposta do cliente), reclamações de clientes, desempenho de atendimento, capacidade de resposta, capacidade de melhoria, atitude, comunicação, velocidade de resposta, uso de tecnologia, profissionalismo do vendedor, capacidade de resposta às necessidades do cliente, solução de problemas, serviços de montagem, pedidos on-line, número de peças necessárias para clientes, foco no cliente, atendimento de pedidos, capacidade de fornecimento, tempo de reparo e oferecimento de garantia.
Continua na próxima página		

Tabela 5 – Principais critérios para seleção de fornecedores.

Item	Critério	Definição
6	Tecnologia e Produto	Tecnologia e produto: taxa de P&D, capacidade do processo, conformidade do produto, codesign (esforço do fornecedor dentro da equipe do projeto), níveis tecnológicos do fornecedor, tempo de ciclo de desenvolvimento do produto, dinamismo do produto, família de produtos, flexibilidade do produto, garantia do produto e recursos para monitoramento do desempenho.
7	Reputação do fornecedor	Infraestrutura, qualidade das instalações e reputação de mercado do fornecedor.
8	Gerenciamento e organização	Controle da organização, planos de negócios, comunicação com o cliente, auditoria interna e administração de dados, honestidade e integridade, ajuste fornecedor-comprador, compromisso estratégico do fornecedor com o comprador, segurança, institucionalização, flexibilidade, experiência de trabalho, prioridade competitiva, compras estratégicas, suporte de alta gerência, tecnologia da informação, fornecedor único ou múltiplo, alternativas de fornecedores, reputação do fornecedor, capacidade do sistema de informação, interoperabilidade com outras partes, capacidade de atualização de hardware ou software, inovação, pesquisa e desenvolvimento.
9	Adaptabilidade	Eficácia de ações corretivas e capacidade de resolução de problemas e conflitos.
10	Ambiental	Emissão de gás carbônico (CO ₂).
11	Fatores de risco	Localização geográfica, estabilidade política, política externa, taxa de câmbio e posição econômica, taxa de criminalidade, terrorismo, restrição de oferta e restrição ao fornecedor comprador.

Fonte: Adaptado de A. Chen, Hsieh e H. Wee (2016)

A maioria das empresas de manufatura gasta um montante considerável das suas receitas em compras de materiais de fornecedores, representando de 60% a 80% dos custos de produção (DEY *et al.*, 2015). Logo, a seleção de fornecedores

adequados tem potencial de não só reduzir os custos de compras de maneira significativa, como também contribuir para a inovação e o alcance de processos produtivos mais eficientes e conseqüentemente contribuir para a manutenção da competitividade das empresas (DEY *et al.*, 2015; YOON *et al.*, 2018).

O estudo da seleção de fornecedores e os fatores que influenciam a tomada de decisão não são recentes, como pode ser observado na publicação datada da década de 60 por Dickson (1966), na qual são elaborados 23 critérios de desempenho após a realização de entrevistas com 170 representantes de empresas nos Estados Unidos e Canadá, principalmente no ramo da manufatura (67,8% do total), podendo se destacar os três critérios de maior importância: qualidade, entregabilidade e histórico de performance. De maneira geral, os 23 critérios apresentados por Dickson (1966) podem ser sintetizados e agrupados em quatro temas: qualidade, entregabilidade, performance e políticas de garantia (P.-S. CHEN; M.-T. WU, 2013).

Dois publicações mais recentes com caráter de revisão que relacionam o gerenciamento de riscos na cadeia de suprimentos e a seleção de fornecedores podem ser destacados – os trabalhos de Hamdi *et al.* (2018) e Rajagopal, Venkatesan e Goh (2017).

Hamdi *et al.* (2018) coletou artigos no período de 2003 a 2014 e realizou uma revisão sistemática da literatura, totalizando 124 artigos a serem analisados com relação ao tema seleção de fornecedores e gerenciamento de riscos na cadeia de suprimentos. Já Rajagopal, Venkatesan e Goh (2017) revisaram trabalhos no período de 2005 a 2016, totalizando 538 publicações, sendo 126 delas selecionadas para a revisão final.

Todos realizam uma revisão de literatura, correlacionando os temas de seleção de fornecedores e gerenciamento de riscos na cadeia de suprimentos, sendo o primeiro como um subconjunto do segundo. A Figura 19, adaptada de Rajagopal, Venkatesan e Goh (2017), ilustra os subconjuntos no contexto da mitigação de riscos na cadeia de suprimentos com suas principais palavras-chave. Ainda, Hamdi *et al.* (2018) subdividem os artigos coletados em cinco tipos de abordagens sendo: 1.) Abordagens Quantitativas, 2.) Abordagens Qualitativas, 3.) Abordagens utilizando ferramentas de Simulação, 4.) Abordagens utilizando Inteligência Artificial e 5.) Híbridas, que envolvem duas ou mais abordagens.

Tomlin (2006) afirma que os riscos podem ser divididos como recorrentes ou disruptivos. Para Jüttner, Peck e Christopher (2003), os riscos na cadeia de suprimentos podem ser classificados como riscos internos, riscos da cadeia de suprimentos ou riscos externos. O primeiro surge dentro da organização, o segundo surge externamente à organização, mas dentro da cadeia de suprimentos e, por fim, o terceiro surge externamente e fora da cadeia de suprimentos, ou seja, afeta diversas cadeias simultaneamente.

Figura 19 – Subconjuntos da mitigação de riscos na cadeia de suprimentos e principais palavras-chave.

Riscos na Cadeia de Suprimentos	<i>Cluster 1: Projeto de redes de suprimento resilientes/robustas e análise propagação de risco</i>	<ol style="list-style-type: none"> 1. Programação estocástica; 2. Programação Linear Inteira Mista; 3. Resiliência; 4. Robustez; 5. Projeto de redes de cadeia de suprimento; 6. Otimização robusta; 7. Riscos múltiplos; 8. Rede de interrupções; 9. Planejamento de cenários; 10. Antecipação.
	<i>Cluster 2: Seleção de fornecedores e alocação de compras sob risco de interrupção</i>	<ol style="list-style-type: none"> 1. Risco de interrupção; 2. Seleção de fornecedores; 3. Programação Linear Inteira Mista; 4. Fornecimento múltiplo; 5. Interrupção de fornecimento; 6. Fornecimento duplo; 7. Programação estocástica; 8. Incerteza de demanda; 9. Valor da condição em risco (CVaR); 10. Fornecimento único.
	<i>Cluster 3: Gerenciamento de inventário sob risco de interrupção e confiabilidade de localização de instalações</i>	<ol style="list-style-type: none"> 1. Interrupção; 2. Localização de instalação; 3. Confiabilidade; 4. Heurística; 5. Aproximação contínua; 6. Competição; 7. Relaxação Lagrangeana; 8. Interrupção da cadeia de suprimentos; 9. Inventário; 10. Jogo de Stackelberg.
	<i>Cluster 4: Coordenação, precificação e contratos de compartilhamento de riscos para mitigação de interrupção de demanda e custos</i>	<ol style="list-style-type: none"> 1. Teoria dos Jogos; 2. Aversão ao Risco; 3. Otimização; 4. Interrupção da demanda; 5. Contrato de receita compartilhada; 6. Canal duplo; 7. Mecanismos de coordenação; 8. Nível de serviço; 9. Políticas de inventário; 10. Modelo estocástico.

Fonte: Adaptado de Rajagopal, Venkatesan e Goh (2017)

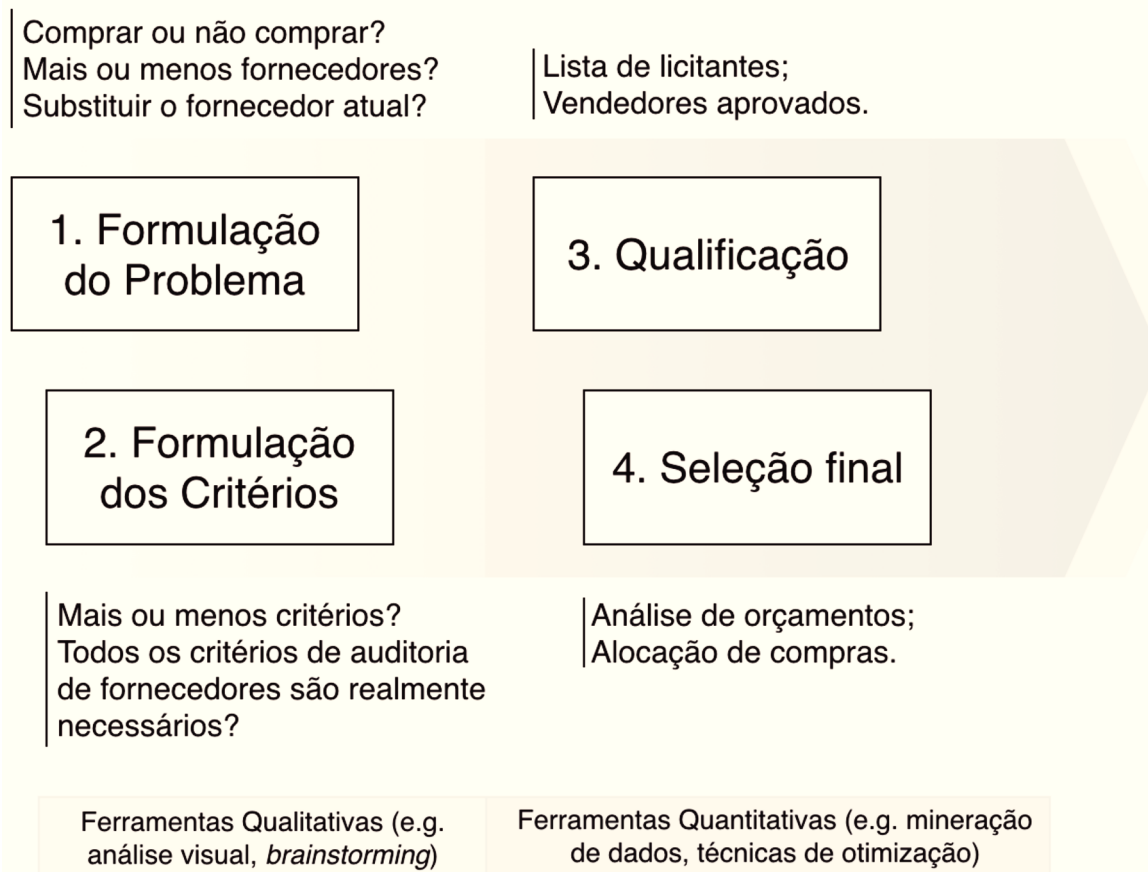
Outro aspecto está na frequência de ocorrência dos três tipos de risco. Os riscos internos possuem maior probabilidade de ocorrência do que os riscos da cadeia de suprimentos/externos mas apesar disso, quando ocorrem, os últimos impactam mais a operação do que o primeiro (HAMDI *et al.*, 2018). Os autores ainda afirmam que os riscos internos podem ser exemplificados como problemas de qualidade, quebra de máquinas, greves trabalhistas, volatilidade da demanda e restrições de capacidade. Já os riscos da cadeia de suprimentos podem ser ilustrados como desastres naturais que afetam uma determinada região, e.g., terremotos, enchentes e incêndios.

Por fim, o risco externo pode ser exemplificado por meio de crises econômicas, ataques terroristas, instabilidade política/econômica e greves generalizadas no setor de transporte (HAMDI *et al.*, 2018). Um recente exemplo na história do Brasil para este tipo de risco foi a paralisação dos caminhoneiros que, iniciada no dia 21 de maio de 2018, afetou diversos setores e cadeias de suprimento simultaneamente.

Em uma abordagem mais holística, De Boer, Labro e Morlacchi (2001) afirmam que a seleção de fornecedores é um processo complexo que consiste em quatro fases:

1.) Formular os problemas dos fornecedores, 2.) Definir os critérios de seleção, 3.) Pré-qualificar os fornecedores adequados e 4.) Selecionar o fornecedor final. A Figura 20 ilustra as fases para seleção de fornecedores.

Figura 20 – Fases do modelo de decisão para seleção de fornecedores.



Fonte: Adaptado de De Boer, Labro e Morlacchi (2001)

Mais recentemente, diferentes métodos foram compilados pela literatura com relação à seleção de fornecedores. No trabalho de A. Chen, Hsieh e H. Wee (2016), os autores apresentam duas figuras no formato mapa de árvore nas quais os métodos são associados a quatro grupamentos: métodos matemáticos, de análise, inteligência artificial e combinados. Já no trabalho de (HAMDI *et al.*, 2018) são apresentados também os grupamentos: métodos matemáticos, de simulação, qualitativos e combinados. Visto que as figuras apresentadas pelos autores se complementam, a Figura 21 sintetiza os itens expostos por ambos autores. Os métodos combinados não foram incluídos na Figura 21, pois resume-se à combinação de quaisquer métodos matemáticos, qualitativos, de simulação e utilizando inteligência artificial, podendo inclusive ser a combinação entre métodos da mesma categoria.

Tendo em vista a elevada quantidade de métodos de seleção de fornecedores reconhecidos pela literatura científica, esta breve revisão de literatura apresenta os métodos mais populares no âmbito das abordagens qualitativas e quantitativas. Entende-se por métodos mais populares como os três que possuem maior número

Figura 21 – Métodos para seleção de fornecedores.

Métodos Matemáticos	Qualitativos	Simulação	Inteligência Artificial
Programação linear inteira mista (MILP); Programação multi-objetiva (MOP); Heurísticas; Programação linear (LP); Técnica para avaliar o desempenho de alternativas através de similaridade com a solução ideal (TOPSIS); Programação por metas (GP).	Árvores de decisão (DT); Análise hierárquica de processo (AHP); Teoria dos jogos (GT); Processo de rede analítica (ANP); Análise do ambiente de dados (DEA); Análise dos modos de falha e efeitos (FMEA); Desdobramento da função qualidade (QFD).	Modelagem baseada em agentes (ABM); Simulação de eventos discretos (DES); Monte Carlo; Redes de Petri; Cadeias de Markov; Jogo da cerveja.	Redes neurais (NN); Teoria dos conjuntos fuzzy (FST); Sistema de colônia de formigas (ACS); Mineração de dados; Particionamento recursivo Sistemas especialistas (ES); Tomada de decisão baseada em critérios múltiplos; Método Delphi.

Fonte: Adaptado de A. Chen, Hsieh e H. Wee (2016) e Hamdi *et al.* (2018)

de publicações para cada tópico, de acordo com a revisão bibliográfica realizada por Hamdi *et al.* (2018).

A maioria das ferramentas para formulação de problemas por meio de métodos qualitativos tem o objetivo de auxiliar especialistas e tomadores de decisões em identificar as alternativas disponíveis para a tomada de decisão (MIRKOUËI; HAAPALA, 2014). As abordagens qualitativas no âmbito da seleção de fornecedores podem ser resumidas com a utilização de pontuações ponderadas de critérios pré-selecionados como significativos para a seleção de fornecedores.

No entanto, estas abordagens ponderadas baseadas em políticas de decisões, apesar de simples e fáceis de se entender, dependem da experiência e entendimento do usuário/pesquisador (Z. LI *et al.*, 2015). Além disso, os autores afirmam que estas abordagens não consideram a natureza estocástica das cadeias de suprimento. Alguns exemplos de fatores de risco qualitativo incluem fatores econômico-ambientais e classificações de fornecedores, expressas por dados fuzzy que são quantificados por um grau de confiança (HAMDI *et al.*, 2018). Três abordagens qualitativas comuns podem ser citadas: *Game Theory*, *Decision Tree* e *Analytic Hierarchy Process*.

Game Theory (GT) ou Teoria dos Jogos consiste na definição de ações que os jogadores podem tomar para assegurar a sua vitória contra jogadores a quem se está jogando contra (HAMDI *et al.*, 2018). Segundo T. Liu, Deng e Chan (2018), um resultado de jogo dito otimizado advém quando nenhum jogador têm o incentivo de desviar de suas estratégias e escolhas iniciais após considerar as decisões do outro jogador. No caso da seleção de suppliers, esta técnica pode representar o comportamento estratégico de barganha e pode representar ações conflituosas entre dois ou mais participantes (T. LIU; DENG; CHAN, 2018).

Decision Tree (DT) ou árvore de Decisão consiste na utilização de um diagrama tipo árvore como ferramenta de suporte para estruturar um problema de tomada de decisão sistematicamente (OMAR, 2011). Esta abordagem, baseada na análise de um diagrama, pode auxiliar a tomada de decisão na presença de diferentes opções e critérios conflituosos, sendo o objetivo desta abordagem a captura do risco e o nível

de risco de cada decisão (HAMDI *et al.*, 2018). Com a abordagem por DT pode-se capturar diferentes cenários de interrupção, completos ou parciais, e determinar um número ótimo de fornecedores (OMAR, 2011; KAMALAHMADI; PARAST, 2017).

Analytic Hierarchy Process (AHP) ou Análise Hierárquica de Processo é uma técnica estruturada que compara e avalia a influência de vários critérios de seleção de fornecedores, sendo o objetivo a seleção da melhor alternativa (HAMDI *et al.*, 2018). Nesta técnica, diferentes níveis de importância são utilizados para classificar os critérios de seleção, i.e. o critério mais importante, os critérios muito importantes, relevantes, e assim por diante, até o menos importante. De maneira geral, o processo pode ser separado em três etapas: 1. Construção de uma estrutura hierárquica por meio da divisão recursiva do problema, 2. Estabelecimento de uma matriz de comparação pareada para indicar a importância relativa das alternativas, e 3. Calcular os pesos prioritários das alternativas de acordo com a matriz de comparação pareada (HOSSEINI; BARKER, 2016). Desta forma, o método prioriza as alternativas de fornecedores baseada nos múltiplos critérios e auxilia na interpretação devido ao fato do problema estar estruturado hierarquicamente por meio de níveis integrados (OMAR, 2011). A AHP possui desvantagens como a dependência do julgamento humano e incapacidade de introduzir restrições nos modelos (HAMDI *et al.*, 2018).

Nas abordagens quantitativas, os fatores ou critérios em estudo podem ser medidos e quantificados numericamente. No caso da seleção de fornecedores, podem-se citar o custo ou confiabilidade de entrega (WETZSTEIN *et al.*, 2016). Hamdi *et al.* (2018) ainda complementam que fatores de risco quantitativos podem ser expressos por dados estocásticos por meio de distribuições estatísticas.

Técnicas de programação matemática são muito importantes para resolução de problemas de otimização e buscam o melhor plano de alocação de recursos possível diante de diferentes restrições impostas pela natureza do problema (HAMDI *et al.*, 2018). Podem-se citar as três abordagens quantitativas mais comuns: *Multi-Objective Programming* (MOP), *Mixed Integer Linear Programming* (MILP) e *Heuristics*.

