

Análise de Dados de Fornecedores e Subcontratados para Apoio à Tomada de Decisão no Planeamento da Produção: O caso do *Cluster* do Calçado
Ricardo Miguel Barros Ferreira

Ricardo Miguel Barros Ferreira. Análise de Dados de Fornecedores e Subcontratados para Apoio à Tomada de Decisão no Planeamento da Produção: O caso do *Cluster* do Calçado

2021/2022

Análise de Dados de Fornecedores e Subcontratados para Apoio à Tomada de Decisão no Planeamento da Produção: O caso do *Cluster* do Calçado
Ricardo Miguel Barros Ferreira

2021/2022



Análise de Dados de Fornecedores e Subcontratados para Apoio à Tomada de Decisão no Planeamento da Produção: O caso do *Cluster* do Calçado

Ricardo Miguel Barros Ferreira

Prof. Doutor Cristóvão Dinis Polido Sousa

Agradecimentos

Primeiramente, fica um agradecimento especial para o meu orientador da tese, Professor Doutor Cristóvão Dinis Polido Sousa, pela sua disponibilidade, pela paciência, e ajuda de forma a conseguir realizar todos os passos necessários para a realização da tese e também para o meu colega que me acompanhou neste projeto, Pedro Pinto, conseguimos os dois juntos ultrapassar todos os obstáculos, e ajudar-nos mutuamente em todo o projeto. Um obrigado ao Eng. Rui Rebelo e ao INESC TEC, pela disposição, apoio prestado durante o desenvolvimento do projeto e oportunidade dada.

Outro agradecimento vai para a Escola Superior de Tecnologia e Gestão, escola que me acolheu nestes cinco anos e me ensinou muito e que levo esses ensinamentos para a minha vida pessoal e profissional. Não deixar de agradecer também a todo o corpo docente que me deixou todos os seus ensinamentos e sempre tiveram disponíveis para me ajudar em tudo.

Também fica o agradecimento aos meus amigos, que foram muito importantes em todo o meu percurso e não me deixaram ir abaixo em nenhum momento, um obrigado a todos eles sem exceção. Por fim, como não podia faltar gostava de agradecer a toda a minha família e namorada, pessoas que me apoiaram em todo o caminho realizado e sempre se preocuparam comigo, mas principalmente aos meus pais, que me ensinaram que com esforço, trabalho e dedicação se consegue tudo na vida, é graças a eles que sou quem sou, um muito obrigado a eles.

Resumo

Nos últimos anos, com a evolução da indústria, o aumento da competitividade no setor e a pressão que existe para que as empresas sejam mais flexíveis para responder ao mercado, têm sido iniciados investimentos em tecnologias com o objetivo de aumentar a produtividade dos seus recursos, aumento do lucro e início do processo de transformação digital. Nesse contexto, as empresas tentam combater um problema que já existe há algum tempo, o planejamento da produção do produto, uma vez que esse processo engloba uma rede colaborativa de diversos recursos, envolvendo uma gestão complicada, demorada e passíveis falhas. Assim, esta dissertação descreve a situação atual do setor do calçado, o problema que enfrenta e as formas de resolver este problema com base em métodos de avaliação e modelos de previsão, que incidirão em toda a rede colaborativa envolvente.

Palavras-chave: Indústria, Sistemas de Informação, Planejamento, Avaliação, Classificação, Previsão, *Machine Learning*.

Abstract

In recent years, with the evolution of the industry, the increased competitiveness in the sector and the pressure that exists for companies to be more flexible to respond to the market, investments in technologies have been initiated with the aim of increasing the productivity of their resources, increase profit and start the process of digital transformation. In this context, companies try to combat a problem that has existed for some time, the product production planning, since this process encompasses a collaborative network of various resources, involving complicated management, time-consuming, and possible failure management. Therefore, this dissertation describes the current situation of the footwear sector, the problem they face and ways to solve this problem based on assessment methods and forecasting models, which will focus on the entire surrounding collaborative network.

Keywords: Industry, Information Technologies, Planning, Evaluation, Classification, Prediction, Machine Learning.

Conteúdo

Agradecimentos	i
Resumo	ii
Abstract	iii
Lista de Figuras	vii
Lista de Tabelas	ix
Siglas e Abreviaturas	xii
1 Introdução	1
1.1 Contexto do problema	1
1.2 Objetivos	3
1.3 Metodologia	4
1.4 Resultados Esperados	6
1.5 Estrutura do Documento	6
2 <i>Planeamento da Produção Orientado a Dados</i>	9
2.1 Abordagens baseadas em Classificação e Avaliação	9
2.1.1 Abordagens de Tomada de Decisão Multicritério (MCDM)	10

2.1.2	Métodos de Programação Matemática (PM)	14
2.1.3	Abordagens <i>Artificial Intelligence</i> (AI)	15
2.2	Abordagens Preditivas	19
2.2.1	<i>Machine Learning</i> (ML))	19
2.3	Reflexão Crítica	22
3	Conceptualização do Domínio	23
3.1	Caracterização do Setor	23
3.2	Visão geral da Rede Colaborativa	25
3.2.1	Caracterização da Rede Colaborativa	25
4	Caso de Uso	31
4.1	Descrição do Problema	32
4.2	Abordagens Desenvolvidas	35
4.2.1	Avaliação de Fornecedores e Subcontratados baseada em AHP	35
4.2.2	Abordagem de <i>Machine Learning</i> na Previsão do Plano de Produção	43
5	Resultados	51
6	Conclusões	57
6.1	Reflexão Crítica	58
6.2	Trabalho Futuro	58
	Bibliografia	59

A Anexos	64
A.1 Anexo A - Definição das Matrizes de Importância	64
A.2 Anexo B - Automatização de atribuição de importância	65
A.3 Anexo C - Análise de Dados dos Fornecedores	66

Lista de Figuras

1	Fases da Metodologia de Trabalho.	4
2	Estruturação das Fases de Trabalho.	5
3	Estrutura da AHP.	10
4	Exemplo da Árvore de Decisão.	16
5	Estrutura da Rede Neuronal.	17
6	Exemplo do particionamento do <i>K-Means</i>	18
7	Estrutura Exemplo da <i>Random Forest</i>	21
8	Rede Colaborativa do Setor do Calçado.	25
9	Processo de Negócio do Setor do Calçado.	27
10	<i>SMART Supply Chain</i>	33
11	Arquitetura do Sistema de Avaliação de Fornecedores e Subcontratados.	33
12	Arquitetura do Sistema de Acompanhamento e Alerta para o Planeamento	34
13	Processo de Construção da AHP.	35
14	Índice de Consistência)	37
15	Rácio de Consistência	37

16	Estrutura da <i>Analytic Hierarchy Process</i> (AHP) baseada no problema.	38
17	Processo de Construção de um algoritmo de <i>Machine Learning</i>	43
18	<i>Mean Absolute Error</i> (MAE)	44
19	<i>Mean Squared Error</i> (MSE)	44
20	<i>Root Mean Squared Error</i> (RMSE)	44
21	Plataforma colaborativa (Classificação da família de materiais usada como exemplo).	52
22	Importância das variáveis (Modelo com melhores resultados).	53
23	Plataforma colaborativa- Exemplo de previsão do tempo de entrega.	54
24	Análise dos Dados dos Fornecedores	66
25	Histograma de tempo de entrega dos Fornecedores	66

Lista de Tabelas

1	Escala de <i>Saaty</i>	36
2	Caracterização do dataset utilizado no algoritmo AHP.	38
3	Matriz de Importância.	39
4	Exemplo da divisão da escala de Saaty no critério de Entrega.	41
5	Exemplo da matriz de importância do critério Entrega.	42
6	Características do dataset utilizado no algoritmo de <i>Machine Learning</i>	45
7	Dados Estatísticos das variáveis numéricas do dataset.	47
8	Parâmetros variáveis dos modelos.	49
9	Exemplo da classificação final relativa a uma família de materiais.	51
10	Precisão dos algoritmos <i>Machine Learning</i> (ML) testados com base em diferentes medidas de erro.	53
11	Exemplo da divisão da escala no critério Origem (Família de Artigo TER).	65
12	Exemplo da divisão da escala no critério Preço (Família de Artigo TER).	65

Exemplos de Código

1	Inserção da matriz de importância no algoritmo.	40
2	Exemplo do modelo de dados do relatório do algoritmo.	40
3	Função para construção da matriz de importância no critério Entrega.	41
4	Utilização da biblioteca <i>ahpy</i> para cálculo do peso.	42
5	Exemplo da análise exploratória do algoritmo.	47
6	Função de remoção de <i>outliers</i>	48
7	Função de <i>Feature Scaling</i>	48
1	Função para construção das matrizes de importância.	64

Siglas e Abreviaturas

AHP	<i>Analytic Hierarchy Process</i>
AI	<i>Artificial Intelligence</i>
ANP	<i>Analytic Network Process</i>
CBR	<i>Case-Based Reasoning</i>
CI	<i>Consistency Index</i>
CR	<i>Consistency Ratio</i>
DEA	<i>Data Envelopment Analysis</i>
DSR	Design Science Research
IoT	<i>Internet of Things</i>
MAE	<i>Mean Absolute Error</i>
MCDM	<i>Multi-Criteria Decision Making</i>
ML	<i>Machine Learning</i>
MSE	<i>Mean Squared Error</i>
NN	<i>Neural Network</i>
PM	<i>Programação Matemática</i>
PME	Pequenas e Médias Empresas
R2	R-Quadrado
RMSE	<i>Root Mean Squared Error</i>

RN Redes Neurais

SVM *Support Vector Machine*

Capítulo 1

Introdução

1.1 Contexto do problema

Nos dias de hoje, com um mercado cada vez mais dinâmico e exigente, é necessário que as indústrias sejam flexíveis de modo a ganharem vantagem competitiva. Este é o principal ingrediente que tem vindo a promover a evolução da indústria ao longo dos tempos, e a indústria do calçado não é exceção. As indústrias têm de ser capazes de flexibilizar a sua capacidade produtiva de modo a responder eficazmente à heterogeneidade cada vez maior das encomendas, ou seja, encomendas de dimensões diferentes, com níveis de personalização cada vez maiores e para um número cada vez mais alargado de clientes. A esta dinâmica do lado da procura, acresce o fator da sazonalidade, ou seja, existe produto (calçado) de inverno e produto de verão, levando a um planeamento atempado da produção de modo a que seja possível ter o produto ao dispor do consumidor no devido tempo.

Para incrementar a flexibilidade dentro da indústria e a garantia de prazos de entrega é necessário conseguir garantir dinâmica no planeamento, de forma a gerir de uma forma mais eficaz a capacidade das empresas, que na maior parte das vezes não dependem apenas delas, mas de toda uma cadeia de fornecimento cada vez mais complexa, incluindo fornecedores, sendo estes fornecedores de matéria-prima e de serviços de produção (i.e., subcontratados). De um ponto de vista mais agregado a dinâmica do planeamento depende de uma rede, onde nem sempre é possível controlar as entidades que nela participam. Consequentemente, o planeamento da produção de uma determinada organização, depende, em certa medida, da qualidade do planeamento de terceiros. Neste sentido, a seleção dos parceiros certos a cada

interação do processo de decisão no âmbito do planeamento de produção é um fator crítico de sucesso. Neste contexto, um parceiro certo seria um fornecedor que, face ao contexto produtivo, teria melhores condições e maior probabilidade de entregar matéria-prima com a qualidade necessária e no horizonte temporal que permita cumprir os prazos de entrega acordados com o cliente. Considerando o exposto, o processo de tomada de decisão para a elaboração do planeamento da produção requer, não apenas informação sobre diversos aspetos internos à organização, mas também informação relevante relacionada com a cadeia de valor em que se insere, nomeadamente no que se refere aos seus fornecedores e subcontratados.