Multi-Objective Programming (MOP) ou Programação Multi-Objetiva é um conjunto de programas matemáticos os quais a solução ótima requer a consideração de dois ou mais objetivos que estão em conflito (HAMDI *et al.*, 2018). Os modelos MOP têm como objetivo selecionar o fornecedor preferencial tendo em vista, por exemplo, a minimização de custos, uma resposta rápida ou uma entrega pontual (D. WU; OLSON, 2008). Ademais, em um outro modelo realizado por Sawik (2016), a formulação do modelo MOP utiliza quatro critérios: o preço, o lead time, o risco de interrupção devido a eventos naturais e o nível de qualidade, sendo estes, os quatro objetivos formulados a serem minimizados.

Mixed Integer Linear Programming (MILP) ou Programação Linear Inteira Mista utiliza uma função objetivo linear sujeita a restrições lineares as quais algumas das

variáveis ou todas elas necessariamente precisam ser números inteiros (HAMDÍ *et al.*, 2018).

Heuristics ou Heurística é uma abordagem utilizada para resolver problemas de maneira mais rápida ou para achar uma solução aproximada quando métodos tradicionais são muito lentos ou falham ao achar uma solução exata, respectivamente (HAMDÍ *et al.*, 2018). Pesquisadores optam por utilizar heurísticas, pois, dependendo do tamanho e complexidade do problema ou pela natureza não linear dos problemas da cadeia de suprimentos, soluções matemáticas não são sempre factíveis, e portanto, é necessário o desenvolvimento de heurísticas e meta-heurísticas, i.e. heurísticas em combinação a métodos exatos, para alcançar resultados satisfatórios entre qualidade da solução e esforço computacional (RAJAGOPAL; VENKATESAN; GOH, 2017).

Após a definição dos principais métodos do estado atual de seleção de fornecedores apresentados pela literatura, a Tabela 6 evidencia as principais vantagens e desvantagens dos métodos até aqui apresentados. A tabela inclui abordagens por simulação e inteligência artificial, que são objeto de estudo principal desta dissertação de mestrado.

Tabela 6 – Vantagens e desvantagens dos métodos de seleção de fornecedores.

Métodos	Vantagens	Desvantagens
Qualitativos (multi-atributos)	Rápido e fácil de utilizar	Depende do julgamento humano
	Leva em conta critérios subjetivos	Não pode introduzir restrições no modelo
	Pouco dispendioso	
Matemáticos	Mono-objetivo: obtém solução ótima	Não leva em conta o julgamento humano
	Mono-objetivo: novas restrições podem ser introduzidas no modelo	
	Multi-objetivo: critérios não necessariamente têm uma dimensão comum	Leva em conta o julgamento humano
	Multi-objetivo: fornecem diversas soluções	Não fornece solução ótima
	Multi-objetivo: possibilidade de introduzir novas restrições	Difícil a análise do resultado
	Heurísticas: fornecem soluções próximas às ótimas	Não fornece solução ótima

Continua na próxima página

Tabela 6 – Vantagens e desvantagens dos métodos de seleção de fornecedores.

Métodos	Vantagens	Desvantagens
Matemáticos	Heurísticas: solucionam problemas complexos	Não fornece solução ótima
Simulação	Gera muitas soluções	Grande número de iterações (muitas vezes não é muito claro)
		Não fornece um valor específico da função objetivo
Inteligência Artificial	Fornecer conhecimento flexível	Coleta informações dos fornecedores
	Leva em conta critérios qualitativos	Acesso restrito à expertise e difícil implementação

Fonte: Adaptado de Hamdi *et al.* (2018)

Após esta breve revisão bibliográfica com relação à ciência da seleção de fornecedores no âmbito da manufatura, é possível constatar que há uma grande variedade de métodos disponíveis que visam mitigar riscos inerentes às incertezas que tangem principalmente: o nível de serviço, a confiabilidade em relação à qualidade e quantidade do fornecimento de suprimentos, e as possibilidades de interrupção da cadeia. O grande volume de publicações e diversidade de métodos e modelos têm sido de grande valia para a comunidade científica para entendimento da dinâmica das relações entre empresas de manufatura e também, internamente, da interrelação dos níveis operacional, tático e estratégico. Porém, naturalmente, a visão gerencial num período de transformação digital da manufatura é fortemente influenciada pela criação efetiva de valor, e portanto, a aferição de benefícios no curto e médio prazo se torna uma característica desejável. Além disso, alguns trabalhos desenvolvidos possuem vieses que, eventualmente, podem não ser sustentáveis diante do novo contexto da Indústria 4.0. Por exemplo, no trabalho realizado por D. Wu e Olson (2008), os autores afirmam que relacionamentos duradouros e permanentes nas cadeias de suprimento geralmente resultam em benefícios como custos de compra mais baixos podendo, conseqüentemente, repassar este preço menor para o cliente final. Embora este ponto de vista seja compartilhado por outros autores como Sheffi e Rice Jr. (2005) e A. Chen, Hsieh e H. Wee (2016), os quais enfatizam que a lealdade na relação

fornecedor-cliente pode beneficiar cadeia tornando-a mais resiliente a crises e a flutuações de demanda, os autores não abordam o fato de que num período de transformação digital com posteriori utilização plena de novas tecnologias, a mentalidade torna-se fundamentada e/ou orientada à dados e sendo baseadas nestes as tomadas de decisão.

Diante deste cenário é possível notar um gap entre os métodos do estado atual da literatura científica e as futuras necessidades e demandas das empresas as quais a relação entre cliente-fornecedor é efêmera com forte influência da utilização de dados. Por fim, neste trabalho busca-se explorar este gap por meio de uma abordagem quantitativa utilizando algoritmos de *machine learning*, que têm potencial para alcançar sistemas produtivos inteligentes (WUEST *et al.*, 2016), com foco na seleção de fornecedores. As técnicas e ferramentas utilizadas neste trabalho são abordadas nas seguintes seções.

2.5 SIMULAÇÃO DE SISTEMAS PRODUTIVOS

Para Will M. Bertrand e Fransoo (2002), a simulação consiste no uso de modelos ou técnicas computacionais que imitam o comportamento de sistemas produtivos, de forma a descrever, analisar e melhor compreender seu funcionamento. De forma análoga, Miyagi (2006) afirma que simulação pode ser entendida como a “imitação” de uma operação ou de um processo do mundo real, e é útil para análise das características operacionais de um sistema.

Historicamente, uma das técnicas mais plausíveis para a geração de soluções em sistemas reais e complexos é por meio da modelagem e simulação, pois aspectos como configurações físicas e regras operacionais podem ser considerados, o que torna a simulação uma poderosa ferramenta para análise e avaliação de sistemas complexos e estocásticos (VIEIRA *et al.*, 2017). Uma grande vantagem na utilização de modelos de simulação é a sua habilidade de lidar com processos complexos, tanto determinísticos como estocásticos, sem nenhuma sofisticação matemática (FIGUEIRA; ALMADA-LOBO, 2014).

De maneira geral, existem três tipos de simulação que são utilizados no contexto da manufatura: dinâmica de sistemas (SD), simulação baseada em agentes (ABM) e simulação de eventos discretos (DES). Na SD, os processos do mundo real são representados como estoques e fluxos e, dentre os três tipos de simulação, é considerado o que possui maior nível de abstração (RONDINI *et al.*, 2017). As áreas de aplicação da SD envolvem o nível estratégico de tomada de decisão, gerenciamento de projetos, gerenciamento de cadeia de suprimento, assim como análises quantitativas de gerenciamento do conhecimento (JAHANGIRIAN *et al.*, 2010).

Segundo Rondini *et al.* (2017), a ABM é a abordagem mais recente de simulação e as entidades possuem comportamento ativo e respondem a estímulos do sis-

tema de acordo com suas próprias regras. Desta forma, este tipo de simulação é influenciada pelo comportamento dos indivíduos por meio da perspectiva *bottom-up* – quando a estratégia da empresa é influenciada pelas experiências e aprendizado diário operacional (SLACK; CHAMBERS; JOHNSTON, 2013) – e portanto, o comportamento geral do sistema é o resultado das várias iterações entre as várias entidades (RONDINI *et al.*, 2017).

A DES é a técnica mais utilizada e caracterizada como “centrada no processo”, aborda tanto a dimensão tática quanto operacional, baseando-se no fluxo de entidades, compartilhamento de recursos e sequência de atividades (RONDINI *et al.*, 2017). Devido à importância da DES na manufatura e por ser a técnica de simulação utilizada nesta dissertação, alguns conceitos fundamentais devem ser definidos, segundo Georges *et al.* (2005):

- Sistema: um sistema é uma coleção de entidades que interagem ao longo do tempo para realizar um, ou mais, objetivos;
- Modelo: é uma representação abstrata de um sistema, usualmente estruturado logicamente com relações matemáticas que descrevem o sistema em termos do estado, entidades, atributos, conjuntos, processos, eventos, atividades e esperas. O modelo deverá ser o mais simples possível, mas complexo o suficiente para responder às questões levantadas sobre o comportamento do sistema a ser estudado.
- Estado do sistema: uma coleção de variáveis que contém toda a informação necessária para descrever o sistema, em qualquer tempo, em um nível de detalhe desejado. Durante o processo de modelagem, muitas simplificações sobre o sistema são feitas. Isto acarreta uma diminuição significativa dos possíveis estados que o sistema pode assumir. Saber fazer estas simplificações sem prejuízo na representação do sistema é o ponto central na modelagem.
- Evento: uma instantânea ocorrência que muda o estado do sistema. Um evento pode ser endógeno ou exógeno, ou seja, interno ou externo. Um evento interno é, por exemplo, o término de uma tarefa, um evento externo pode ser a chegada de mais um consumidor ao sistema.
- Entidade: qualquer objeto ou componente no sistema que requer explícita representação no modelo (consumidores, máquinas, equipamento, etc.). A entidade pode ser dinâmica, ou seja, se move na simulação, como veículos de transportes e consumidores, ou então a entidade pode ser estática, como um servidor ou uma máquina.
- Atributo: uma propriedade de uma dada entidade. Esta propriedade pode pertencer a uma única entidade ou pertencer a um conjunto de entidades. Por exemplo,

um atributo pode ser o tempo de chegada dos consumidores, ou então o tempo de atendimento de um caixa, ou uma característica do consumidor (homem, mulher, idoso ou preferencial), ou do produto (produto A, produto B, etc.).

- Recursos: um recurso é uma entidade que presta serviço para uma (ou mais) entidade dinâmica. Um produto (que é uma entidade) requer uma determinada máquina para ser processado, a máquina é um recurso (e também uma entidade). Há vários possíveis estados para os recursos, mas os mais usuais são: ocupado, ocioso, em manutenção, quebrado e em preparação (setup).
- Lista: as entidades são gerenciadas alocando-as aos recursos que as processam através de uma ordem de atendimento, representada pelas listas. As listas são usadas para representar filas de entidades à espera do recurso, ordenadas em alguma lógica, FIFO (*First In, First Out*), LIFO (*Last In, First Out*), SPT (*Shortest Processing Time*) de acordo com o atributo da entidade (preferencial ou não preferencial), ou mesmo aleatoriamente.
- Atividade: a duração especificada de tempo, com sabido início e término. Uma atividade geralmente é associada a um recurso, cujo atributo é conhecido. Quando a atividade se inicia, sabe-se, portanto, quando será seu fim.
- Lista de eventos futuros: quando ocorre um evento que denota o início de uma atividade, o evento que denotará seu fim é conhecido, pois o tempo da atividade é conhecido, assim é possível compor uma lista de eventos futuros.
- Espera: a duração de tempo não é especificada, ou seja, não se sabe quando terminará. Geralmente a espera é associada a uma fila. Por exemplo, quando uma atividade entra em uma fila de espera de um determinado recurso não se sabe quanto tempo esta entidade permanecerá na fila. Embora se saiba o tempo de processamento do recurso, geralmente o que não se sabe é o tamanho da fila e possibilidade de que ocorra algum evento externo que atrase ainda mais essa fila. De modo geral, o tempo de espera nas filas é uma variável de interesse que se deseja conhecer através da simulação.
- Carregamentos: componentes do modelo que representam pessoas ou objetos que se movimentam pelo sistema alterando seu status. Podem ser permanentes ou temporárias. São objetos dinâmicos da simulação, sendo normalmente criados, circulados pelo sistema, e então, retirados, deixando o sistema. Exemplos de carregamentos reais são peças, paletes, papéis, etc. Podem, ainda, representar elementos intangíveis como chamadas, correios eletrônicos, projetos, etc.
- Acumuladores estatísticos: são definidos para fornecer as medidas de desempenho do sistema como resposta. Quando ocorre algo na simulação, os acumu-

ladores estatísticos são atualizados, realizando um acompanhamento de uma determinada medida e fornecendo-a ao final da simulação.

- **Sistemas terminais:** aqueles que sempre retornam para uma condição inicial fixa; como exemplo, pode-se tornar um banco onde o começo do expediente bancário possui sempre a mesma condição inicial: a abertura dos caixas e a formação de filas pelos clientes que aguardavam a abertura do banco.
- **Sistemas não terminais:** aqueles que não possuem uma posição inicial fixa e nem um ponto natural de término; como exemplo, pode-se tomar um sistema formado por um forno que é alimentado continuamente durante 24 horas por dia.
- **Inicialização (Warm-up):** designa os momentos iniciais da simulação, onde o sistema ainda não atingiu o seu estado estacionário, tal como os primeiros minutos do início das operações de uma linha de produção, quando nem todas as máquinas estão em operação porque as peças a serem processadas ainda estão sendo processadas pela primeira das muitas máquinas que as processarão. De um modo geral este período de inicialização, ou aquecimento, é desprezado, e somente após todas as máquinas passarem a receber peças e, portanto, o sistema atingir o estado estacionário, é que se utilizam os dados coletados.
- **Relógio:** uma variável que representa o tempo na simulação.
- **Corrida:** é o ato de executar uma simulação do modelo construído a fim de se coletar as medidas de desempenho geradas pela simulação e efetuar as subsequentes análises.
- **Replicação:** é o ato de gerar uma nova corrida, coletando as medidas de desempenho do sistema simulado a fim de se fazer uma análise estatística sobre a coleção de dados de saída obtidos em cada replicação. Note que a natureza estocástica do sistema a ser simulado gerará diferentes dados de saída a cada nova replicação feita.
- **Semente de simulação:** é um número ou vetor que inicializa um gerador de números pseudoaleatórios. A semente de simulação é útil para garantir a repetitividade e reprodutibilidade dos experimentos.

A seguir são apresentadas as vantagens e desvantagens do uso da simulação discreta como ferramenta de análise a problemas reais. Os tópicos aqui apresentados são uma compilação dos trabalhos realizados por Miyagi (2006) e Chrispim (2007).

Vantagens:

- Experimentação em curto espaço de tempo: como o modelo é simulado computacionalmente, o tempo real pode ser simulado em segundos;
- Permitir que sistemas possam ser analisados dinamicamente com representação do tempo real: o avanço dos softwares contribuiu para que sistemas complexos pudessem ser modelados sem a necessidade de experiência em programação computacional por parte do analista;
- Modelos de fácil demonstração: a capacidade de representar graficamente de forma dinâmica um processo é considerada um dos recursos mais valiosos;
- Uso de animação durante uma apresentação estabelece a credibilidade do modelo: ela pode ser usada para descrever a operação e a interação dos processos do sistema simultaneamente;
- Novas políticas, procedimentos operacionais, regras de decisão, fluxos de informação e procedimentos organizacionais podem ser estudados sem interferência nas operações do sistema real;
- Novos equipamentos, arranjos físicos e sistemas de transporte podem ser testados antes de se investir recursos com as aquisições envolvidas;
- Hipóteses de como e por que certos fenômenos ocorrem podem ser avaliados;
- Gargalos onde as informações ou materiais têm seus fluxos comprometidos podem ser identificados.

Desvantagens:

- A simulação pode não dar resultados acurados com dados de entrada de má qualidade: não importa o quanto um modelo é bem desenvolvido, se ele não possuir os dados de entrada adequados, dados plausíveis não serão obtidos;
- A simulação não gera respostas simples para problemas complexos: se o sistema analisado tiver muitos componentes e interações, a melhor alternativa de operação é considerar cada elemento do sistema pois caso sejam ignorados, os resultados obtidos não serão satisfatórios;
- Somente a simulação não resolve problemas: a simulação não resolve problemas por si só, ela fornece ao tomador de decisão soluções potenciais para o problema. Cabe ao analista avaliar e implementar as mudanças que são propostas;
- A construção de modelos requer um treinamento especial: pode ser considerada uma arte que se aprende ao longo do tempo e que envolve o bom uso da experiência;

- Os resultados da simulação podem ser difíceis de interpretar: como as saídas da simulação podem incluir variáveis aleatórias, não é trivial determinar se os resultados observados resultam de inter-relações efetivas das partes do sistema ou se são fruto da aleatoriedade do sistema;
- A modelagem do sistema e a análise dos dados podem consumir muito tempo e muitos recursos, no entanto, economizar tempo e recursos na modelagem e na análise podem resultar em cenários que não representam a realidade.

De acordo com Rondini *et al.* (2017), o caso da simulação de eventos discretos, as entidades possuem comportamento passivo e é categorizado este tipo de simulação como nível médio de abstração (WEIDMANN *et al.*, 2015). Os diversos softwares de DES do tipo *open source* podem ser encontrados em Dagkakis e Heavey (2016). A Tabela 7 ilustra os principais softwares comerciais de DES:

Tabela 7 – Principais softwares comerciais de DES.