Algumas das estratégias encontradas na literatura para lidar com este problema, focam a redução da rede, através de parcerias de longa duração com os fornecedores mais confiáveis, com o objetivo de reduzir o risco da compra e melhorar o serviço, indo assim de encontro à filosofia “*just-in-time*” [1]. Neste contexto surge a problemática da seleção dos melhores fornecedores, uma vez que, um fornecedor económico pode não responder às necessidades em tempo útil, e um fornecedor pontual pode não entregar a qualidade desejada para combater possíveis falhas na seleção dos fornecedores é necessário, fazer uma seleção correta dos critérios de avaliação [2] [3]. O modelo que estrutura e articula tais critérios, necessita de acomodar dados intra e inter-organizacionais de modo, a proporcionar uma análise capaz de aferir os níveis de interação da empresa com a rede colaborativa de fornecedores e subcontratados para que, a seleção do parceiro (fornecedor/subcontratado) seja eficaz e contribua para um planeamento mais eficiente. Sendo assim, o desenvolvimento de um sistema de classificação de redes colaborativas automático e eficaz requer uma abordagem sociotécnica, considerando, por um lado, os sistemas de informação que suportam a indústria do calçado e as características da rede colaborativa e seus relacionamentos. Nesse sentido, os principais desafios para a construção do modelo de classificação, com o objetivo de apoiar o planeamento da produção, estão relacionados aos critérios e à qualidade dos dados disponíveis. Assim, o sistema de classificação automática da rede colaborativa a desenvolver deverá estar em condições de responder aos seguintes requisitos principais:

- Classificar automaticamente os vários intervenientes na cadeia de valor (nomeadamente fornecedores e subcontratados);
- Apresentar sugestões dos fornecedores/subempreiteiros mais preparados para fornecer ou prestar os serviços requeridos, bem como uma previsão do prazo de entrega.

Juntamente com a classificação de fornecedores e subcontratados, o planeamento proativo exige atividades de previsão, que envolvem toda a rede colaborativa. Uma das atividades identificadas é a previsão de prazos de entrega de matéria-prima, para indicar ao planeador se ele terá todos os materiais a tempo de iniciar as operações de fabrico. Portanto, o verdadeiro desafio deste projeto pode ser colocado através da seguinte questão de investigação: **Como uma correta avaliação de fornecedores e subcontratados pode contribuir para o aumento da flexibilidade da indústria, especificamente na capacidade de realizar um planeamento dinâmico das operações para otimizar os prazos de entrega para clientes?** A questão de investigação proposta fundamenta a seguinte hipótese: **Através do aprimoramento do processo de seleção de fornecedores e do uso de métodos preditivos para monitorização da rede de operações colaborativas, é possível melhorar o processo de previsão e, assim, aumentar a capacidade de cumprir a entrega nos prazos acordados com os clientes. Na seleção de fornecedores e na escolha de métodos preditivos de monitoramento das operações envolvendo toda a rede colaborativa, é possível um consequente aumento da eficiência no trabalho de planeamento.**

1.2 Objetivos

Este trabalho tem como principal propósito a melhoria do planeamento do setor do calçado, tendo em conta toda a rede colaborativa envolvente. Sendo assim, o principal objetivo deste projeto passa pela utilização dos dados existentes com o intuito de desenvolver sistemas que auxiliem na tomada de decisão. De forma mais concreta os objetivos são:

1. Recolher e analisar os dados da empresa relacionados com pedidos de compras, suas encomendas, seus produtos, materiais e processo produtivo;
2. Identificar padrões nos dados e usar técnicas de ML para apoiar as atividades de planeamento de produção por meio de previsão;
3. Classificar os fornecedores como um, meio de apoio à tomada de decisão nas atividades de planeamento da produção.

Com a concretização deste objetivo, os resultados esperados são: um sistema capaz de avaliar os intervenientes da rede colaborativa da empresa, e assim construir uma ferramenta que

permite apoiar o planeador no processo de tomada de decisão, na escolha dos fornecedores, bem como um modelo de previsão capaz de realizar a previsão de datas de entrega, fornecendo ao planeador uma forte ferramenta que, lhe permite construir uma escala de risco às tarefas que possam não ser concluídas dentro do prazo previsto.

1.3 Metodologia

A estruturação da metodologia de trabalho é uma fase crucial na elaboração de trabalhos técnico-científicos, de modo a obter artefactos com valor científico e simultaneamente úteis para o domínio do problema. Neste sentido o foco do trabalho não é o desenvolvimento de soluções de engenharia para um problema específico, mas sim o desenho de artefactos técnicos para um domínio de problemas. Assim, a metodologia seguida para este trabalho, segue princípios de *Design Science Research (DSR)* [4].

De acordo com a Figura 1, a metodologia seguida contempla 4 fases, nomeadamente: i) Estudo e Compreensão do Domínio, ii) Compreensão do Problema, iii) Desenho dos Artefactos, iv) Validação .