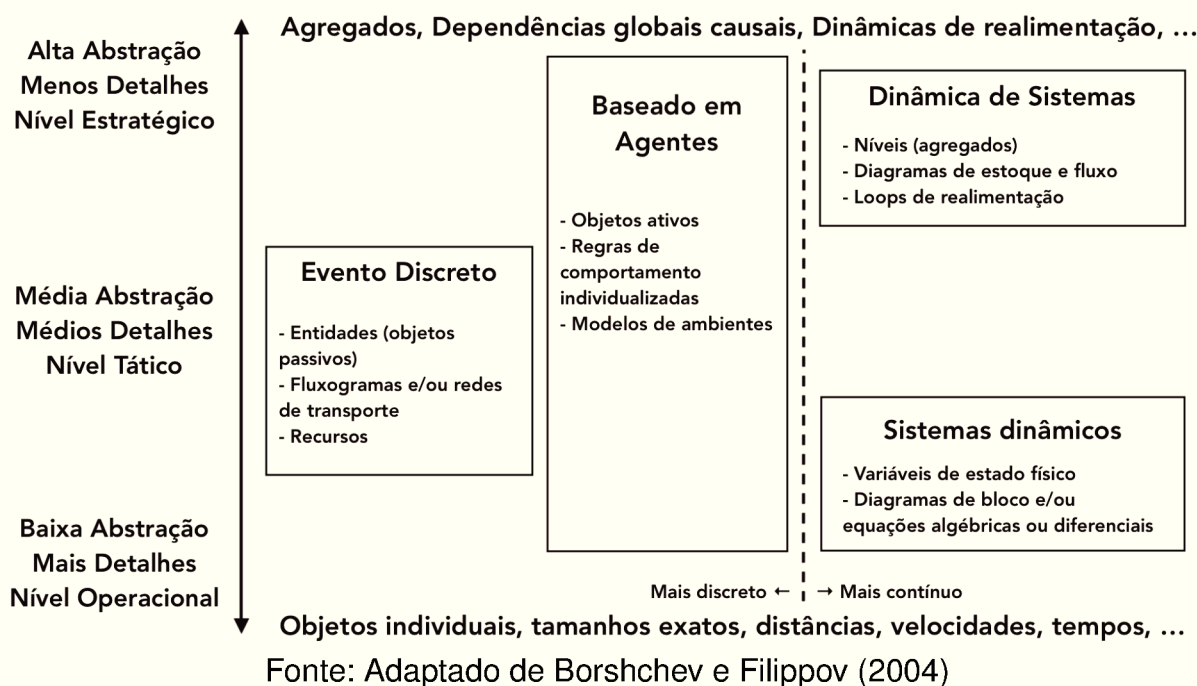
Software	Empresa	Descrição	Atualização
Anylogic	The Anylogic Company	Uma ferramenta gráfica de simulação de propósito geral que suporta eventos discretos, dinâmica de sistemas e abordagens de modelagem baseadas em agentes	25 de maio de 2015
Arena	Rockwell Automation	Ambiente gráfico integrado de simulação, que contém todos os recursos para modelagem de processos, desenho e animação, análise estatística e resultados	14 de agosto de 2017
Enterprise Dynamics	Incontrol Simulation Solutions	Uma plataforma de software de simulação	24 de maio de 2016
ExtendSim	Imagine That Inc.	Um pacote de software de simulação de propósito geral	13 de fevereiro de 2015
Continua na próxima página			

Tabela 7 – Principais softwares comerciais de DES.

Software	Empresa	Descrição	Atualização
FlexSim	FlexSim Software Products Inc.	Um software de DES com uma interface intuitiva para modelar simulações em 3D	6 de setembro de 2016
GoldSim	GoldSim Technology Group LLC	Combina dinâmica de sistemas com aspectos de DES, incorporados em um framework de Monte Carlo	21 de setembro de 2015
Plant Simulation	Siemens PLM Software	Software que permite a simulação e otimização de sistemas e processos de produção	19 de maio de 2016
ProModel	ProModel Inc.	Uma ferramenta de DES que também permite a modelagem de processos contínuos	1 de junho de 2014
Simcad Pro	CreatASoft Inc.	Software de DES e contínua dinâmica. Interface visual sem ambiente de codificação	11 de agosto de 2016
SimEvents	MathWorks	Adiciona DES ao ambiente MATLAB / Simulink	14 de setembro de 2016
Simio	Simio LLC.	Orientado a objetos e um software de simulação baseado em agentes	1 de junho de 2016
SIMUL8	SIMUL8 Corporation	Software de simulação baseado em objetos	22 de março de 2016

Um trabalho clássico que aborda os três tipos de simulações é apresentado por Borshchev e Filippov (2004), no qual é apresentada a Figura 22, que mostra o nível de abstração em relação aos tipos de simulação e suas características principais.

Figura 22 – Abordagens em modelagem e simulação na escala de nível de abstração.



é válido lembrar que, apesar das abordagens serem diferentes, elas não são excludentes na implementação, ou seja, podem ser utilizadas de maneira integrada.

Por fim, é possível concluir que, com a atual disponibilidade de recursos tecnológicos em combinação às exigências cada vez mais restritas dos clientes, sistemas com comportamentos complexos tornam-se comuns em sistemas produtivos como um todo. Deste modo, o uso da simulação tem potencial para atuar na compreensão destes sistemas frente às novas expectativas da sociedade, fortemente influenciada pelo paradigma da Indústria 4.0.

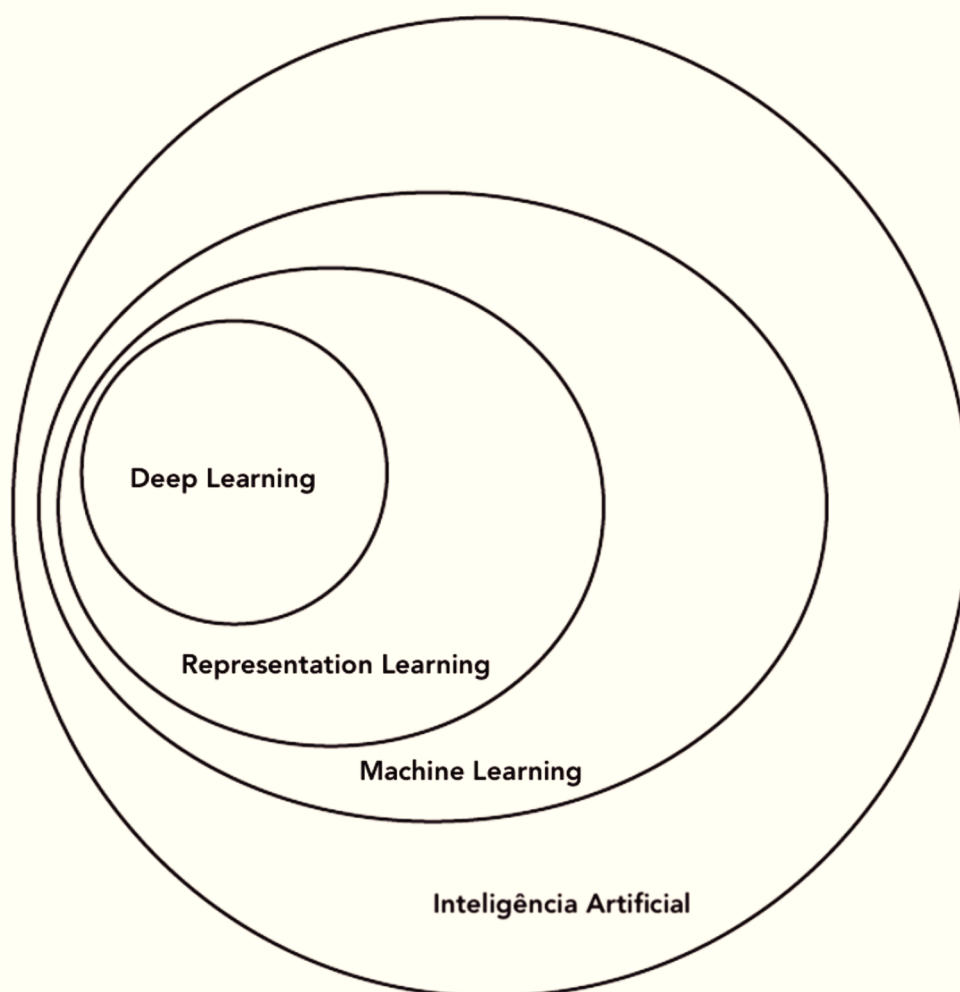
2.6 MACHINE LEARNING

Machine Learning (ML) é o campo de estudo que se preocupa em como construir programas de computador que automaticamente melhoram com a experiência (MITCHELL *et al.*, 1997). Ademais, o ML é uma disciplina focada na interrelação de duas perguntas: i.) Como construir sistemas computacionais que automaticamente melhoram por meio da experiência? ii.) Quais são as leis estatísticas, computacionais, informacionais e teóricas que governam todos estes sistemas de aprendizado incluindo computadores, humanos e organizações? (JORDAN; MITCHELL, 2015). Inserida na inteligência artificial (IA), o ML tem progredido substancialmente nas últimas duas décadas, abrangendo desde a curiosidade laboratorial à tecnologia prática de uso comercial e emergiu como um método para desenvolvimento de softwares de visão computacional, reconhecimento de voz, processamento de linguagem natural, controle de robôs e outras aplicações (JORDAN; MITCHELL, 2015). Analogamente,

Ribeiro, Grolinger e Capretz (2015) afirmam que o ML faz parte da IA e é uma coleção de técnicas estatísticas para construção de modelos matemáticos que podem fazer inferências em amostragem de dados.

Goodfellow, Bengio e Courville (2016) apresentam um diagrama de Venn, representado na Figura 23, no qual são correlacionados os temas Inteligência Artificial (IA), ML, *Representation Learning* e *Deep Learning* (DL).

Figura 23 – Diagrama de Venn relacionando sistemas de aprendizado relacionados à IA.

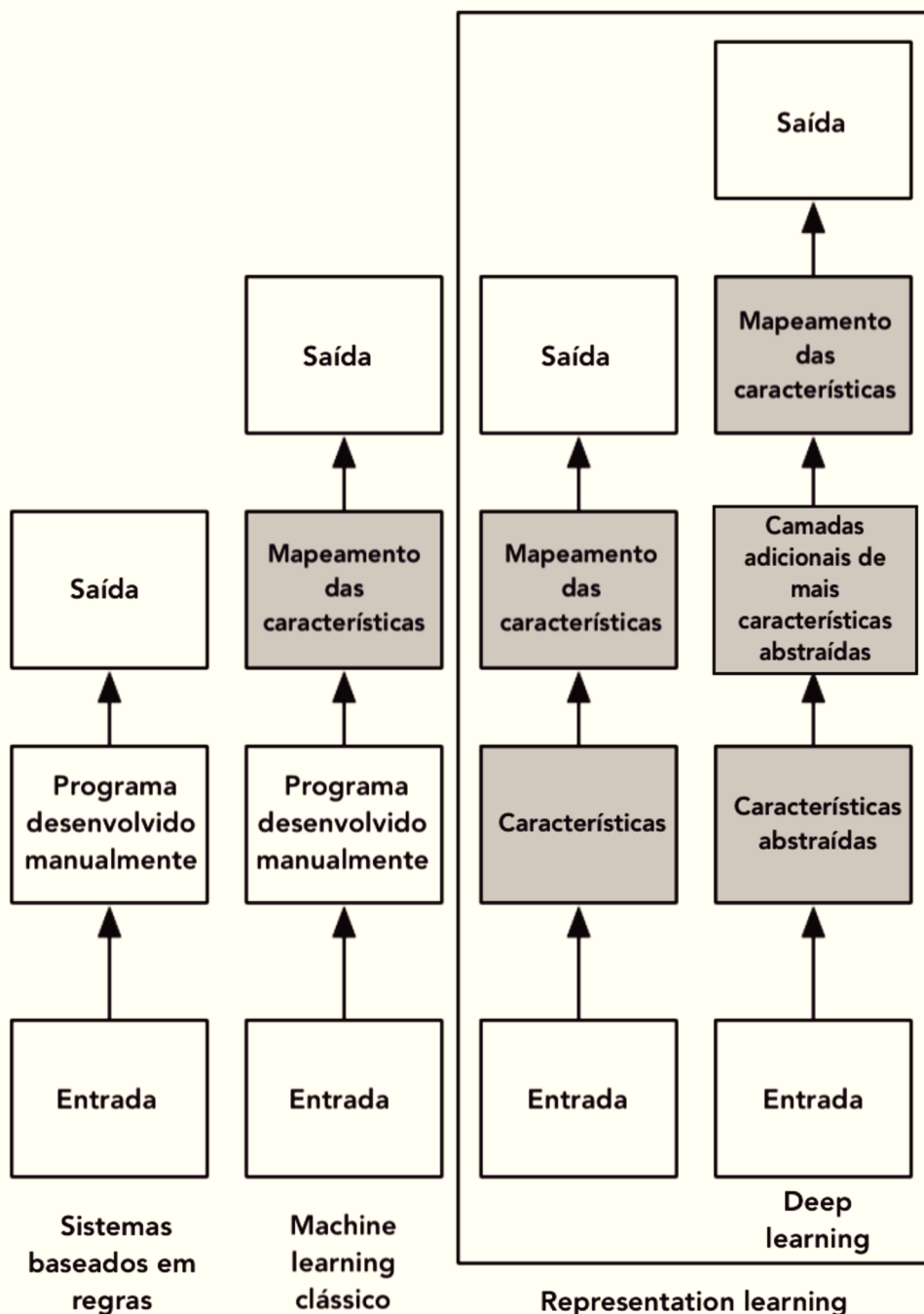


Fonte: Adaptado de Goodfellow, Bengio e Courville (2016)

Além disso, os autores complementam o entendimento dos diferentes termos com um fluxograma, representado pela Figura 24, que exemplifica as diferentes partes de cada sistema de aprendizado e também os sistemas baseados em regras. As caixas representadas em cinza são as etapas onde ocorrem aprendizados baseados em dados (GOODFELLOW; BENGIO; COURVILLE, 2016).

Segundo S. J. Russell e Norvig (2016), o primeiro trabalho reconhecido como IA foi realizado por McCulloch e Pitts (1943). Eles se basearam em três fontes: i.) Conhecimento da fisiologia básica e função dos neurônios no cérebro, ii.) Uma aná-

Figura 24 – Fluxograma mostrando como as diferentes partes de um sistema de IA se relacionam entre si nas diferentes disciplinas de IA.



Fonte: Adaptado de Goodfellow, Bengio e Courville (2016)

lise formal da lógica proposicional de Whitehead e B. Russell (1912) e iii.) A teoria da computação de Turing (1937). Eles propuseram um modelo de neurônios artificiais em que cada neurônio é caracterizado como estando ligado ou desligado, com uma mudança para “ON” ocorrendo em resposta à estimulação por um número suficiente de neurônios vizinhos. Eles mostraram, por exemplo, que qualquer função computável poderia ser computada por alguma rede de neurônios conectados, e que todos os co-

nectivos lógicos (e, ou, não, etc.) poderiam ser implementados por estruturas de rede simples. Dois estudantes de graduação em Harvard, Marvin Minsky e Dean Edmonds, construíram o primeiro computador de rede neural em 1950. O SNARC, como era chamado, usou 3.000 tubos de vácuo e um mecanismo piloto automático excedente de um bombardeiro B-24 para simular uma rede de 40 neurônios (S. J. RUSSELL; NORVIG, 2016).

De acordo com Marr (2018), o sistema de aprendizado especialista XCON da empresa *Digital Equipment Corporation* foi implantado em 1980 e, em 1986, foi creditado como gerador de economia anual para a empresa de US\$ 40 milhões. Isso é significativo porque até esse ponto os sistemas de IA eram geralmente considerados feitos tecnológicos impressionantes com utilidade limitada ao mundo real. No geral, a indústria da IA cresceu de alguns milhões de dólares em 1980 para bilhões de dólares em 1988, incluindo centenas de empresas que criam sistemas especialistas, sistemas de visão, robôs e software e hardware especializados para esses fins (S. J. RUSSELL; NORVIG, 2016). Logo depois disso veio um período chamado “*AI Winter*”, no qual muitas empresas caíram no esquecimento quando não cumpriram suas promessas extravagantes.

Nos últimos anos, houve uma revolução tanto no conteúdo quanto na metodologia do trabalho na inteligência artificial. Agora, é mais comum basear-se em teorias existentes do que propor novas, baseando-se em teoremas consolidados, ao invés de utilizar a intuição, além de ser necessário mostrar a relevância para as aplicações do mundo real em vez de exemplos abstraídos (S. J. RUSSELL; NORVIG, 2016). Os autores ainda comentam que a IA foi fundada em parte como uma rebelião contra as limitações dos campos existentes, como a teoria de controle e a estatística, mas agora está adotando essas áreas de estudo.

Por fim, David McAllester, professor emérito de ciência da computação da *North Carolina State University*, conclui: “no período inicial da IA, parecia plausível que novas formas de computação simbólica, por exemplo, quadros e redes semânticas, tornassem obsoleta a teoria clássica. Isso levou a uma forma de isolacionismo em que a IA se separou em grande parte do resto da ciência da computação. Este isolacionismo está atualmente sendo abandonado. Há um reconhecimento de que o ML não deve ser isolado da teoria da informação, que o raciocínio incerto não deve ser isolado da modelagem estocástica, que a busca não deve ser isolada da otimização/controlado clássico e que o raciocínio automatizado não deve ser isolado dos métodos formais e análise estática” (S. J. RUSSELL; NORVIG, 2016).

Ao longo dos 60 anos de história da ciência da computação, a ênfase tem sido no algoritmo como o principal assunto de estudo, mas alguns trabalhos recentes em IA sugerem que, para muitos problemas, faz mais sentido se preocupar com os dados e ser menos exigente quanto ao algoritmo a ser aplicado (S. J. RUSSELL; NORVIG,

2016). Esse tipo de trabalho sugere que o gargalo de conhecimento na IA – o problema de como expressar todo o conhecimento de que um sistema necessita – pode ser resolvido em muitas aplicações por métodos de aprendizado, ao invés de engenharia de conhecimento codificada manualmente, desde que os algoritmos de aprendizado tenham dados suficientes para o processo de aprendizagem (HALEVY; NORVIG; PEREIRA, 2009).

A crença de que por trás de um grande e complexo volume de dados podem existir modelos relativamente simples (ALPAYDIN, 2014), que explicam o seu comportamento tem alimentado o interesse por esta área de estudo. Outro motivo pelo crescente interesse em ML é o fato do avanço do processamento computacional ter possibilitado a escala das aplicações ao nível comercial. Sistemas que utilizam ML são usados para, por exemplo, identificar objetos em imagens, transcrever fala em texto, combinar produtos, notícias e publicações de acordo com o interesse do usuário e selecionar resultados de pesquisas relevantes (LECUN; BENGIO; HINTON, 2015).

De maneira geral, as técnicas de ML podem ser divididas em: aprendizado não supervisionado, aprendizado por reforço e aprendizado supervisionado. Visto que nesta dissertação de mestrado será implementado o aprendizado supervisionado, este será explicado mais detalhadamente.

2.6.1 Aprendizado não supervisionado

Na técnica de ML não supervisionada, as observações não possuem rótulos. Então, os algoritmos são utilizados para identificar padrões ocultos nas variáveis de entrada (FLATH; STEIN, 2018; RIBEIRO; GROLINGER; CAPRETZ, 2015). Para Wuest *et al.* (2016), se as instâncias não estiverem rotuladas, pode-se considerar como aprendizado não supervisionado e o principal objetivo destes algoritmos é a descoberta das classes dos itens pela técnica de *clustering*.

Analogamente, Ribeiro, Grolinger e Capretz (2015) afirmam que o aprendizado não supervisionado é conhecido como estimativa de densidade em estatística e é representado principalmente por algoritmos de *clustering* – técnica de mineração de dados para fazer agrupamento de dados de maneira automática. O principal atributo do aprendizado não supervisionado é a ausência de feedback do agente externo, pois algoritmo por si só identifica os possíveis agrupamentos na base de dados (WUEST *et al.*, 2016).

2.6.2 Aprendizado por reforço

O aprendizado por reforço é aprender o que fazer para que seja maximizada uma recompensa numérica, logo, ao aprendiz não é dito que ações tomar, e sim, a ele cabe descobrir quais ações tomar para que a maior recompensa seja alcançada (SUTTON; BARTO, 2018). Analogamente, Wuest *et al.* (2016) afirmam que o aprendizado

por reforço é definido pelo fornecimento de informação de treinamento pelo ambiente. Neste caso, o usuário deve descobrir quais ações implicam em melhores resultados por tentativa e erro ao invés de receber recomendações.