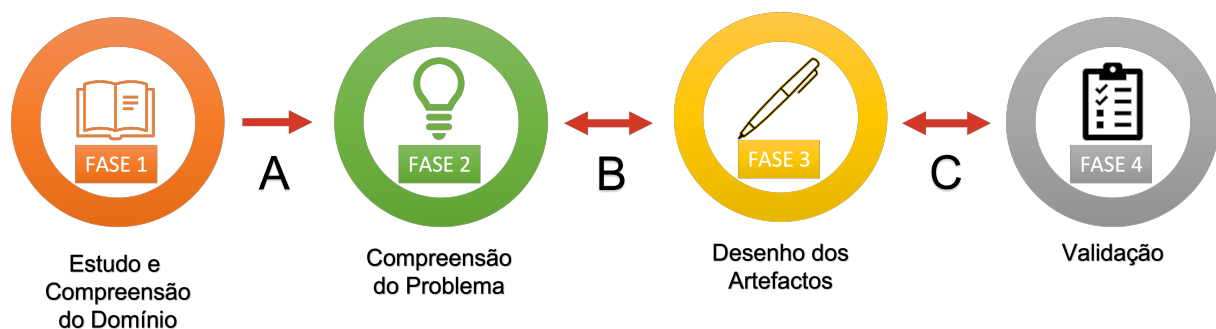


Figura 1: Fases da Metodologia de Trabalho.

Com base na metodologia DSR, para além das fases, foram incluídos com 3 milestones fundamentais, representados pelas letras A, B e C respetivamente. Estes milestones representam ciclos de interação entre as fases, de modo a garantir a entrega de artefactos técnico-científicos válidos e úteis para o utilizador. A metodologia reflete, deste modo, uma abordagem iterativa e incremental em detrimento de uma abordagem sequencial. De igual modo as iterações não ocorrem apenas nas fases mas entre as fases. A compreensão e conceptualização do problema, por exemplo, é garantida quer pela interação entre a estudo da literatura (que é constante ao longo do desenvolvimento do trabalho), e o feedback da avaliação dos artefac-

tos. Isto significa que, para cada fase contribui o resultado da anterior mais o feedback da seguinte. De igual modo entre as fases 2 a 4 existe uma interação bi-direcional, uma vez que, a validação dos desenvolvimentos pode afetar o desenho dos artefactos e surge a necessidade de compreender melhor o problema e voltar a realizar o desenho dos artefactos. Deste modo, o milestone A, assegura o clico de relevância do problema. Já o milestone B, garante o desenho de artefactos de acordo com as necessidades do domínio, úteis para o problema específico mas generalizáveis para problemas similares. Por fim, o milestone C, assegura o ciclo de rigor. Os artefactos são avaliados e evoluem progressivamente até à sua versão final.

Concretamente, nesta abordagem metodológica os artefactos são avaliados qualitativamente, no final, por intermédio de casos de estudo e, ao longo do seu processo de desenvolvimento, através de protótipos discutidos com os especialistas do domínio. As fases apresentadas são constituídas por certas tarefas, da realização dessas tarefas temos resultados que dão um contributo importante para todo o desenvolvimento do trabalho. Na Figura 2 é ilustrado estes conjuntos.

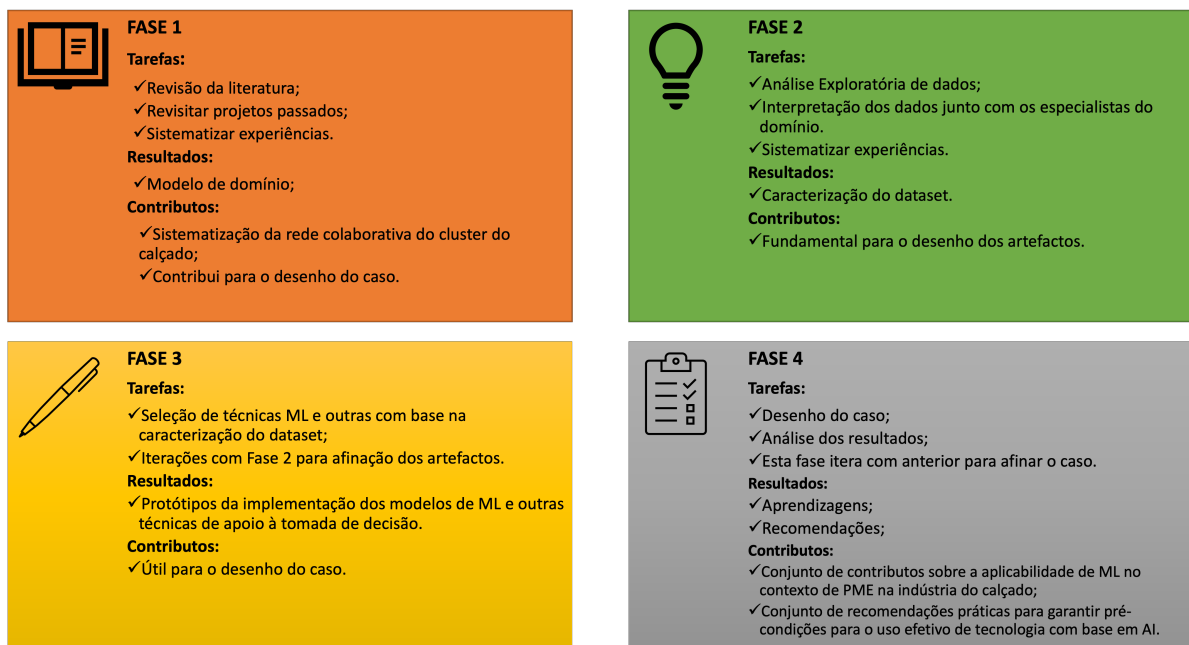


Figura 2: Estruturação das Fases de Trabalho.

1.4 Resultados Esperados

Perante os objetivos apresentados anteriormente, numa fase inicial foi necessário o compreender de todo o domínio do negócio em questão, com vista a facilitar a interpretação dos dados do negócio, bem como a adequada seleção dos mesmos. Numa fase seguinte foi possível realizar um estudo sobre as várias abordagens que podem melhorar o processo de planeamento produtivo, nomeadamente abordagens de classificação e avaliação, abordagens preditivas baseadas em ML, abordagens matemáticas.

Realizado o estudo foi necessário a extração e tratamento de dados e a construção de mecanismos suportados pelas abordagens referidas anteriormente. A construção destes tem como objetivo, a classificação dos intervenientes externos da rede colaborativa envolvente, e também o auxílio do planeamento na previsão dos tempos de entregas dos mesmos.

Por fim toda a solução apresentada foi aplicada a um caso de estudo, no âmbito do projeto CNI4.0-Foot¹. Esta solução foi integrada no sistema de informação associado ao projeto, bem como testada em ambiente real, sendo possível discutir as alternativas identificadas por estes mecanismos no suporte à decisão no decorrer das tarefas de planeamento, em termos de escolha dos parceiros correctos e também na previsão temporal na realização das mesmas tarefas.

1.5 Estrutura do Documento

O documento encontra-se dividido em 6 capítulos, cada capítulo aborda um tema relevante para a dissertação.

- **Capítulo 1 - Introdução:** é apresentada a motivação e o contexto do trabalho, bem como os seus objetivos, os resultados esperados, metodologia de trabalho e como é estruturado o documento.
- **Capítulo 2 - Planeamento da Produção Orientado a Dados:** são apresentadas as diversas abordagens que visam resolver os problemas relacionados com a seleção de fornecedores e a utilização dos dados de forma a realizar previsão das tarefas existentes

¹<https://www.inesctec.pt/pt/projetos/cni40foot>

no setor.

- **Capítulo 3 - *Conceptualização do Domínio***: realização de uma descrição dos pontos principais do setor em causa, desde a sua caracterização, a descrição da rede colaborativa envolvente e uma overview a todo o seu processo de negócio.
- **Capítulo 4 - *Caso de Uso***: é apresentado o projeto e problema onde se enquadra o trabalho desenvolvido, bem como a descrição práticas de todas as abordagens utilizadas.
- **Capítulo 5 - *Resultados***: são apresentados todos os resultados obtidos bem como a sua interpretação.
- **Capítulo 6 - *Conclusões e Trabalho Futuro***: é efetuada uma conclusão de todo o trabalho realizado, assim como a sua reflexão crítica e trabalhos futuros.

No final do documento encontra-se um apêndice, onde estão ilustrados alguns anexos relativos à plataforma associada ao trabalho, bem como alguns exertos de código.

Capítulo 2

Planeamento da Produção Orientado a Dados

Uma vez definidos os problemas a serem abordados no projeto, é necessário a realização de uma revisão de literatura, com o objetivo de perceber as formas já existentes e quais as abordagens utilizadas na solução dos tipos problemas em causa. Portanto foi realizada uma divisão nos tipos de abordagens, abordagens baseadas em classificação e avaliação e abordagens preditivas.

2.1 Abordagens baseadas em Classificação e Avaliação

A avaliação de fornecedores é um tema já debatido uma vez que, a dificuldade da seleção correta de fornecedores já é uma questão antiga, o que faz com que existam grandes quantidades de modelos definidos. Junyi Chai em 2013, realizou uma revisão de literatura relativa às abordagens já retratadas. Este autor com o objetivo de as categorizar, realiza uma separação em três tipos, sendo eles: *Multi-Criteria Decision Making Multi-Criteria Decision Making* (MCDM), *Mathematical Programming* **MP!** (**MP!**) e *Artificial Intelligence* AI [5].

2.1.1 Abordagens de Tomada de Decisão Multicritério (MCDM)

A Tomada de Decisão Multicritério, em inglês *Multi-Criteria Decision Making* MCDM, é das abordagens mais reconhecidas neste tipo de problemas de decisão, sendo por volta dos anos 60 que existiram alguns avanços relativos a esta temática [6]. Estas abordagens têm como objetivo recomendar aos tomadores de decisão, dentro de um conjunto finito de alternativas, qual a indicada, esta escolha é realizada através da avaliação dos vários critérios selecionados [7]. O problema da seleção de fornecedores, recorrendo à literatura, é geralmente classificado como um problema de MCDM [8].

Estes tipos de abordagens seguem na sua generalidade certas etapas sendo elas: i) Formular o problema, ii) Identificar os requisitos, iii) Definir os objetivos, iv) Identificar o conjunto das alternativas, v) Desenvolver critérios de avaliação, vi) Identificar a técnica de tomada de decisão e aplicar utilizando os critérios definidos anteriormente [9]. Os métodos baseados neste tipo de abordagens podem ser classificados em quatro tipos: Métodos de utilidade multiatributo, Métodos de *Outranking*, Métodos de Compromisso e outros métodos [5].