O aprendizado por reforço é visto por alguns pesquisadores como um caso especial de aprendizado supervisionado, mas diferem quanto à aplicação. Os problemas que utilizam aprendizado por reforço são descritos pela falta de exemplos rotulados como comportamentos bons e ruins, e é mais adequado em situações onde não há um supervisor experiente (WUEST *et al.*, 2016). Para Ribeiro, Grolinger e Capretz (2015), o aprendizado por reforço é caracterizado quando os resultados do aprendizado implicam em ações que mudam o ambiente.

2.6.3 Aprendizado supervisionado

O aprendizado supervisionado é o aprendizado a partir de um conjunto de dados de treinamento rotulado fornecido por um supervisor externo (SUTTON; BARTO, 2018). Analogamente, a técnica de ML supervisionada é a tarefa de inferência de uma função de dados de treinamento rotulados, ou seja, cada exemplo é um par de objetos de entrada (na maioria dos casos um vetor) e um valor de saída (FLATH; STEIN, 2018; RIBEIRO; GROLINGER; CAPRETZ, 2015).

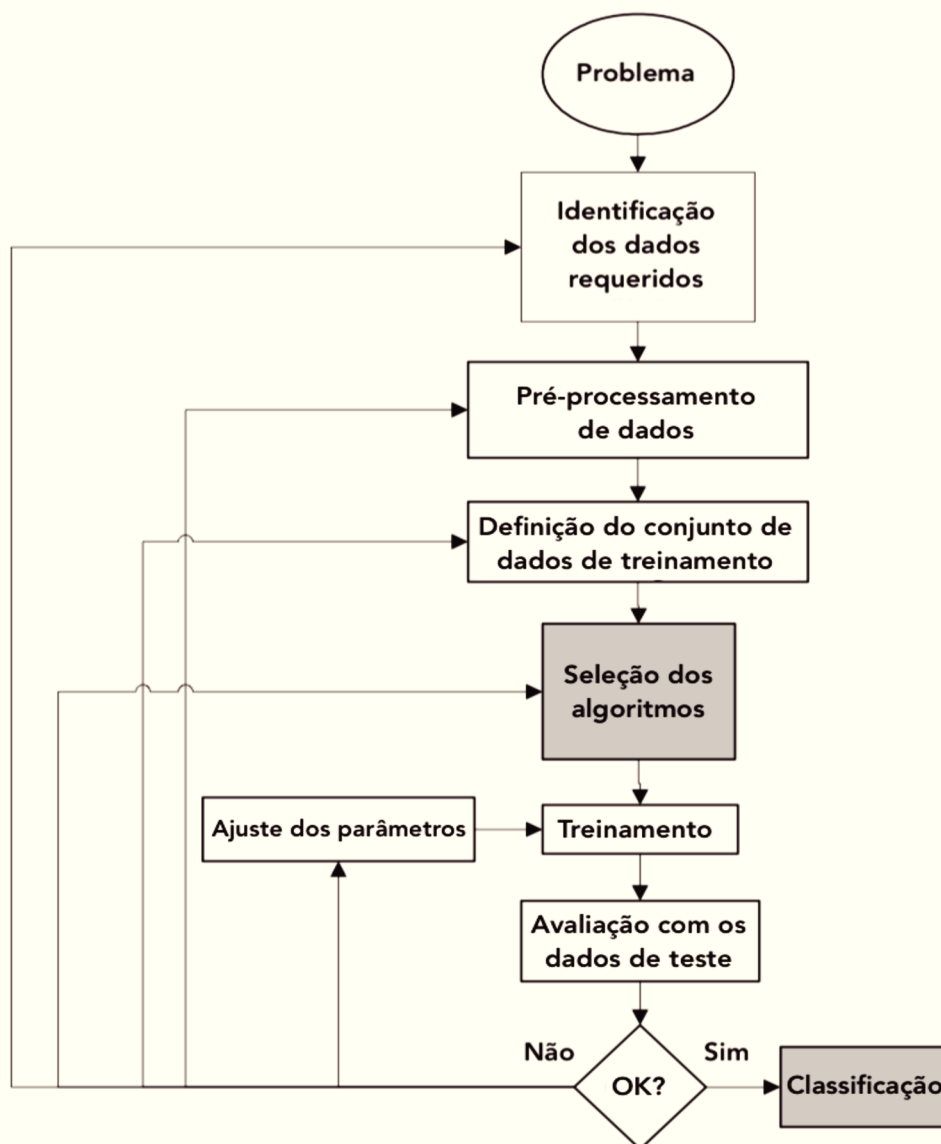
Wuest *et al.* (2016) exemplifica que no aprendizado supervisionado, o algoritmo aprende com exemplos fornecidos por um supervisor experiente e isto é possível devido à disponibilidade de feedbacks especializados e dados rotulados. O processo para aplicação do aprendizado supervisionado para um problema real pode ser subdividido em seis etapas: i.) coleta dos dados, ii.) pré-processamento dos dados, iii.) definição do conjunto de dados de treinamento, iv.) seleção do algoritmo, v.) avaliação no conjunto nos dados de teste e ajuste dos parâmetros vi.) resultados. A Figura 25 proposta por Sotiris B Kotsiantis, Zaharakis e Pintelas (2007) ilustra o processo de aprendizado supervisionado.

A etapa de coleta de dados diz respeito à coleta do conjunto de dados com potencial para descrever o problema que se está tentando resolver.

A etapa de pré-processamento têm impacto significativo na performance do algoritmo de ML e consiste na organização destes dados por meio de sub-etapas como limpeza, normalização, transformação, seleção e extração de características, etc. (S. KOTSIANTIS; KANELLOPOULOS; PE PINTELAS, 2006). Durante a etapa de pré-processamento é quando são eliminados conjunto de dados que não fazem sentido, e.g., um valor negativo para o número de operadores disponíveis em uma estação de trabalho, ou também para aquisições incompletas de dados, e.g., falta de dados devido à falha de um dispositivo de medição.

A etapa de definição do conjunto de dados de treinamento refere-se à utilização de parte da base de dados disponível para o treinamento dos algoritmos de ML.

Figura 25 – Processo de aprendizado supervisionado.



Fonte: Adaptado de Sotiris B Kotsiantis, Zaharakis e Pintelas (2007)

Embora não exista uma regra geral, Sotiris B Kotsiantis, Zaharakis e Pintelas (2007) afirmam que pode-se dividir o conjunto de treinamento como dois terços da base de dados completa e um terço para teste do modelo.

Na etapa de seleção do algoritmo é contabilizada a performance de cada algoritmo por meio da acurácia dos modelos. Neste sentido, adicionalmente à etapa anterior, é possível utilizar a técnica de validação cruzada, que consiste na divisão do conjunto de dados em subconjuntos mutuamente exclusivos para então avaliar a capacidade de generalização do modelo por meio da avaliação da acurácia dos testes realizados nesses subconjuntos (KOHAVI *et al.*, 1995).

Por fim, nas próximas etapas são realizados ajustes dos parâmetros do modelo de ML com o objetivo de maximizar a sua acurácia para que na última etapa, este modelo com os parâmetros ajustados seja utilizado para obter os resultados.

O aprendizado supervisionado possui diferentes aplicações na manufatura sendo o monitoramento e controle as principais que podem ser destacadas (WUEST *et al.*, 2016). De acordo com Wuest *et al.* (2016), o aprendizado supervisionado é a técnica mais adequada para resolução dos problemas enfrentados na manufatura, pois os descreve com grandes detalhes visto que os dados geralmente são rotulados e feedbacks especializados estão disponíveis. Os rótulos no aprendizado supervisionado podem ser do tipo discreto ou contínuo e são tratados por algoritmos de classificação ou regressão, respectivamente (RIBEIRO; GROLINGER; CAPRETZ, 2015). A classificação é utilizada para predição, reconhecimento de padrões e detecção de valores anômalos enquanto que a regressão é usada para predição e ranqueamento.

Dois algoritmos de ML supervisionado são utilizados neste trabalho: i.) K-Nearest Neighbors (KNN) e ii.) Logistic Regression (LR). Apesar de existirem outros algoritmos de ML supervisionado, e.g., *linear regression, support vector machine (SVM), decision tree, rede neural* e etc., os dois algoritmos foram escolhidos pois são utilizados para predição de classes (classificação). Ademais, apesar de outros algoritmos (como o SVM) também serem utilizados para classificação, eles não foram experimentados devido à limitação de tempo e certamente podem ser objeto de estudo de pesquisas futuras.

Antes de apresentar os fundamentos de cada algoritmo é importante entender os conceitos de instâncias (*instances*), características (*features/attributes*) e rótulos (*labels/targets*). As instâncias, ou também amostras, são observações únicas das quais o modelo irá aprender, sendo cada instância um vetor de características e rótulos. As características são os atributos que descrevem cada instância, podendo ser valores contínuos ou discretos. Os rótulos são os resultados que desejam-se fazer predições podendo assumir valores contínuos ou discretos. A Tabela 8 ilustra um exemplo de uma base de dados com instâncias, características e rótulos.

Tabela 8 – Exemplo de uma base de dados com n instâncias, n características e um rótulo com valores binários.

Itens	Carac. 1	Carac. 2	...	Carac. n	Rótulo
Instância 1	i_{11}	i_{12}	...	i_{1n}	Sim
Instância 2	i_{21}	i_{22}	...	i_{2n}	Não
Instância 3	i_{31}	i_{32}	...	i_{3n}	Não
...
Instância n	i_{n1}	i_{n2}	...	i_{nn}	Sim

Fonte: Autoria própria (2019)

2.6.3.1 *K-Nearest Neighbors*

O algoritmo *k-Nearest Neighbors* (KNN) é o procedimento não-paramétrico, i.e., não assume conhecimento prévio das distribuições estatísticas, de decisão mais simples que atribui à instância não classificada o rótulo da instância mais próxima utilizando distâncias geométricas (COVER; HART, 1967).

Dados os vetores (x_1, x_2) , (y_1, y_2) pertencentes aos números reais, a distância geométrica entre cada par de valores estará contida dentro dos números reais. A distância entre dois pontos no espaço Euclidiano para duas dimensões é calculada de acordo com a Equação 1.

$$D_{(x,y)} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (1)$$

De maneira abstraída, o algoritmo KNN busca identificar o vizinho mais próximo a um par de dados desconhecidos que deseja-se predizer o rótulo, e é atribuída a esta instância o rótulo do vizinho mais próximo. Por exemplo, dado um conjunto de dados compostos por duas características e um rótulo (x_1, x_2, r_1) , por meio do cálculo da distância D , será encontrado um par de dados (y_1, y_2, r_2) em que a distância é minimizada. Logo, será atribuído a r_1 o valor de r_2 .

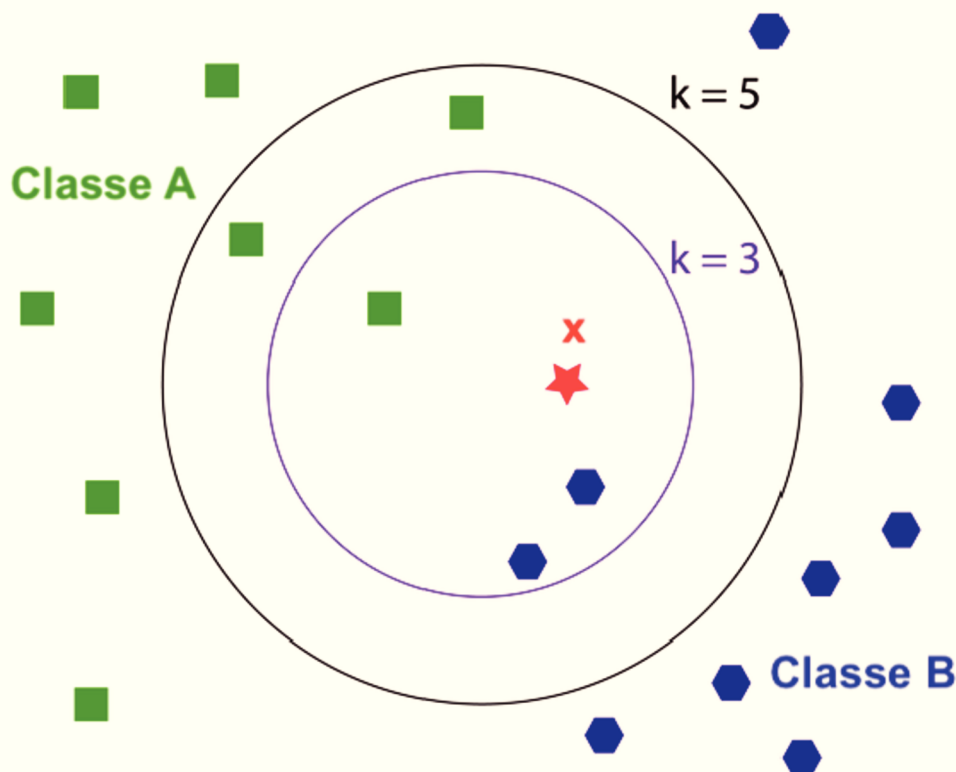
Seguindo este raciocínio, nota-se que, para um conjunto de dados grande, é necessária uma grande capacidade de processamento, pois para cada nova instância que se deseja fazer uma predição deverá-se iterar toda base de dados. Além disso, dependendo da aplicação, não faz sentido avaliar apenas um único vizinho mais próximo. Portanto, utiliza-se o parâmetro “ k ” para avaliar os k -vizinhos mais próximos daquela instância desconhecida e então atribui-se o rótulo dos k -vizinhos mais próximos ao rótulo da instância desconhecida. De maneira gráfica, a Figura 26 ilustra um exemplo prático da classificação de uma instância desconhecida “ x ” sendo possíveis os rótulos de Classe A ou Classe B.

Na Figura 26 observa-se que dependendo da escolha do parâmetro k , diferentes resultados podem ser obtidos, fazendo com que a escolha do parâmetro k seja extremamente relevante para a performance do classificador. Por exemplo, para $k = 1, 3$ e 5 o classificador inferiria que o rótulo de x seria Classe B, Classe B e Classe A, respectivamente. Nota-se também que k é geralmente escolhido como um número ímpar para evitar situações de empate entre dois ou mais rótulos.

2.6.3.2 *Logistic Regression*

Apesar do nome do algoritmo conter a palavra regressão, a regressão logística (LR) é um algoritmo de aprendizado utilizado para classificação ou predição da probabilidade de ocorrência de um evento por adaptação dos dados a uma função logística e pode ser utilizada para situações em que a variável dependente é de natureza bi-

Figura 26 – Ilustração da predição de uma instância x utilizando k-vizinhos.



Fonte: Adaptado de Cover e Hart (1967)

nária (BURKOV, 2019; YU; J. LIU, 2011). Além disso, a LR é um recurso que permite estimar a probabilidade associada à ocorrência de determinado evento em face de um conjunto de variáveis explanatórias, i.e., variáveis que afetam a resposta do sistema e podem ser definidas pelo experimentador. O algoritmo utiliza da função logística padrão ou função sigmoide, ilustrada pela Equação 2, para, por exemplo, classificar instâncias entre 0 e 1.

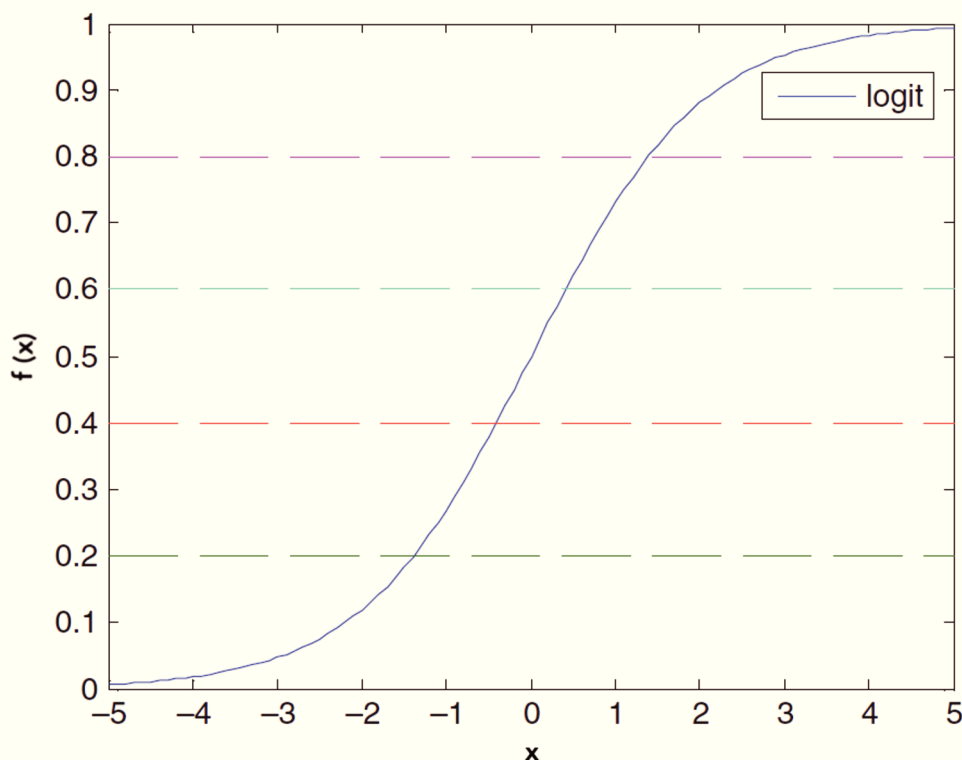
$$Prob(evento) = P(x) = \frac{1}{1 + e^{-g(x)}} \quad (2)$$

Onde a Prob (evento) é a função que descreve a probabilidade de um evento estar sob controle (0) e fora de controle (1) (YU; J. LIU, 2011). Além disso, os autores complementam com um exemplo representado pela Figura 27, a qual indica que há uma relativa baixa probabilidade do evento estar fora de controle até que um limite seja atingido ($x = 0$), limite no qual a probabilidade do sistema estar fora de controle aumenta drasticamente.

A variável $g(x)$ é usualmente definida como (Equação 3):

$$g(x) = \frac{\log P(x)}{1 - P(x)} \quad (3)$$

Figura 27 – Função logística representando a probabilidade de um sistema estar sob ou fora de controle.



Fonte: Adaptado de Yu e J. Liu (2011)

Onde $g(x)$ é uma combinação linear de variáveis independentes $x_i (i=1,2,\dots,k)$ (Equação 4):

$$g(x) = \alpha + \beta_1 * x_1 + \beta_2 * x_2 + \dots + \beta_k * x_k \quad (4)$$

Onde $\beta_i (i=1,2,\dots,k)$ são chamados os coeficientes de regressão de x_i .

A pré condição para calcular $P(x)$ é a determinação dos parâmetros α e β previamente, e a estimativa para α e β são obtidos utilizando o método da máxima verossimilhança (YU; J. LIU, 2011). Estes estimadores são utilizados para prever a probabilidade $P(x)$ da ocorrência de um evento de acordo com a Equação 2.

2.6.4 Ciência de Dados e Manufatura

A utilização de uma grande quantidade de dados nas empresas de todos os setores faz parte de um escopo maior de transformação digital, sendo a disseminação e facilidade do uso do ML, um dos grandes responsáveis para esta eminente transformação. Inserido no escopo da transformação digital, uma nova área vem se consolidando como suporte para a inovação devido ao uso intensivo de ML: a ciência de dados.

A ciência de dados (*data science*) é o estudo da extração generalizada de conhecimento de dados e implica o envolvimento de dados em suas diversas formas

– e.g. texto, imagens, vídeos –, estatística e um sistemático estudo da organização (DHAR, 2013). O autor ainda afirma que, no domínio das tomadas de decisão, um requisito epistêmico comum ao avaliar se um novo conhecimento é acionável é o seu poder preditivo, e não apenas sua capacidade de explicar o passado. E o campo da ciência de dados possui esta capacidade. Não coincidentemente o termo “ciência” se faz presente na denominação desta área de estudo. O conhecimento gerado por meio dos dados acontece por meio de estudos sistemáticos da empresa, que constrói e organiza o conhecimento na forma de justificativas e previsões testáveis (DHAR, 2013).

Ao cientista de dados, é requerida uma habilidade integrada em ferramentas de programação, matemática, IA, estatística, base de dados e otimização (GRÖGER, 2018; DHAR, 2013). Os dois autores divergem quanto à necessidade de conhecimento profundo para formulação de problemas. Enquanto Gröger (2018) afirma que os cientistas de dados necessitam apenas de conhecimentos básicos do negócio, Dhar (2013) afirma que é necessário um profundo conhecimento dos problemas e processos e para isto ocorrer, invariavelmente deverá ter conhecimentos profundos do negócio como um todo.

Ambos estão corretos, a diferença está na maturidade do mercado e da empresa. Portanto, em um contexto de transformacional, Gröger é mais realista ao assumir que o cientista de dados não conseguirá ter todas as qualificações, pois os profissionais estão se adaptando a esta mudança de era no campo da manufatura.

Gröger (2018) utiliza o termo *citizen data scientist* (cientista de dados cidadão ou cientista de dados leigo) para definir o profissional que consegue fazer a conexão entre o mundo dos negócios e o mundo da ciência de dados. A Figura 28 ilustra as características utilizando a abordagem de *citizen data scientists*.

Moore (2017) ainda complementa que *citizen data scientists* são capazes de realizar análises sofisticadas que antes exigiam mais conhecimento, permitindo que eles fornecessem análises avançadas sem ter as habilidades que caracterizam os cientistas de dados. Eles podem atuar na automação de tarefas que são repetitivas, intensivas manualmente e que não exigem conhecimento profundo em ciência de dados, como por exemplo na etapa de pré-processamento de dados.

Partindo da premissa que a transformação digital é dependente da disponibilidade de dados, é natural inferir que os resultados das implementações também são dependentes da qualidade destes dados – algo que nem sempre é observado nas empresas de manufatura.

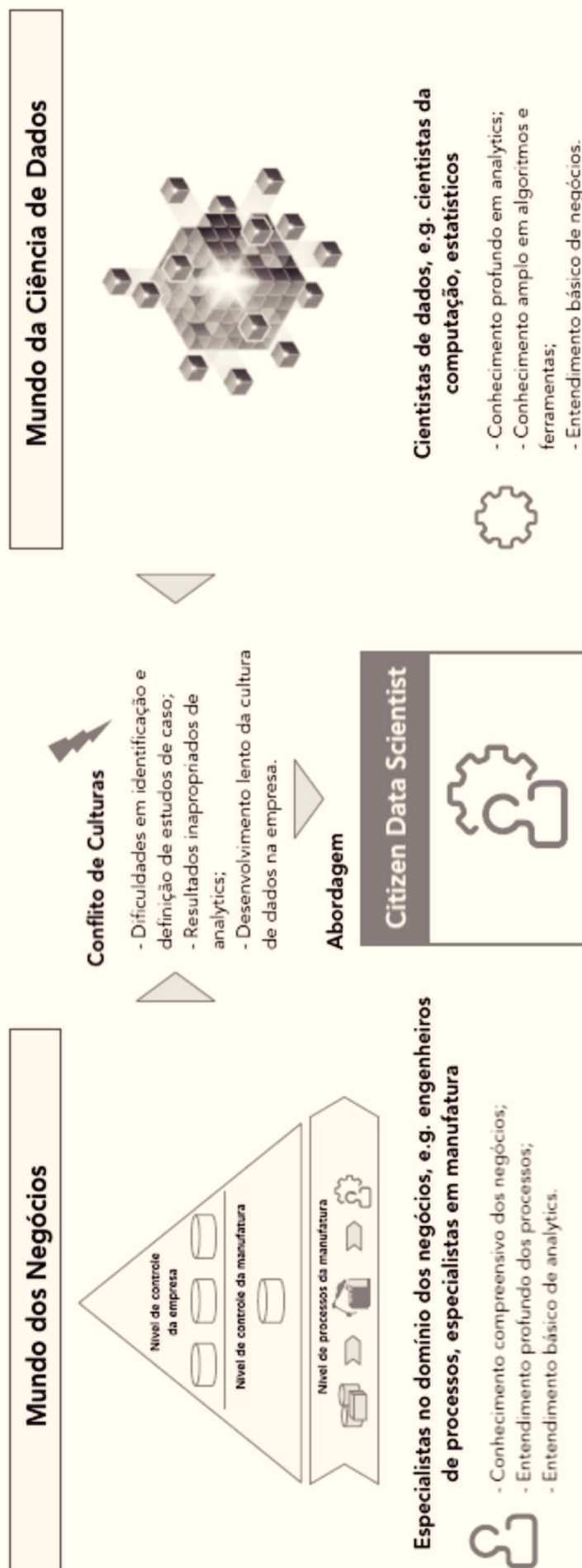
Em situações onde a concepção de uma base de dados é complexa, uma prática que vêm sendo utilizada em visão computacional é a utilização de dados sintéticos para treinamento de algoritmos de ML. Por exemplo, objetos são modelados em softwares CAD e suas imagens renderizadas são utilizadas para o treinamento de algoritmos e objetos são reconhecidos em vídeos onde há a presença de apenas

objetos reais.

Seguindo este mesmo raciocínio, dados sintéticos podem ser gerados por meio de simulações e serviços digitais podem ser experimentados seja para validação de hipóteses quanto para estudos exploratórios de treinamento de algoritmos em cenários fictícios.

No próximo capítulo, uma simulação que performa a escolha inteligente de fornecedores com utilização de dois algoritmos de ML (KNN e LR) é apresentado, na qual o objetivo é a maximização do nível de serviço, inclusive, diante de um cenário de interrupção da cadeia de suprimentos.

Figura 28 – A abordagem de citizen data scientists.



Fonte: Adaptado de Gröger (2018)

3 EXPERIMENTAÇÃO DIGITAL

Nesta seção é apresentado o modelo de simulação e o modelo de ML que foram desenvolvidos no contexto da seleção de fornecedores. Posteriormente é apresentado como estes dois modelos são integrados com o intuito de mensurar o desempenho de um serviço digital de seleção de fornecedores aplicado à manufatura.

3.1 MODELO DE SIMULAÇÃO

O objetivo do modelo de simulação é a geração de dados sintéticos que simulam um sistema produtivo do tipo *make-to-order* que possui até quatro fornecedores disponíveis. O modelo de simulação é realizado com o software Anylogic.

O modelo considera que os pedidos aos fornecedores de matéria-prima somente acontecem após a consolidação de um pedido do cliente e que a matéria-prima é o único suprimento necessário para manufaturar o produto final. Analogamente, os pedidos de compra são caracterizados por distribuições normais como apresentado em Bodaghi, Jolai e Rabbani (2018). A incerteza da demanda também é modelada com distribuições normais.

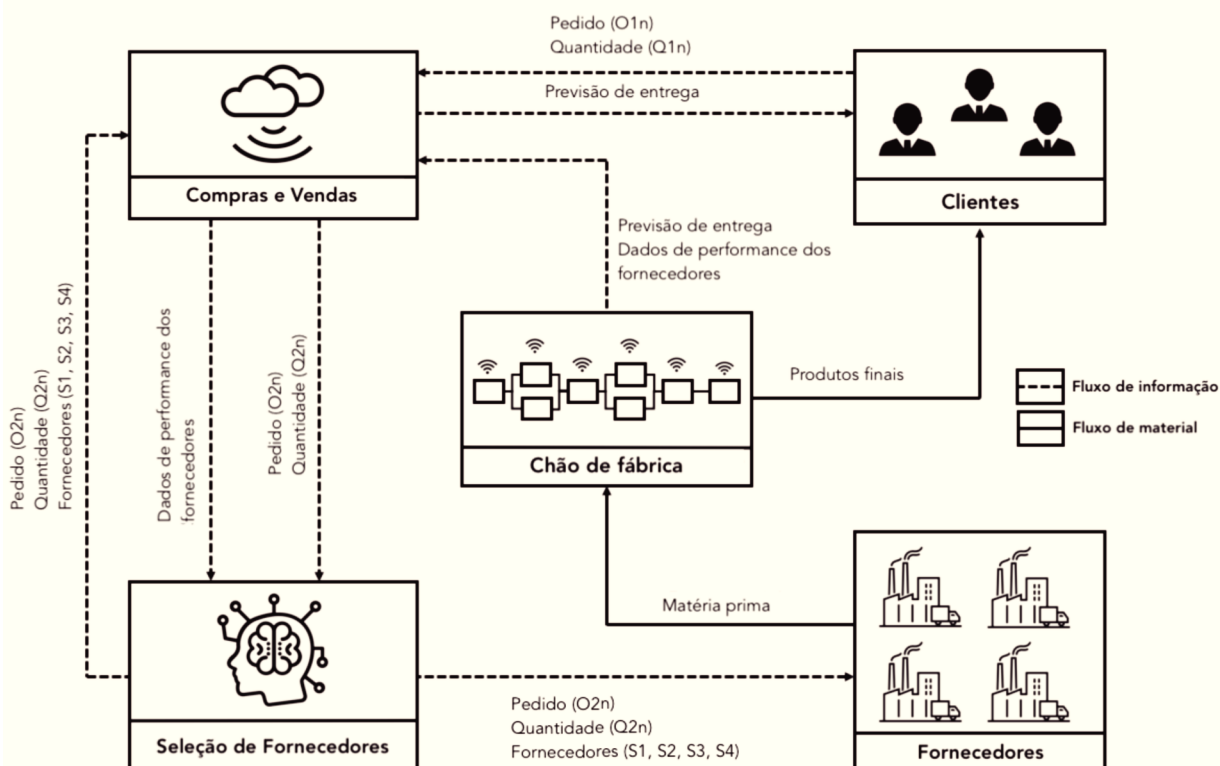
A performance dos fornecedores é modelada de maneira análoga ao trabalho apresentado por Tomlin (2006), que possui a premissa que um fornecedor pode ser instável em um determinado período de tempo e também pode apresentar limitações quanto à sua capacidade. Neste trabalho, quatro possíveis fornecedores são modelados e as restrições previamente mencionadas influenciam a performance de entrega no prazo dos fornecedores, que é modelada como uma distribuição normal.

Outra consideração do modelo de simulação é que há uma data de entrega prometida pelo fornecedor. Esta data é considerada como o acordo de negociação entre o cliente e fornecedor, e qualquer pedido entregue após esta data é considerado como atrasado. A Figura 29 ilustra o modelo de simulação com o fluxo de materiais e de informação.

3.2 MODELO DE ML

O objetivo do modelo de ML é realizar a seleção de fornecedores com as maiores chances de entregar um determinado pedido num tempo pré-acordado. A tomada de decisão considera, exclusivamente, dados passados de performance dos fornecedores.

O modelo de ML apresentado nesta dissertação é implementado utilizando linguagem Python e o pacote Scikit-Learn. Este pacote é definido por Pedregosa *et al.* (2011) como um módulo Python que integra uma ampla gama de algoritmos no estado da arte de ML para problemas de média complexidade do tipo supervisionado

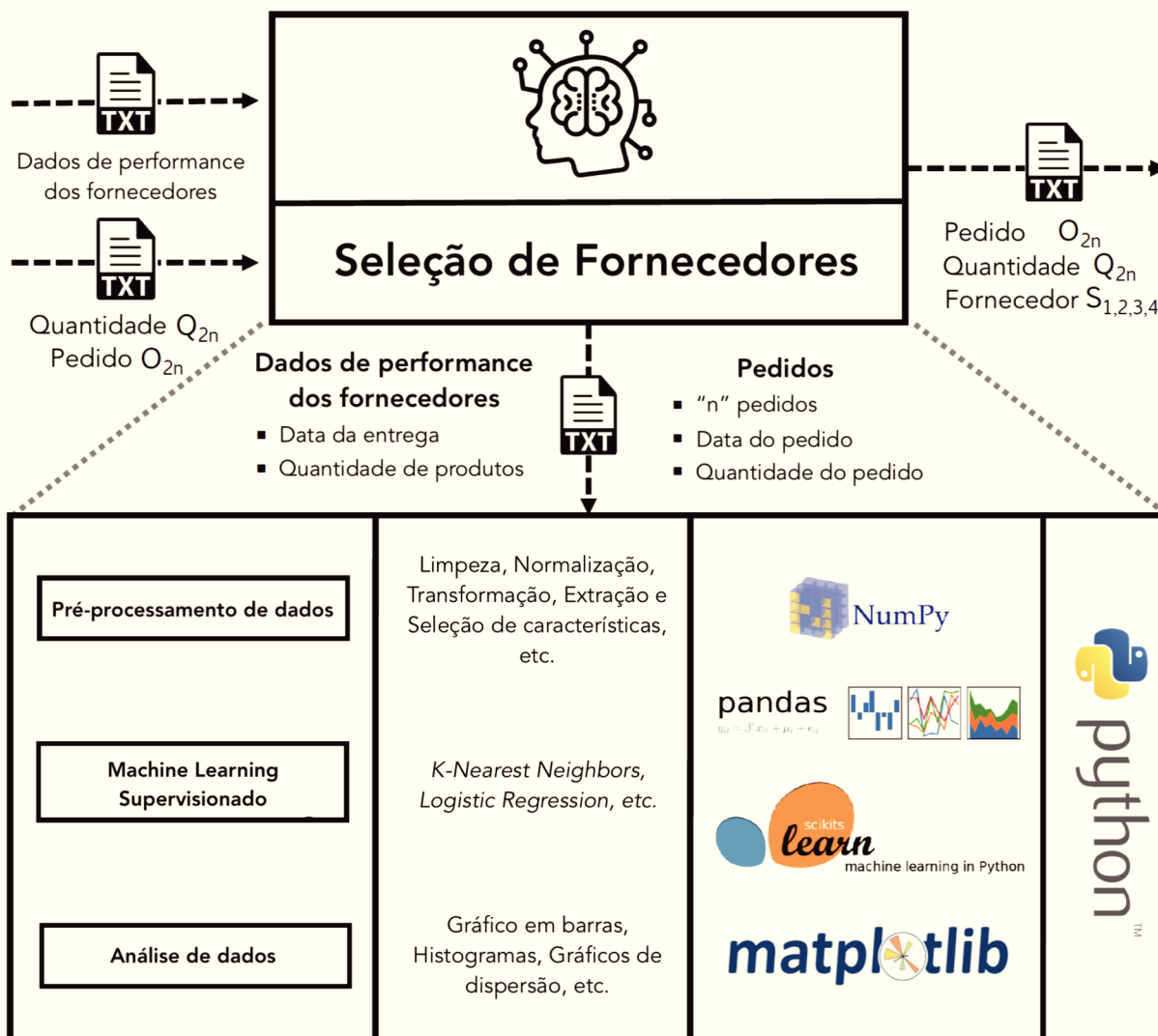
Figura 29 – Modelo de simulação *make-to-order*.

ou não-supervisionado. Além disso, outros pacotes como Numpy, Matplotlib e Pandas são utilizados para execução de tarefas como pré-processamento, análise e visualização de dados. A Figura 30 apresenta o modelo de seleção de fornecedores utilizando ML.

A etapa de pré-processamento frequentemente tem um impacto significativo na performance dos algoritmos de ML e pode incluir sub-etapas como limpeza de dados, normalização, transformação, extração e seleção de características, etc. (S. KOTSISANTIS; KANELLOPOULOS; PE PINTELAS, 2006). Neste trabalho, como os dados são oriundos de uma simulação, a base de dados pode ser considerada como de boa qualidade pois alguns problemas recorrentes em bases de dados reais como: ausência de observações, combinação de dados impossível (e.g. número negativo de produtos), valores nulos, etc. dificilmente acontecem, pois trata-se de um ambiente controlado por dados sintéticos. Portanto, a etapa de pré-processamento é mais simples tratando-se de modelos de simulação em comparação a bases de dados reais.

Dois algoritmos de ML, i.e., KNN e LR, são utilizados no experimento. O algoritmo KNN é o algoritmo mais comum de classificação em casos que não há conhecimento prévio da distribuição estatística dos dados (N. LI *et al.*, 2016). A LR é baseada em aprendizado supervisionado e se organiza de acordo com a natureza das observações, sendo necessária pouca informação sobre as características dos dados de entrada (YU; J. LIU, 2011).

Figura 30 – Modelo de seleção de fornecedores utilizando ML.



Fonte: Autoria própria (2019)

Apesar de ter sido definido que no modelo de simulação serão utilizadas distribuições normais com o intuito de melhor caracterizar as incertezas do sistema, em casos reais é altamente provável que os processos se comportem de maneira diferente de distribuições normais bem definidas. Logo, o modelo de ML não tira vantagem de qualquer conhecimento prévio acerca do comportamento do sistema com o objetivo de melhor representar um caso real.

Neste trabalho, os dados passados são rotulados como i.) Entregas no prazo e ii.) Entregas atrasadas. O algoritmo KNN é aplicado separadamente em cada uma das bases de dados e mapeia a performance dos fornecedores de acordo com as características previamente mencionadas do modelo: período e quantidade de itens pedidos. Na LR, as duas bases de dados são utilizadas como entrada e o resultado esperado para cada pedido do cliente é a probabilidade de cada fornecedor entregar o pedido de acordo com o prazo pré-acordado.

Além disso, uma combinação destas duas técnicas de ML é apresentada. A primeira, Híbrida A, confronta os resultados de ambos algoritmos de classificação sem considerar a acurácia de cada técnica. A segunda, Híbrida B, utiliza a mesma abordagem mas considerando a acurácia de cada técnica.