2.1.1.1 Métodos de utilidade multiatributo: AHP e ANP

Os métodos de utilidade multiatributo são utilizados com o intuito de atribuir um valor de utilidade a cada alternativa existente. Este valor pode representar um valor de referência para a realização de uma classificação das alternativas, este é obtido através da avaliação dos critérios definidos [10]. Alguns dos exemplos destes tipos de métodos são a AHP e a *Analytic Network Process* (ANP), a sua estrutura é ilustrada abaixo na Figura 3

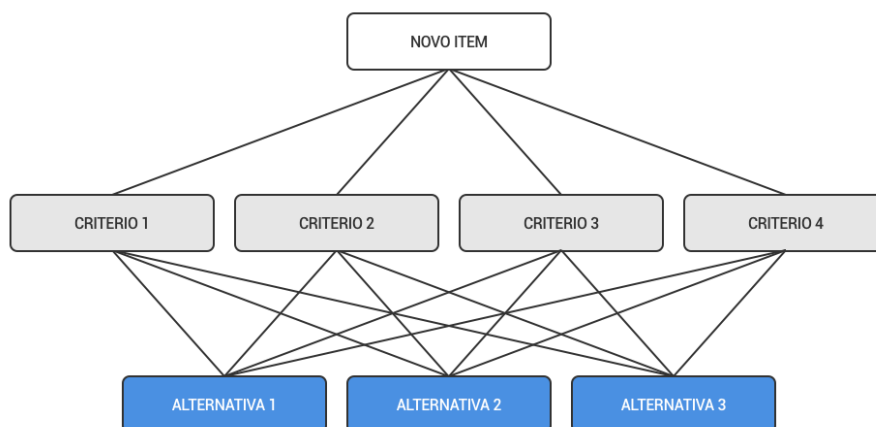


Figura 3: Estrutura da AHP.

O *Analytic Hierarchy Process* AHP, método desenvolvido por Saaty no início da década de 1980 [11], baseado na matemática e na psicologia humana, visa auxiliar a tomada de decisão sobre qual é a melhor escolha a ser feita diante de um problema. É um método de decisão que prioriza alternativas quando há multicritérios a serem considerados, esta priorização é identificada através da avaliação dos critérios definidos, como já referido anteriormente [12]. É amplamente utilizado neste tipo de problemas devido, ao facto de ser um método de fácil implementação, fácil análise de decisões, possibilidade de avaliar critérios quantitativos e qualitativos utilizando uma escala de preferência [13].

No que diz respeito *Analytic Network Process* ANP, é uma extensão da AHP, apenas difere em duas situações: i) em termos da estrutura, sendo que a AHP é disposta em uma estrutura hierárquica, enquanto que a ANP utiliza uma estrutura em rede; ii) e relativo à independência entre as alternativas e atributos, uma vez que, em contraste com a AHP não assume que as alternativas, atributos e critérios sejam independentes entre si [14]. Alguns dos exemplos da implementação destes métodos estão retratados no artigo de Chan em 2010, utilizando o método na seleção de fornecedores na indústria da moda [15], e no artigo de Amin em 2019, com a utilização da ANP na gestão dos intervenientes da *Supply Chain* [16].

2.1.1.2 Métodos de *Outranking*: ELECTRE e PROMETHEE

Os modelos de *Outranking* são utilizados, geralmente, quando existe a necessidade de lidar com decisões de múltiplos critérios, onde coabitam tanto critérios qualitativos como quantitativos [17]. Segundo Roy [18], este tipo de métodos define uma relação binária S num conjunto de X alternativas tais que xSy , sendo x e y critérios de decisão. Dado o que se sabe sobre as preferências do decisor e dada a qualidade das avaliações das ações e a natureza do problema, existem argumentos suficientes para declarar que x é pelo menos melhor que y , enquanto não há razão essencial para refutar esta decisão [19]. Alguns dos exemplos deste tipo de métodos são a ELECTRE e a PROMETHEE.

Com a utilização do ELECTRE (*Elimination and Choice Expressing Reality*), é fornecido um conjunto preciso e adequado de ações, eliminando as alternativas que são superadas por outras, com a utilização de multicritérios. Este método segue de forma rigorosa a metodologia dos métodos *Outranking*, baseando-se na opinião dos especialistas e nas informações disponíveis relacionadas ao problema subjacente [20]. Ao longo do tempo este método foi sendo reestruturado dando origem aos métodos, ELECTREI, ELECTREII, ELECTREIII, ELECTREIV,

ELECTRE IS, ELECTRE TRI [20]. Todos estes métodos têm por base os mesmo conceitos fundamentais, mas diferem em termos operacionais e diferentes tipos de problema. O ELECTREI é destinado a problemas de seleção, o ELECTRE TRI a problemas de afetação e os restantes a problemas de classificação [7].

O PROMETHEE, segue os conceitos fundamentais dos métodos *Outranking*, mas induz uma função preferencial para descrever a diferença de preferência entre os pares de alternativas em cada critério. Sendo assim, funções de preferência sobre diferença numérica entre pares de alternativas são construídas para, descrever a diferença de preferência do ponto de vista do tomador da decisão. Este valor varia entre 0 e 1, quanto maior o valor da função, maior é a diferença da preferência. No caso deste valor ser 0, não existe a diferença preferencial entre alternativas [21].

Tal como os métodos anteriores, existem alguns artigos onde estão descritos resultados da implementação destes métodos, *Vahdani* em 2010 utilizou o método ELECTRE com o intuito de realizar a seleção de fornecedores com pesos e dados de intervalo [22], também *Chen* em 2011 utilizou PROMETHEE com o objetivo de selecionar os fornecedores correctos, com vista a melhorias na organização, no que diz respeito a aumento de competitividade, redução de custos, entre outros [23].

2.1.1.3 Métodos de Compromisso: TOPSIS e VIKOR

Este tipo de método é geralmente a solução ideal para um problema com critérios incompatíveis. A sua metodologia tem por base a escolha de uma solução viável, sendo esta a mais próxima do ideal, e estabelece um acordo por concessões mútuas [24]. Esta solução é obtida através do uso de um índice de classificação baseado, em uma medida de proximidade da solução ideal positiva. O principal conceito deste tipo de método foca-se em soluções ideias positivas e negativas, sendo que a solução ótima é definida como a alternativa mais próxima da solução ideal positiva e mais distante da solução ideal negativa [25]. Os métodos principais que seguem esta metodologia são o TOPSIS e o VIKOR.