A acurácia do modelo KNN é representada pela taxa de acerto de predição. Por exemplo, caso a predição do modelo em uma determinada instância seja de “entrega atrasada”, e de fato houver um atraso na entrega, é considerado que o modelo acertou a predição. A acurácia é representada por acc_a e acc_b referem-se as entregas no prazo e atrasadas, respectivamente. Ademais, R_{ka} e R_{kb} referem-se aos resultados do classificador KNN, e.g., Fornecedor 1 selecionado, utilizando as bases de dados com rótulos de entregas no prazo e atrasadas, respectivamente.

Para a acurácia do modelo de LR, é considerada a área abaixo da curva característica de operação do receptor (ROC). A área abaixo da curva ROC (AUC) pode variar entre 0 e 1, sendo o valor de 0,5 considerado como uma performance randômica de predição. O trabalho apresentado por Fawcett (2006) é recomendado para uma explicação detalhada da curva ROC. Ademais, R_{lr1} e R_{lr2} representam os resultados para o primeiro e o segundo fornecedor mais provável de realizar uma entrega no prazo de acordo com o classificador LR, respectivamente. Ambos pseudo-códigos são apresentados a seguir por meio dos Algoritmos 1 e 2.

Algoritmo 1 Híbrida A

```

1: procedimento SELEÇÃO( $\tau, \gamma$ )                                     ▷ data and quantidade
2:   se  $R_{lr1} = R_{kb}$  então
3:     retornar  $R_{ka}$ 
4:   senão
5:     retornar  $R_{lr1}$ 
6:   fim se
7: fim procedimento

```

3.3 INTEGRAÇÃO DOS MODELOS

A integração dos modelos de simulação e ML acontece por meio da troca de informação entre os dois modelos com arquivos em formato texto, isto é, como cada modelo é capaz de ler e escrever arquivos texto, a comunicação se dá a partir da escrita de um arquivo de um modelo e leitura deste arquivo pelo outro modelo. O formato texto foi escolhido, pois o software de simulação Anylogic utiliza a linguagem de programação Java, enquanto que o modelo de ML utiliza a linguagem de programação Python. Como havia a necessidade de comunicá-los entre si, a maneira encontrada foi por meio de arquivos texto que ambos conseguem interpretar.

A sequência de atividades que envolvem esta integração entre os dois modelos é resumida em três passos, e é exemplificada por meio da Figura 31.

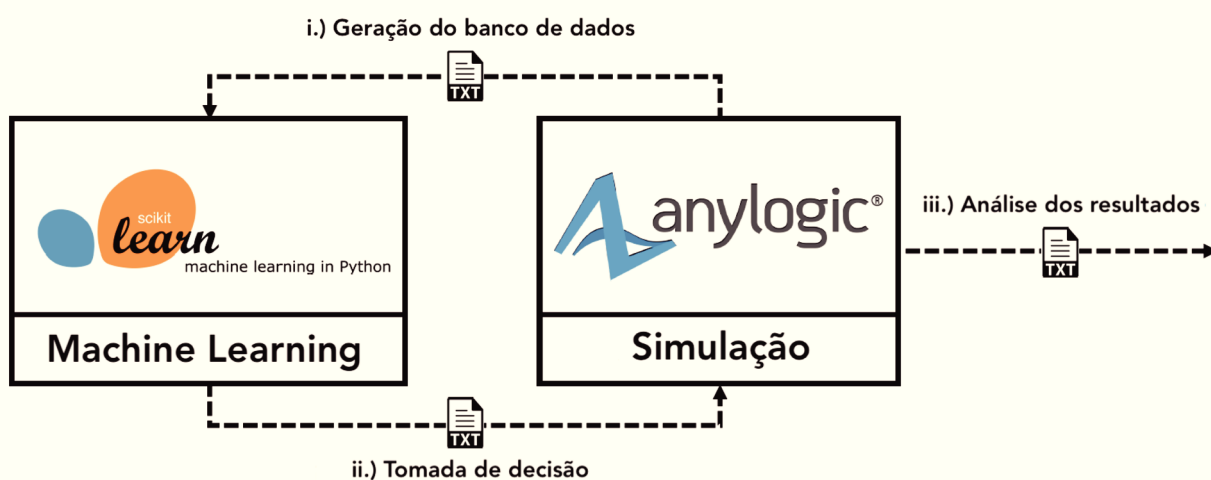
Algoritmo 2 Híbrida B

```

procedimento SELEÇÃO( $\tau, \gamma$ ) ▷ data and quantidade
2:   se  $AUC \geq acc_a$  então
      se  $AUC \geq acc_b$  então
4:     retornar  $R_{lr1}$ 
      senão
6:       se  $R_{kb} = R_{lr1}$  então
          retornar  $R_{lr2}$ 
7:       senão
8:         retornar  $R_{lr1}$ 
9:       fim se
10:    fim se
11:  fim se
12:  senão
      se  $AUC \geq acc_b$  então
14:    se  $AUC \geq acc_b$  então
        se  $R_{ka} = R_{kb}$  então
16:          se  $R_{ka} = R_{lr1}$  então
              retornar  $R_{lr2}$ 
17:          senão
18:            retornar  $R_{lr1}$ 
19:          fim se
20:        fim se
21:      senão
22:        retornar  $R_{kb}$ 
23:      fim se
24:    senão
25:      retornar  $R_{kb}$ 
26:    fim se
27:  fim se
28: fim procedimento

```

Figura 31 – Integração entre os modelos de simulação e ML.



Fonte: Autoria própria (2019)

O primeiro passo consiste na i.) Geração da base de dados por meio do modelo de simulação. Esta base de dados é utilizada como dado de entrada no modelo de ML e então, no passo dois, ii.) Os resultados da tomada de decisão orientada a dados são compilados em um arquivo de saída. Posteriormente, este arquivo de saída se torna o arquivo de entrada para o experimento teste de simulação. Finalmente, no passo três, iii.) Os resultados da simulação teste são compilados e analisados.

3.4 CONSIDERAÇÕES DO MODELO

Com o intuito de avaliar o desempenho da abordagem desenvolvida, um experimento foi conduzido em uma simulação com uma janela de quatro anos. Nesta janela de experimentação, 50% do período é utilizado para treinamento dos algoritmos, 25% do tempo para aperfeiçoamento e validação do modelo e os últimos 25% para teste do modelo.

No primeiro cenário, é assumida a total disponibilidade dos quatro fornecedores. No segundo cenário, dois fornecedores estão indisponíveis devido a uma interrupção no sistema, logo, sobram apenas dois fornecedores para a fase teste do modelo.

Os fornecedores possuem características pré-determinadas e suas performances são estocasticamente definidas com relação à entrega dos pedidos, significando que cada um possui restrições quanto à capacidade de produção e estão sujeitos a sazonalidades. Os fornecedores S_1 e S_3 possuem ambas restrições, isto é, dado um pedido de “p” produtos, dependendo da data a que este pedido se refere, há uma probabilidade maior ou menor do fornecedor entregar o pedido no prazo pré-acordado. Analogamente, os fornecedores S_2 e S_4 também possuem restrições, no entanto, somente com relação à quantidade de produtos dos pedidos, isto é, dependendo da quantidade de “p” produtos e independentemente da data.

A seguir, um exemplo numérico abstraído para auxiliar a compreensão do papel das características dos fornecedores no modelo de simulação. Defina-se “p” igual a 100 produtos e “data” como o período entre os meses de dezembro a março, para o fornecedor S_1 . Este, ao receber um pedido de 150 produtos em janeiro, terá uma performance diferente caso o mesmo pedido fosse realizado em junho. Analogamente, o pedido sendo realizado na mesma data (janeiro), porém de 50 unidades, também teria performance diferente (entrega dentro ou fora do prazo). Portanto, por meio desta modelagem, tenta-se retratar casos os quais a empresa está sujeita a sazonalidades, não possui capacidade de produção para atender um determinado pedido ou também quando há despriorização de um pedido por ser muito pequeno. Nestes exemplos apresentados, haveria maior probabilidade de entregar fora do prazo e o que define esta performance são distribuições estatísticas normais.

Por fim, a performance da seleção de fornecedores em ambos os cenários é avaliada comparando i.) a escolha randômica de fornecedores, o que significa que a

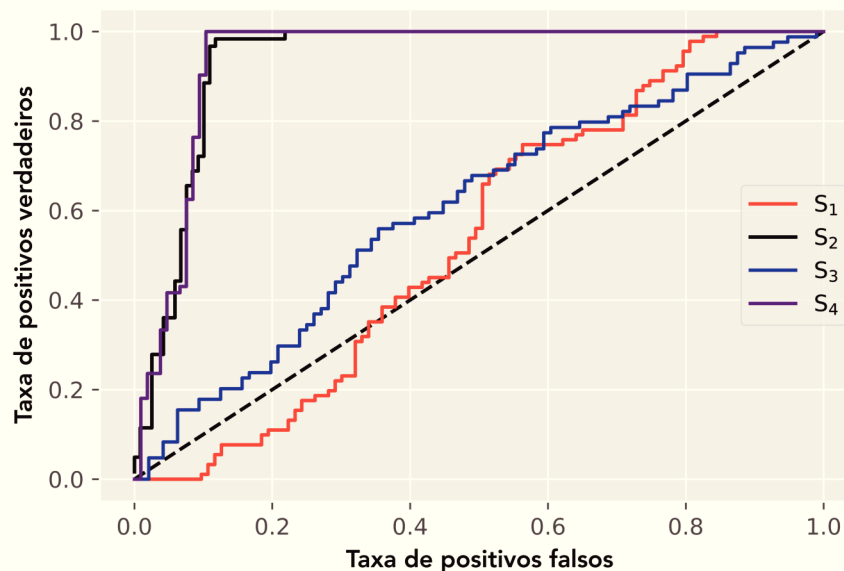
escolha não é baseada em dados, ou seja, não é considerado que pode haver uma correlação entre as características data e quantidade do pedido, ii.) utilizando os algoritmos KNN e iii.) LR, assim como com a combinação de ambos algoritmos por meio dos métodos iv.) Híbrida A e v.) Híbrida B.

4 RESULTADOS E DISCUSSÕES

Nesta seção, os resultados do experimento são apresentados utilizando a combinação dos modelos de simulação ao modelo de ML para uma seleção de fornecedores orientada a dados.

Primeiramente, a performance da LR é apresentada nas Figuras 32, 33, 34 e 35 utilizando quatro diferentes sementes de simulação (n_1 a n_4). A área abaixo da curva ROC traduz a acurácia da predição para cada fornecedor avaliado. É importante recordar que as características em estudo nesta experimentação são o período em que o pedido foi realizado e a quantidade de produtos do pedido.

Figura 32 – Curva ROC de todos os fornecedores com a semente de simulação n_1 .

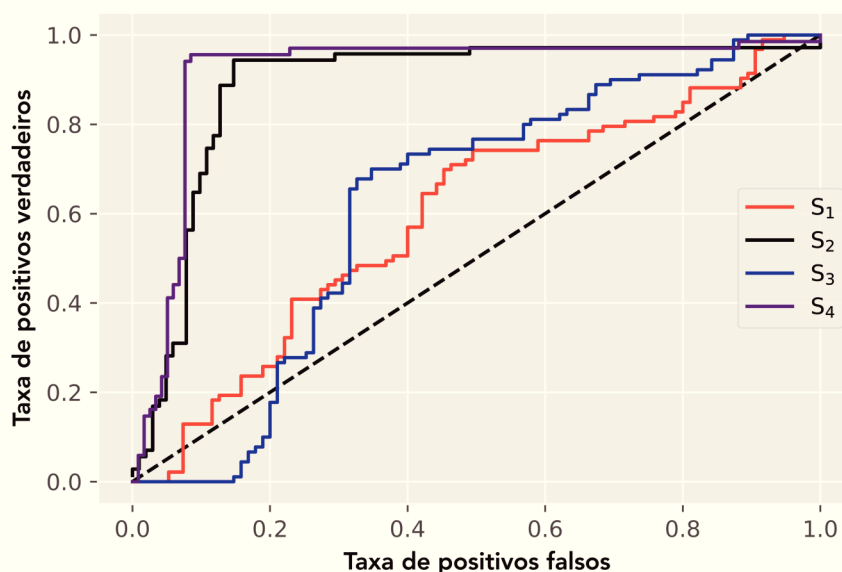


Fonte: Autoria própria (2019)

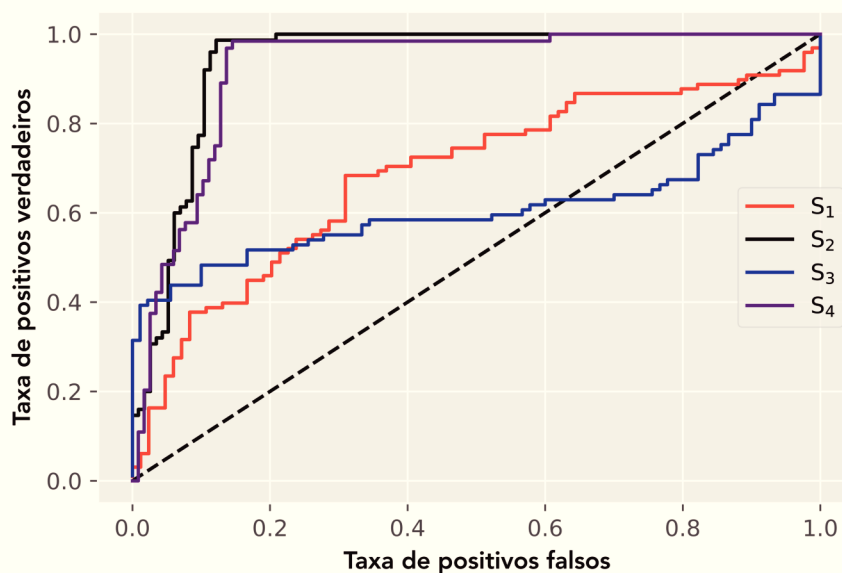
Observa-se que como o área abaixo da curva ROC dos fornecedores 1 e 3 (S_1 e S_3) é inferior a área abaixo da curva ROC dos fornecedores 2 e 4 (S_2 e S_4), é possível inferir que, baseado em dados históricos de performance, o algoritmo LR prediz a performance de S_2 e S_4 melhor que a performance de S_1 e S_3 . Então, dependendo das características dos fornecedores, modelos mais acurados podem ser encontrados utilizando o mesmo algoritmo.

Além disso, um extrato das curvas ROC de S_1 e S_2 utilizando diferentes sementes de simulação (n_1 a n_4) é apresentado nas Figuras 36 e 37. É possível notar que os valores são aproximados nas curvas de S_2 em comparação à curva de S_1 . Isto pode ser explicado pela melhor definição da restrição de capacidade que foi modelada para S_2 em comparação a S_1 , o que faz com que S_2 seja mais predizível.

Uma amostra da classificação executada pelo algoritmo LR é mostrada na Ta-

Figura 33 – Curva ROC de todos os fornecedores com a semente de simulação n_2 .

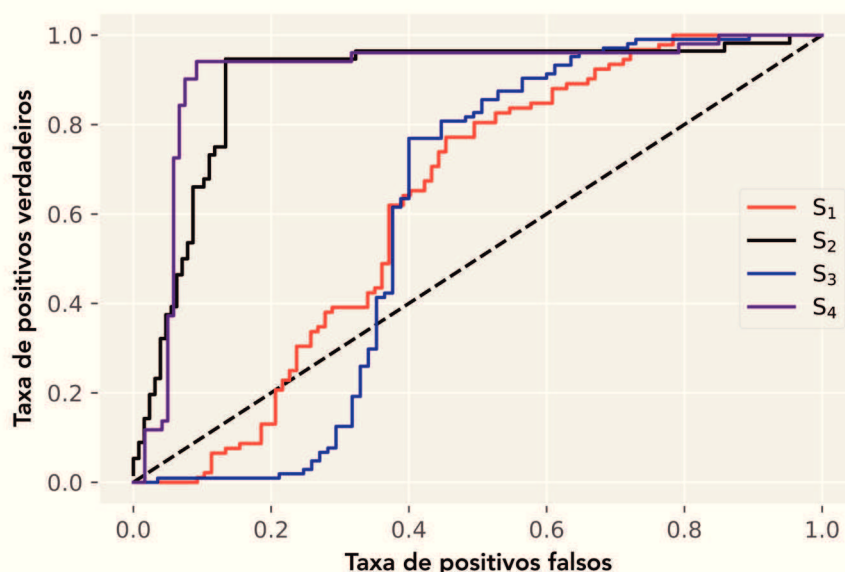
Fonte: Autoria própria (2019)

Figura 34 – Curva ROC de todos os fornecedores com a semente de simulação n_3 .

Fonte: Autoria própria (2019)

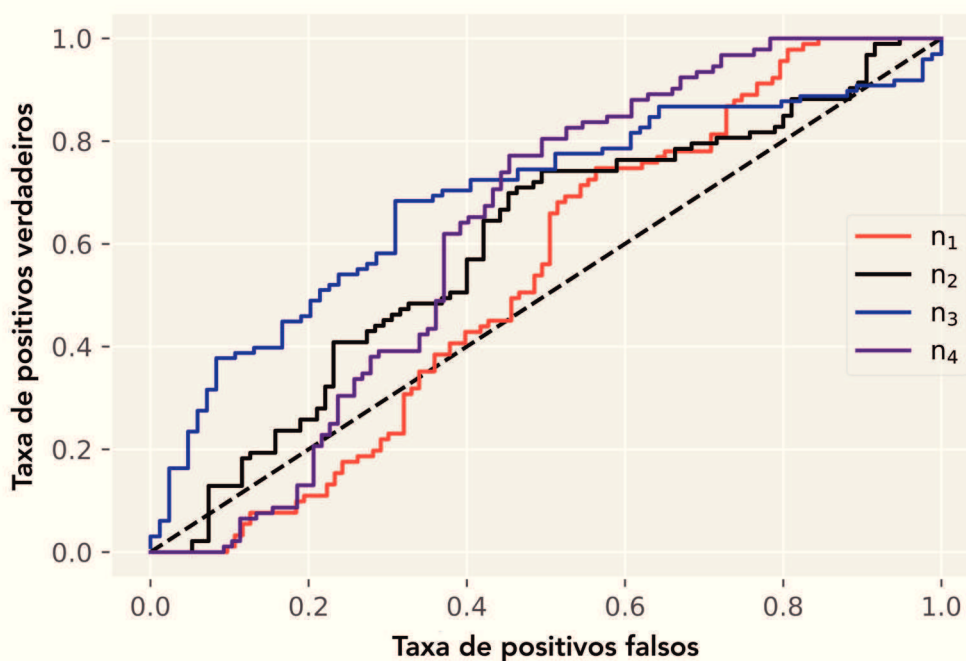
bela 9. O algoritmo quantifica a probabilidade de cada fornecedor entregar cada pedido no prazo especificado. Esta abordagem pode atuar como suporte ao processo de tomada de decisão pois possui um aspecto importante de evitação do viés humano, já que a decisão é baseada exclusivamente em dados quantitativos históricos referentes à performance da entrega. Nesta dissertação, o algoritmo LR realiza a seleção de fornecedores por si só e também em combinação com o algoritmo KNN.