O TOPSIS (*Technique for order preference by similarity to an ideal solution*), assume que cada critério selecionado pode aumentar ou diminuir de forma repetitiva a utilidade, facilitando a definição dos valores positivos e negativos da solução ideal. O método utiliza a abordagem da

distância euclidiana ¹ para calcular a distância relativa das alternativas à solução ideal. Sendo assim, com a obtenção destes valores o método realiza uma comparação entre as alternativas e conclui uma classificação das alternativas [7].

O método VIKOR foi elaborado com o intuito de resolver o problema de tomada de decisão em que os critérios são comensuráveis e incompatíveis. Este concentra-se em classificar e identificar, a partir de um conjunto de alternativas, a solução compromisso, auxiliando os tomadores de decisão na definição da solução ideal. Esta escolha é realizada através da medida da "proximidade" da solução ideal, realizando de seguida uma ordenação das alternativas [27].

Estas duas técnicas apesar de seguirem os conceitos fundamentais dos métodos de compromisso, realizam diferentes tipos de normalização, o VIKOR usa uma normalização linear, enquanto que o TOPSIS utiliza uma normalização vetorial [27]. Algumas das utilizações destes métodos na seleção de fornecedores são descritas em 2009 por Crispim e De Sousa [28], com a utilização da TOPSIS na seleção de fornecedores em empresas virtuais.

2.1.1.4 Outros métodos

Existem outros dos métodos MCDM tais como o SMART (*Simple Multi-Attribute Rating Technique*), um método capaz de lidar com métodos quantitativos e qualitativos, apesar de não conseguir lidar de forma eficaz, quando existe incerteza na informação, como valores definidos em intervalos. Como todos os outros métodos o seu objetivo passa pela identificação da melhor solução dentro de um grupo de alternativas para isso ele realiza um cálculo de desempenho para todas as alternativas através da soma dos pesos atribuídos a cada critério, multiplicando depois pela importância definida para cada um desses critérios [29]. Outro dos métodos é DEMATEL, este é apropriado e útil para examinar relações interdependentes entre fatores num sistema complexo. Os fatores identificados são depois classificados e são obtidas as prioridades para o auxílio na tomada de decisão. O método consegue converter a relação entre os fatores causa e efeito em um modelo estrutural inteligente do sistema [30].

¹Distância entre dois pontos no espaço, podendo ser calculada através das coordenadas cartesianas dos pontos usando o Teorema de Pitágoras [26].

2.1.2 Métodos de Programação Matemática (PM)

Os métodos de *Programação Matemática* (PM), dependem não só dos critérios definidos pela empresa, mas também pelo objetivo desta seleção. Este modelo permite que a empresa crie um problema em torno de uma função objetivo, função esta que poderá ser em termos de maximizar ou de minimizar algo, variando os valores das variáveis na função [31]. Este tipo de métodos ao contrário dos anteriores não lida com métodos quantitativos. Na secção seguinte, são descritos alguns dos métodos que se regem por este conceito, tais como Programação Linear, *Data Envelopment Analysis* (DEA), Programação Multi-Objetivo, entre outros.

2.1.2.1 Data Envelopment Analysis (DEA)

O DEA é um método que tem por base o conceito de “eficiência” de uma alternativa de decisão. Esta avaliação é realizada através de critérios de benefício, critérios estes que representam o desempenho do fornecedor na realização dos serviços, acompanhada pelos critérios de custo, que apresentam os custos associados à escolha do fornecedor. A eficiência da alternativa é realizada através do rácio entre a soma dos pesos dos critérios de benefício e soma dos pesos dos critérios de custo [32]. Como *Weber*, em 1998 [33], já tinha citado na sua publicação, este método consegue ajudar a classificar o fornecedor em duas categorias sendo elas: fornecedores eficientes e fornecedores não eficientes, *Weber* dizia que este método não apenas categoriza o fornecedor mas também pode ser utilizado como ferramenta de negociação com fornecedores não eficientes [33].

2.1.2.2 Programação Linear e Não Linear

A Programação Linear é um método de otimização matemática para, determinar qual a solução para alcançar o melhor resultado em um determinado modelo matemático, sob um conjunto de requisitos estruturados como relações lineares [5]. Este método realiza as etapas essenciais dos métodos de programação matemática, a essas adiciona a etapa de identificação do conjunto de restrições, sendo através destas que é construída uma região viável, obtendo soluções válidas para a função objetivo. A Programação Não-linear tem por base a mesma ideologia do método de Programação linear, mas com uma diferença, a hipótese de conseguir lidar com algumas restrições ou função objetivo que seja não-lineares [34].

2.1.2.3 Programação Multi-Objectivo

Este método é caracterizado para problemas de decisão por funções objetivo múltiplas e incompatíveis, com a possibilidade de serem otimizadas sobre um conjunto de soluções viáveis [5]. O desafio na utilização deste tipo de métodos passa pela existência de não apenas uma solução ótima, mas sim um conjunto de soluções. Sendo assim, é necessário informações por parte do tomador da decisão com o objetivo de identificar a solução mais satisfatória [35].

2.1.2.4 Programação por metas (*Goal Programming*)

A programação por metas é um ramo da Programação Multi-Objectivo, que tem como objetivo ser uma extensão, de forma ser possível lidar com várias medidas objetivas normalmente incompatíveis. O método coloca uma meta a cada uma destas medidas como um valor-alvo a ser alcançado, sempre em vista de obter uma solução que melhor atinja as metas estabelecidas [36].

2.1.2.5 Programação Estocástica

A Programação Estocástica é um método mais direcionado para a resolução de problemas de otimização de incertezas, em que as distribuições de probabilidade que dominam os dados são conhecidas ou facilmente estimadas, apesar da existências de vários parâmetros desconhecidos [5]. A utilização deste no ambiente real traz algumas vantagens, uma vez que a existência de parâmetros desconhecidos é inevitável.