Figura 35 – Curva ROC de todos os fornecedores com a semente de simulação n_4 .



Fonte: Autoria própria (2019)

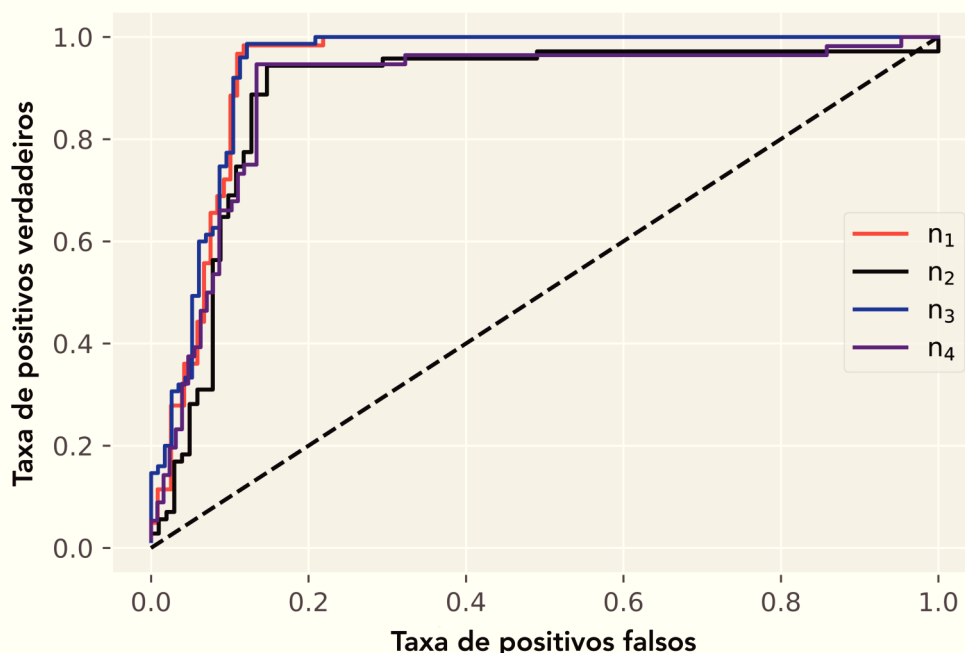
Figura 36 – Curva ROC do fornecedor S_1 após simulação com quatro diferentes sementes de simulação (n_1 to n_4).



Fonte: Autoria própria (2019)

Como mencionado anteriormente, a performance do algoritmo KNN é medido pela sua acurácia. Cinco sementes diferentes são apresentadas na Figura 38, na qual é mostrada a acurácia do algoritmo KNN neste modelo. Neste estudo, duas predições são realizadas utilizando o modelo KNN, no caso com o i.) Rótulo das entregas no prazo e ii.) Rótulo das entregas atrasadas. Em outras palavras, em i.) o modelo su-

Figura 37 – Curva ROC do fornecedor S_2 após simulação com quatro diferentes sementes de simulação (n_1 to n_4).



Fonte: Autoria própria (2019)

gere qual o fornecedor com maior probabilidade de entrega no prazo dado um pedido específico, e em ii.) o modelo sugere qual o fornecedor com maior probabilidade de entrega atrasada dado um pedido específico.

Os resultados apresentados ilustram o potencial uso de modelos de ML como ferramentas para suporte à tomada de decisão. As simulações utilizando diferentes sementes de simulação foram realizadas para testar a performance destes modelos baseada na confiabilidade de entrega, i.e., a taxa de entregas no prazo.

Duas simulações foram realizadas baseadas nos dois cenários previamente mencionados. Cada simulação foi repetida utilizando cinco diferente sementes de simulação e os resultados são apresentados na Figura 39 por meio dos valores médios.

Tabela 9 – Predições de probabilidade de entrega no prazo utilizando o algoritmo LR para cada fornecedor e os resultados da classificação.

Pedido	S_1	S_2	S_3	S_4	R_{lr1}	R_{lr2}
1	0.4361	0.1972	0.4980	0.1887	S_3	S_1
2	0.3319	0.3402	0.6014	0.3642	S_3	S_4
...
454	0.6236	0.2755	0.3248	0.2475	S_1	S_3
455	0.3313	0.9232	0.6015	0.9493	S_4	S_2
...

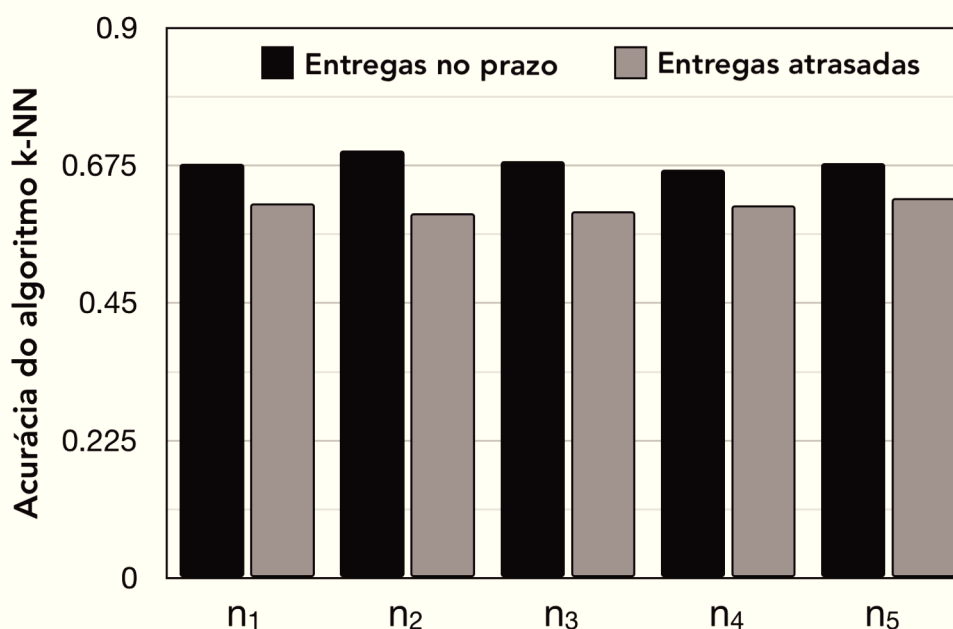
Continua na próxima página

Tabela 9 – Predições de probabilidade de entrega no prazo utilizando o algoritmo LR para cada fornecedor e os resultados da classificação.

Pedido	S_1	S_2	S_3	S_4	R_{lr1}	R_{lr2}
729	0.3230	0.0730	0.6108	0.0672	S_3	S_1
730	0.4156	0.2006	0.5178	0.1947	S_3	S_1

Fonte: Autoria própria (2019)

Figura 38 – Comparação da acurácia do modelo que utiliza o algoritmo KNN com cinco sementes de simulação diferentes.

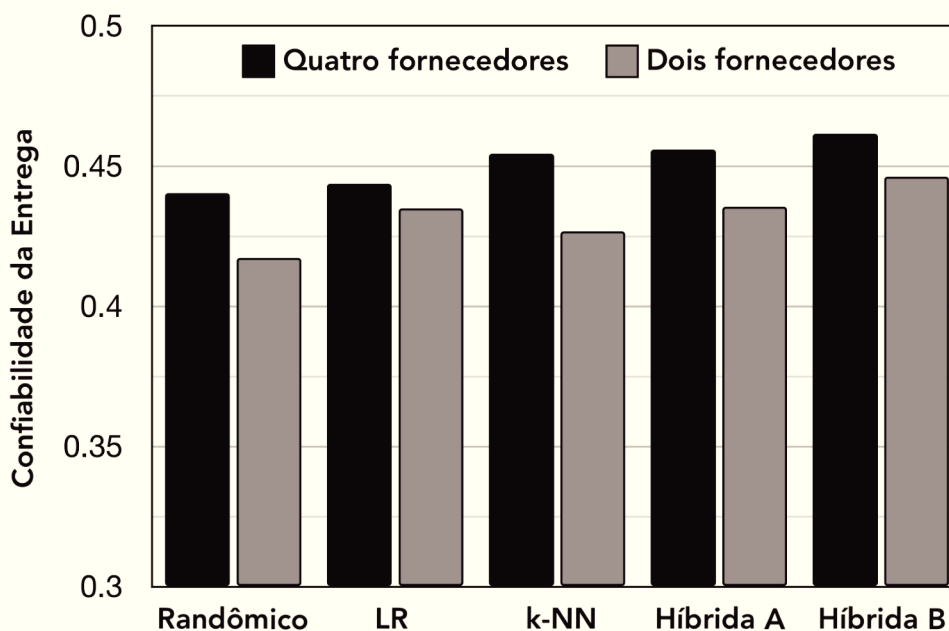


Fonte: Autoria própria (2019)

Os resultados do experimento sugerem que um maior número de fornecedores leva a um sistema mais resiliente para lidar com riscos recorrentes. Em parte, isto se deve ao fato de que existem modelos assertivos que identificam os fornecedores adequados para cada pedido específico. Portanto, há mais opções para serem avaliadas com o objetivo de fazer uma boa seleção.

No entanto, é importante mencionar que existe um *trade-off*, como o número de pedidos durante um período não muda, um maior número de fornecedores significa que menos dados estão sendo analisados para cada um deles. Por exemplo, no caso deste experimento o número total de pedidos é 2921, logo, como esta base de dados não muda, haveria menos informação para cada um dos fornecedores caso ao invés de 2 ou 4 fornecedores estivessem sendo avaliados 6 ou 8 fornecedores. Então, com menos informação para cada fornecedor, há uma tendência de piora na representação

Figura 39 – Performance da confiabilidade de entrega após a seleção de fornecedores utilizando as técnicas de ML supervisionado, LR e KNN.



Fonte: Autoria própria (2019)

real pelo modelo de ML, o que se traduz em menos acurácia, e com menos acurácia, piores resultados seriam obtidos para uma seleção de fornecedores.

Também é possível observar que o uso misto de dois algoritmos de ML culminou em uma melhoria na confiabilidade de entrega dos fornecedores. Por exemplo, utilizando uma abordagem randômica para seleção de fornecedores – o que significa que este tipo de dados não está sendo analisado por modelos de ML –, a confiabilidade de entrega foi de 44,03%, e adotando o modelo Híbrido B, o resultado aumentou para 46,16%. Num universo de 2921 pedidos, isto significaria que 62 entregas atrasadas foram evitadas com a simples utilização de dados disponíveis que a princípio não seriam analisados.

Analisando e generalizando os resultados deste estudo, é possível observar que, de fato, ter um *Big Data* é inútil se não for aplicado na tomada de decisão (GANDOMI; HAIDER, 2015). E em uma sociedade cada vez mais influenciada por tomadas de decisão orientada a dados, o uso de todo e qualquer dado disponível tem potencial de gerar novos mecanismos de negociação e novas formas de tomadas de decisão.

Serviços emergentes utilizando analytics, i.e., análise computacional sistemática de dados, incluindo tecnologias como armazém de dados, ML e visualizações (LARSON; CHANG, 2016) são uma nova forma de criar valor na era da manufatura digital. Portanto, uma nova realidade complexa e disruptiva emerge e decisões em níveis táticos e estratégicos devem considerar o impacto da fragmentação digital em todos os aspectos do negócio (ABBOSH; NUNES; OVANESSOFF, 2017), incluindo a fragmentação dos relacionamentos entre empresas de manufatura e seus fornecedo-

res.

Relacionamentos ad-hoc entre clientes e fornecedores podem se tornar uma realidade mais comum com a adoção de uma cultura orientada a dados em empresas de manufatura. Por exemplo, isto pode aumentar o poder de barganha das empresas que realizam contratos inteligentes baseado nos modelos preditivos de seleção de fornecedores. Além disso, com o auxílio de algoritmos de ML para realizar análises de performance dos fornecedores, pode ser possível contribuir para o desenvolvimento de modelos mais robustos de gerenciamento de riscos que, quando aplicados em situações reais, podem resultar em cadeias de suprimento mais resilientes a eventos de falha recorrentes e até disruptivos.

5 CONCLUSÕES

A principal contribuição deste trabalho está no fato de integrar a simulação de sistemas produtivos e técnicas de ML para resolver um problema de manufatura, neste caso, a seleção de fornecedores. Além disso, foram discutidos conceitos fundamentais relacionados ao paradigma da manufatura digital e os potenciais serviços digitais orientados a dados que podem emergir para apoio à tomada de decisão. Portanto, esta pesquisa foi motivada pelo grande leque de oportunidades de pesquisa relacionados aos diversos serviços digitais orientados a dados que estão na eminência de surgir devido, principalmente, à disponibilidade e viabilidade de novas tecnologias e técnicas de análises de dados.

Nesta dissertação de mestrado, a seleção de fornecedores por meio de algoritmos de ML supervisionado foi discutida. Um foco específico da análise foi direcionada para a seleção de fornecedores no contexto da manufatura digital. Os casos de teste foram realizados em um ambiente de manufatura digital utilizando uma ferramenta de simulação. Portanto, conclui-se que:

- Os resultados sugerem que o uso de algoritmos de ML supervisionados podem apoiar o processo de tomada de decisões na seleção de fornecedores, levando a uma maior previsibilidade da performance de entrega no prazo. No experimento, os fornecedores tiveram restrições na capacidade de produção em determinados períodos e foram representados em um conjunto de dados dividido por categorias como data e quantidade do pedido. O modelo de ML supervisionado foi capaz de prever o desempenho dos fornecedores quando ocorreram variações nestas características.
- Na experimentação digital realizada nesta pesquisa, foi constatado que a utilização dos modelos de ML ocasionou uma melhor performance de entregas dos suprimentos no prazo. Foi observado que a utilização combinada dos algoritmos LR e KNN obtiveram uma performance superior à utilização dos algoritmos individualmente. Isto pode ser um indício que modelos de aprendizagem podem ser combinados por meio de sistemas baseados em regras para obter melhores resultados na classificação dos fornecedores. Além disso, como esperado, os resultados mostraram que a consideração da acurácia dos algoritmos nos modelos combinados torna os resultados das previsões melhores.
- O uso de ML supervisionado pode contribuir para a seleção de fornecedores como uma estratégia de mitigação de riscos que pode auxiliar os modelos de gerenciamento de riscos na cadeia de suprimentos. Com o advento da disponibilidade de *Big Data*, a tomada de decisões na manufatura se tornará cada vez mais dependente de métodos estatísticos. Portanto, é essencial pavimentar

o caminho de substituição de abstrações por modelos de ML em processos de gerenciamento de risco na manufatura, para que a criação de valor possa ser percebida por profissionais da indústria e dados reais possam ser compartilhados, levando a um ciclo virtuoso de melhoria.

Finalmente, algumas limitações e futuros caminhos de pesquisa podem ser destacados:

- Embora o modelo considere variações estocásticas para se aproximar de um caso real, o modelo ainda é baseado em dados não reais, portanto, os resultados estão sujeitos a variações em cenários reais.
- Para aplicações reais de empresas orientadas a dados, um maior número de características nas bases de dados existirá devido ao aumento da disponibilidade de dados, e portanto, nesses casos, uma pré-seleção de características para identificar as mais relevantes para o modelo de previsão pode ser utilizada, bem como o uso de técnicas de DL. Para esse fim, o modelo de simulação pode ser estendido adicionando variabilidade de produto, custos de transporte, entre outras características personalizadas.
- Além disso, é possível investigar diferentes algoritmos de ML supervisionados, bem como novos métodos combinando dois ou mais deles, considerando a acurácia dessas técnicas para obter melhores resultados.

REFERÊNCIAS

ABBOSH, O.; NUNES, P.; OVANESSOFF, A. Adapting Your Digital Business to a Fragmented World. **Harvard Business Review**, v. 95, p. 56–61, 2017.

ALPAYDIN, Ethem. **Introduction to Machine Learning**. [S.l.]: MIT press, 2014.

ARTHUR, Charles. Tech giants may be huge, but nothing matches big data. **The Guardian**, v. 23, 2013.

BAINES, Tim S; LIGHTFOOT, Howard W; BENEDETTINI, Ornella *et al.* The servitization of manufacturing: A review of literature and reflection on future challenges. **Journal of manufacturing technology management**, Emerald Group Publishing Limited, v. 20, n. 5, p. 547–567, 2009.

BAINES, Tim S; LIGHTFOOT, Howard W; EVANS, Steve *et al.* State-of-the-art in product-service systems. **Proceedings of the Institution of Mechanical Engineers, Part B: journal of engineering manufacture**, Sage Publications Sage UK: London, England, v. 221, n. 10, p. 1543–1552, 2007.

BAINES, Tim; W. LIGHTFOOT, Howard. Servitization of the manufacturing firm: Exploring the operations practices and technologies that deliver advanced services. **International Journal of Operations & Production Management**, Emerald Group Publishing Limited, v. 34, n. 1, p. 2–35, 2013.

BERKOVICH, Marina; LEIMEISTER, Jan Marco; KRCMAR, Helmut. Requirements engineering for product service systems. **Business & Information Systems Engineering**, Springer, v. 3, n. 6, p. 369–380, 2011.

BODAGHI, Gholamreza; JOLAI, Fariborz; RABBANI, Masoud. An integrated weighted fuzzy multi-objective model for supplier selection and order scheduling in a supply chain. **International Journal of Production Research**, Taylor & Francis, v. 56, n. 10, p. 3590–3614, 2018.

BOEHM, Matthias; THOMAS, Oliver. Looking beyond the rim of one's teacup: a multidisciplinary literature review of Product-Service Systems in Information Systems, Business Management, and Engineering & Design. **Journal of Cleaner Production**, Elsevier, v. 51, p. 245–260, 2013.

BORSHCHEV, Andrei; FILIPPOV, Alexei. From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools. *In*: CITESEER. PROCEEDINGS of the 22nd international conference of the system dynamics society. [S.l.: s.n.], 2004.

BURKOV, Andriy. **The Hundred-Page Machine Learning Book**. [S.l.: s.n.], 2019.

CENAMOR, Javier; SJÖDIN, D Rönnerberg; PARIDA, Vinit. Adopting a platform approach in servitization: Leveraging the value of digitalization. **International Journal of Production Economics**, Elsevier, v. 192, p. 54–65, 2017.

CHEN, Amy; HSIEH, Chih-Ying; WEE, HM. A resilient global supplier selection strategy – a case study of an automotive company. **The International Journal of Advanced Manufacturing Technology**, Springer, v. 87, n. 5-8, p. 1475–1490, 2016.

CHEN, Ping-Shun; WU, Ming-Tsung. A modified failure mode and effects analysis method for supplier selection problems in the supply chain risk environment: A case study. **Computers & Industrial Engineering**, Elsevier, v. 66, n. 4, p. 634–642, 2013.

CHENG, Colin CJ; KRUMWIEDE, Dennis. What makes a manufacturing firm effective for service innovation? The role of intangible capital under strategic and environmental conditions. **International Journal of Production Economics**, Elsevier, v. 193, p. 113–122, 2017.

COVER, Thomas; HART, Peter. Nearest neighbor pattern classification. **IEEE Transactions On Information Theory**, IEEE, v. 13, n. 1, p. 21–27, 1967.

DAGKAKIS, Georgios; HEAVEY, Cathal. A review of open source discrete event simulation software for operations research. **Journal of Simulation**, Taylor & Francis, v. 10, n. 3, p. 193–206, 2016.

DAVIS, Jim *et al.* Smart manufacturing, manufacturing intelligence and demand-dynamic performance. **Computers & Chemical Engineering**, Elsevier, v. 47, p. 145–156, 2012.

DE BOER, Luitzen; LABRO, Eva; MORLACCHI, Pierangela. A review of methods supporting supplier selection. **European journal of purchasing & supply management**, Elsevier, v. 7, n. 2, p. 75–89, 2001.

DESMET, Steven *et al.* **Servitization: or why services management is relevant for manufacturing environments**. [S.l.]: Pearson Education Limited, 2013.

DEY, Prasanta Kumar *et al.* Strategic supplier performance evaluation: A case-based action research of a UK manufacturing organisation. **International Journal of Production Economics**, Elsevier, v. 166, p. 192–214, 2015.

DHAR, Vasant. Data science and prediction. **Communications of the ACM**, ACM, v. 56, n. 12, p. 64–73, 2013.

DICKSON, Gary W. An analysis of vendor selection systems and decisions. **Journal of Purchasing**, Wiley Online Library, v. 2, n. 1, p. 5–17, 1966.

ECK, Nees van; WALTMAN, Ludo. Software survey: VOSviewer, a computer program for bibliometric mapping. **Scientometrics**, Akadémiai Kiadó, co-published with Springer Science+ Business Media BV, v. 84, n. 2, p. 523–538, 2009.

ECONOMIST, The. The world most valuable resource - Data and the new rules of competition. **The Economist**, European Union, 2017.

FALAGAS, Matthew E *et al.* Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. **The FASEB journal**, Federation of American Societies for Experimental Biology, v. 22, n. 2, p. 338–342, 2008.

FAWCETT, Tom. An introduction to ROC analysis. **Pattern Recognition Letters**, Elsevier, v. 27, n. 8, p. 861–874, 2006.

FIGUEIRA, Gonçalo; ALMADA-LOBO, Bernardo. Hybrid simulation–optimization methods: A taxonomy and discussion. **Simulation Modelling Practice and Theory**, Elsevier, v. 46, p. 118–134, 2014.

FLATH, Christoph M; STEIN, Nikolai. Towards a data science toolbox for industrial analytics applications. **Computers in Industry**, Elsevier, v. 94, p. 16–25, 2018.

FRAZZON, Enzo Morosini *et al.* Towards socio-cyber-physical systems in production networks. **Procedia Cirp**, Elsevier, v. 7, p. 49–54, 2013.

GANDOMI, Amir; HAIDER, Murtaza. Beyond the hype: Big data concepts, methods, and analytics. **International Journal of Information Management**, Elsevier, v. 35, n. 2, p. 137–144, 2015.

GEBAUER, Heiko; FLEISCH, Elgar; FRIEDLI, Thomas. Overcoming the service paradox in manufacturing companies. **European management journal**, Elsevier, v. 23, n. 1, p. 14–26, 2005.

GEORGES, Marcos Ricardo Rosa *et al.* Metodologia para modelagem em simulação de sistemas: aplicação em manufatura discreta. [sn], 2005.

GOEDKOOOP, Mark. Product service systems. **Ecological and economic basis**, 1999.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

GRÖGER, Christoph. Building an Industry 4.0 Analytics Platform. **Datenbank-Spektrum**, Springer, v. 18, n. 1, p. 5–14, 2018.

HALEVY, Alon; NORVIG, Peter; PEREIRA, Fernando. The unreasonable effectiveness of data. **IEEE Intelligent Systems**, IEEE, v. 24, n. 2, p. 8–12, 2009.

- HAMDI, Faiza *et al.* Optimization of a supply portfolio in the context of supply chain risk management: literature review. **Journal of Intelligent Manufacturing**, Springer, v. 29, n. 4, p. 763–788, 2018.
- HOSSEINI, Seyedmohsen; BARKER, Kash. A Bayesian network model for resilience-based supplier selection. **International Journal of Production Economics**, Elsevier, v. 180, p. 68–87, 2016.
- HOUAISS, Antônio; VILLAR, Mauro de Salles; FRANCO, Francisco Manoel de Mello. Dicionário da língua portuguesa. **Rio de Janeiro: Objetiva**, p. 353, 2001.
- JAHANGIRIAN, Mohsen *et al.* Simulation in manufacturing and business: A review. **European Journal of Operational Research**, Elsevier, v. 203, n. 1, p. 1–13, 2010.
- JORDAN, Michael I; MITCHELL, Tom M. Machine learning: Trends, perspectives, and prospects. **Science**, American Association for the Advancement of Science, v. 349, n. 6245, p. 255–260, 2015.
- JÜTTNER, Uta; PECK, Helen; CHRISTOPHER, Martin. Supply chain risk management: outlining an agenda for future research. **International Journal of Logistics: Research and Applications**, Taylor & Francis, v. 6, n. 4, p. 197–210, 2003.
- KAGERMANN, Henning. Change through digitization? Value creation in the age of Industry 4.0. *In*: MANAGEMENT of permanent change. [S.l.]: Springer, 2015. p. 23–45.
- KAGERMANN, Henning *et al.* **Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry; final report of the Industrie 4.0 Working Group**. [S.l.]: Forschungsunion, 2013.
- KAGERMANN, H *et al.* Smart service welt: recommendations for the strategic initiative web-based services for businesses. **Berlin: Acatech-National Academy of Science and Engineering**, 2014.
- KAMALAHMADI, Masoud; PARAST, Mahour Mellat. An assessment of supply chain disruption mitigation strategies. **International Journal of Production Economics**, Elsevier, v. 184, p. 210–230, 2017.
- KANG, Hyoung Seok *et al.* Smart manufacturing: Past research, present findings, and future directions. **International Journal of Precision Engineering and Manufacturing-Green Technology**, Springer, v. 3, n. 1, p. 111–128, 2016.
- KARLSSON, Anna; LARSSON, Lisa; ÖHRWALL RÖNNBÄCK, Anna. Product-service system innovation capabilities: linkages between the fuzzy front end and subsequent development phases. **International Journal of Production Research**, Taylor & Francis, v. 56, n. 6, p. 2218–2232, 2018.

KLAUS, Schwab. **The fourth industrial revolution**. [S.l.: s.n.], 2016.

KOCH, Volkmar *et al.* Industry 4.0: Opportunities and challenges of the industrial internet. **Strategy & PwC**, 2014.

KOHAVI, Ron *et al.* A study of cross-validation and bootstrap for accuracy estimation and model selection. *In*: MONTREAL, CANADA, 2. IJCAI. [S.l.: s.n.], 1995. p. 1137–1145.

KOTSIANTIS, SB; KANELLOPOULOS, D; PINTELAS, PE. Data preprocessing for supervised learning. **International Journal of Computer Science**, Citeseer, v. 1, n. 2, p. 111–117, 2006.

KOTSIANTIS, Sotiris B; ZAHARAKIS, I; PINTELAS, P. Supervised machine learning: A review of classification techniques. **Emerging artificial intelligence applications in computer engineering**, v. 160, p. 3–24, 2007.

KUSIAK, Andrew. Smart manufacturing. **International Journal of Production Research**, Taylor & Francis, v. 56, n. 1-2, p. 508–517, 2018.

KUSIAK, Andrew. Smart manufacturing must embrace big data. **Nature News**, v. 544, n. 7648, p. 23, 2017.

LARSON, Deanne; CHANG, Victor. A review and future direction of agile, business intelligence, analytics and data science. **International Journal of Information Management**, Elsevier, v. 36, n. 5, p. 700–710, 2016.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **nature**, Nature Publishing Group, v. 521, n. 7553, p. 436, 2015.

LEE, Jay *et al.* Predictive manufacturing system-Trends of next-generation production systems. **IFAC Proceedings Volumes**, Elsevier, v. 46, n. 7, p. 150–156, 2013.

LEOPOLD, Alexander Till; ZAHIDI, Saadia; RATCHEVA, Vesselina. The future of jobs: Employment, skills and workforce strategy for the fourth industrial revolution. *In*: WORLD Economic Forum. [S.l.: s.n.], 2016.

LI, Ni *et al.* Human performance modeling for manufacturing based on an improved KNN algorithm. **The International Journal of Advanced Manufacturing Technology**, Springer, v. 84, n. 1-4, p. 473–483, 2016.

LI, ZP *et al.* Supplier selection decision-making in supply chain risk scenario using agent based simulation. *In*: IEEE. INDUSTRIAL Engineering and Engineering Management (IEEM), 2015 IEEE International Conference on. [S.l.: s.n.], 2015. p. 900–904.

LIU, Tianyu; DENG, Yong; CHAN, Felix. Evidential supplier selection based on DEMATEL and game theory. **International Journal of Fuzzy Systems**, Springer, v. 20, n. 4, p. 1321–1333, 2018.

LU, Qi. Microsoft touts developer tools, business software at Build. *Seattle Times*, 2017.

MANZINI, Ezio; VEZZOLI, Carlo. A strategic design approach to develop sustainable product service systems: examples taken from the ?environmentally friendly innovation? Italian prize. **Journal of cleaner production**, Elsevier, v. 11, n. 8, p. 851–857, 2003.

MARR, Bernard. **The Most Amazing Artificial Intelligence Milestones So Far**. [S.l.: s.n.], 2018. <https://www.forbes.com/sites/bernardmarr/2018/12/31/the-most-amazing-artificial-intelligence-milestones//so-far/>. Acessado em 08 de janeiro de 2019.

MCCULLOCH, Warren S; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.

MIRKOU EI, Amin; HAAPALA, Karl R. Integration of machine learning and mathematical programming methods into the biomass feedstock supplier selection process. **Flexible Automation and Intelligent Manufacturing**, 2014.

MITCHELL, Tom M *et al.* Machine learning. 1997. **Burr Ridge, IL: McGraw Hill**, v. 45, n. 37, p. 870–877, 1997.

MIYAGI, Paulo E. Introdução a simulação discreta. **Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos. São Paulo**, 2006.

MONOSTORI, László *et al.* Cyber-physical systems in manufacturing. **CIRP Annals**, Elsevier, v. 65, n. 2, p. 621–641, 2016.

MONT, Oksana. **Product-service systems: panacea or myth?** [S.l.]: IIIIEE, Lund University, 2004.

MOORE, Susan. **Gartner Says More Than 40 Percent of Data Science Tasks Will Be Automated by 2020**. [S.l.: s.n.], 2017. <https://www.gartner.com/en/newsroom/press-releases/2017-01-16-gartner-says-more-than-40-percent-of-data-science-tasks-will-be-automated-by-2020>. Acessado em 11 de janeiro de 2019.

NEELY, Andy. Exploring the financial consequences of the servitization of manufacturing. **Operations management research**, Springer, v. 1, n. 2, p. 103–118, 2008.

OMAR, Mohamed K. The solutions of DT and AHP for supplier selection problem. *In: IEEE. QUALITY and Reliability (ICQR), 2011 IEEE International Conference on. [S.l.: s.n.], 2011. p. 506–510.*

OPRESNIK, David; TAISCH, Marco. The value of big data in servitization. **International Journal of Production Economics**, Elsevier, v. 165, p. 174–184, 2015.

PEDREGOSA, F. *et al.* Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PINEDO, Michael L. **Scheduling: theory, algorithms, and systems**. [S.l.]: Springer, 2016.

PITELIS, Christos N. The co-evolution of organizational value capture, value creation and sustainable advantage. **Organization studies**, SAGE Publications Sage UK: London, England, v. 30, n. 10, p. 1115–1139, 2009.

RABETINO, Rodrigo; KOHTAMÄKI, Marko; GEBAUER, Heiko. Strategy map of servitization. **International Journal of Production Economics**, Elsevier, v. 192, p. 144–156, 2017.

RADDATS, Chris; EASINGWOOD, Chris. Services growth options for B2B product-centric businesses. **Industrial Marketing Management**, Elsevier, v. 39, n. 8, p. 1334–1345, 2010.

RAJAGOPAL, Varthini; VENKATESAN, Shanmugam Prasanna; GOH, Mark. Decision-making models for supply chain risk mitigation: A review. **Computers & Industrial Engineering**, Elsevier, v. 113, p. 646–682, 2017.

RANDHAWA, Krithika; SCERRI, Moira. Service innovation: A review of the literature. *In: THE handbook of service innovation. [S.l.]: Springer, 2015. p. 27–51.*

REN, Guangjie; GREGORY, MJ. Servitization in manufacturing companies: a conceptualization, critical review, and research agenda, 2007.

RIBEIRO, Mauro; GROLINGER, Katarina; CAPRETZ, Miriam AM. Mlaas: Machine learning as a service. *In: IEEE. MACHINE Learning and Applications (ICMLA), 2015 IEEE 14th International Conference on. [S.l.: s.n.], 2015. p. 896–902.*

ROMETTY, VM. **A Conversation with Ginni Rometty**. [S.l.: s.n.], 2013.

RONDINI, Alice *et al.* Hybrid simulation modelling as a supporting tool for sustainable product service systems: a critical analysis. **International Journal of Production Research**, Taylor & Francis, v. 55, n. 23, p. 6932–6945, 2017.

RUSSELL, Stuart J; NORVIG, Peter. **Artificial intelligence: a modern approach**. [S./]: Malaysia; Pearson Education Limited, 2016.

RÜSSMANN, Michael *et al.* Industry 4.0: The future of productivity and growth in manufacturing industries. **Boston Consulting Group**, v. 9, 2015.

RYMASZEWSKA, Anna; HELO, Petri; GUNASEKARAN, Angappa. IoT powered servitization of manufacturing—an exploratory case study. **International Journal of Production Economics**, Elsevier, v. 192, p. 92–105, 2017.

SAWIK, Tadeusz. On the risk-averse optimization of service level in a supply chain under disruption risks. **International Journal of Production Research**, Taylor & Francis, v. 54, n. 1, p. 98–113, 2016.

SCHLÄPFER, Ralf C; KOCH, Markus; MERKHOFER, P. Industry 4.0 challenges and solutions for the digital transformation and use of exponential technologies. **Deloitte, Zurich**, 2015.

SHEFFI, Yossi; RICE JR., James B. A supply chain view of the resilient enterprise. **MIT Sloan Management Review**, Massachusetts Institute of Technology, Cambridge, MA, v. 47, n. 1, p. 41, 2005.

SLACK, Nigel; CHAMBERS, Stuart; JOHNSTON, Robert. **Operations management**. [S./]: Pearson education, 2013.

SUTTON, Richard S; BARTO, Andrew G. **Reinforcement learning: An introduction**. [S./]: MIT press, 2018.

TOMLIN, Brian. On the value of mitigation and contingency strategies for managing supply chain disruption risks. **Management Science**, INFORMS, v. 52, n. 5, p. 639–657, 2006.

TUKKER, Arnold. Eight types of product–service system: eight ways to sustainability? Experiences from SusProNet. **Business strategy and the environment**, Wiley Online Library, v. 13, n. 4, p. 246–260, 2004.

TUKKER, Arnold. Product services for a resource-efficient and circular economy—a review. **Journal of cleaner production**, Elsevier, v. 97, p. 76–91, 2015.

TUKKER, Arnold; TISCHNER, Ursula. **New business for old Europe: product-service development, competitiveness and sustainability**. [S./]: Routledge, 2017.

TURING, Alan M. On computable numbers, with an application to the Entscheidungsproblem. **Proceedings of the London mathematical society**, Wiley Online Library, v. 2, n. 1, p. 230–265, 1937.

VANDERMERWE, Sandra; RADA, Juan. Servitization of business: adding value by adding services. **European management journal**, Elsevier, v. 6, n. 4, p. 314–324, 1988.

VIEIRA, Guilherme Ernani *et al.* Evaluating the Robustness of Production Schedules using Discrete-Event Simulation. **IFAC-PapersOnLine**, Elsevier, v. 50, n. 1, p. 7953–7958, 2017.

WEE, D *et al.* Industry 4.0-how to navigate digitization of the manufacturing sector. **McKinsey & Company**, v. 58, 2015.

WEIDMANN, Dominik *et al.* Product-Service System Development with Discrete Event Simulation. *In: 2015 IEEE International Systems Conference (SysCon 2015)*. [S.l.: s.n.], 2015.

WETZSTEIN, Anton *et al.* A systematic assessment of supplier selection literature—state-of-the-art and future scope. **International Journal of Production Economics**, Elsevier, v. 182, p. 304–323, 2016.

WHITE, Allen L; STOUGHTON, Mark; FENG, Linda. Servicizing: the quiet transition to extended product responsibility. **Tellus Institute, Boston**, v. 97, 1999.

WHITEHEAD, Alfred North; RUSSELL, Bertrand. **Principia mathematica**. [S.l.]: University Press, 1912. v. 2.

WILL M. BERTRAND, J; FRANSOO, Jan C. Operations management research methodologies using quantitative modeling. **International Journal of Operations & Production Management**, MCB UP Ltd, v. 22, n. 2, p. 241–264, 2002.

WISE, Richard; BAUMGARTNER, Peter. Go downstream: The new profit imperative in manufacturing. **IEEE Engineering Management Review**, IEEE, v. 28, n. 1, p. 89–96, 2000.

WU, Desheng; OLSON, David L. Supply chain risk, simulation, and vendor selection. **International Journal of Production Economics**, Elsevier, v. 114, n. 2, p. 646–655, 2008.

WUEST, Thorsten *et al.* Machine learning in manufacturing: advantages, challenges, and applications. **Production & Manufacturing Research**, Taylor & Francis, v. 4, n. 1, p. 23–45, 2016.

YOON, Jiho *et al.* Models for supplier selection and risk mitigation: a holistic approach. **International Journal of Production Research**, Taylor & Francis, v. 56, n. 10, p. 3636–3661, 2018.

YU, Jianbo; LIU, Jianping. LRProb control chart based on logistic regression for monitoring mean shifts of auto-correlated manufacturing processes. **International Journal of Production Research**, Taylor & Francis, v. 49, n. 8, p. 2301–2326, 2011.

ZHANG, Fuqiang *et al.* Modeling and analyzing of an enterprise collaboration network supported by service-oriented manufacturing. **Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture**, Sage Publications Sage UK: London, England, v. 226, n. 9, p. 1579–1593, 2012.

ZHANG, Min *et al.* Innovating through services, co-creation and supplier integration: Cases from China. **International Journal of Production Economics**, Elsevier, v. 171, p. 289–300, 2016.