2.1.3 Abordagens AI

Auxiliados por um computador, estes modelos podem ser treinados por um especialista na tomada de decisões de seleção e avaliação de fornecedores, ou através dos dados de decisões anteriormente realizadas, e assim é possível que seja tomada uma decisão sustentada, através deste modelo, mesmo que por uma pessoa sem conhecimento dos processos de seleção [32]. Chai em 2020 [37], numa nova revisão da literatura, relativa ao tema em questão, realizou algumas diferenças no que diz respeito a este tipo de abordagens. Em comparação à revisão realizada anteriormente existe uma maior ênfase relativamente aos propósitos

funcionais de dados, no que diz respeito à classificação e agrupamento. Sendo assim, alguns dos métodos associados à classificação são Árvores de Decisão, Redes Neurais e o *Case-Based Reasoning* (CBR), em termos de métodos de agrupamentos temos o *K-means*, e por último temos o Algoritmo Genético que por norma não são afiliados a classificação ou agrupamento [37].

2.1.3.1 Métodos de Classificação

Os métodos de classificação visam reunir um conjunto de modelos ou classificadores, com o objetivo de representar as classes de dados importantes. Este método efetua normalmente duas etapas, uma primeira na realização de uma fase de aprendizagem ou treino através de um conjunto de dados, e uma segunda, que com auxílio do conjunto de modelos ou classificadores reunidos, classifica outro conjunto de dados ainda não classificados [37].

2.1.3.1.1 Árvores de Decisão

As árvores de decisão são um método que utiliza um modelo de decisões em forma de árvore, como é possível observar na Figura 4.

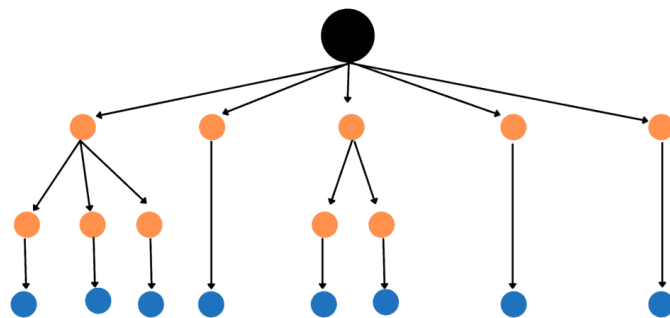


Figura 4: Exemplo da Árvore de Decisão.

Esta tem uma estrutura idêntica a um fluxograma em que cada nó representa um "teste" em um atributo ou subconjunto de atributos, cada ramo representa o resultado de cada teste e o no final, normalmente chamado de "folha", representa um rótulo e classe. Todos os caminhos existentes da raiz à folha representam regras de classificação [38]. Alguns de exemplos de árvores de decisão são: CART [39], C4.5 [39].

2.1.3.1.2 Rede Neuronal

As redes neuronais refletem o comportamento do cérebro humano, fornecendo a possibilidade de que as máquinas reconheçam padrões e resolvam problemas. A rede é um conjunto de entradas e saídas conectadas, em que cada conexão tem um peso atribuído [37]. Na Figura 5 é possível observar um exemplo da estrutura de uma rede.

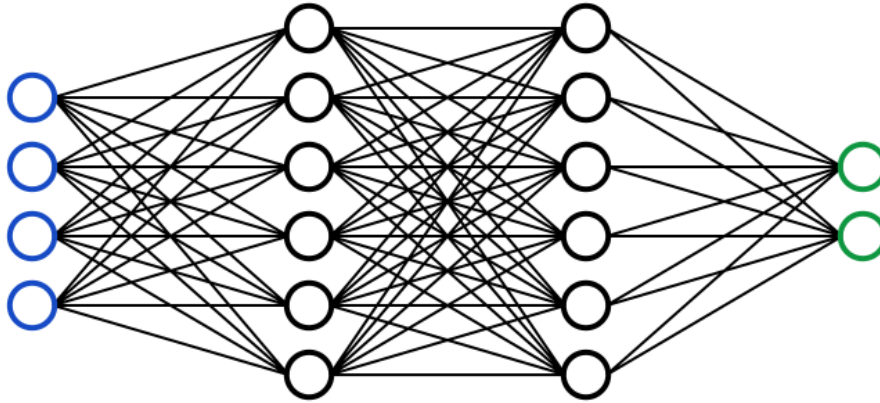


Figura 5: Estrutura da Rede Neuronal.

A Rede Neuronal é composta por camadas de nós, camada de entrada, normalmente chamada de *Input layer*, existe depois uma camada intermédia, podendo esta ser constituída por uma ou mais camadas, de seu nome *Hidden Layers*, e no final da rede temos a camada de saída, chamada de *Output Layer*. De forma a identificar a solução ótima, a rede necessita ser treinada, com o objetivo de adquirir conhecimento e melhorar a sua precisão [40].

2.1.3.1.3 Case-Based Reasoning (CBR) System

Os métodos de CBR usam uma coleção de soluções para resolver os novos problemas, e também armazenar todos os tuplos de treino, como casos. Estes são usados quando existe um novo problema a resolver, numa fase inicial o método realiza uma comparação do problema com todos os casos armazenados, se encontrar algum idêntico a solução do problema adapta-se. Caso contrário o método realiza a procura junto do dados de treino fornecidos [37].

2.1.3.2 Métodos de Segmentação (*Clustering Methods*)

Os métodos de agrupamento realizam um processo de particionamento de um determinado conjunto de dados em subconjuntos, normalmente chamados de *clusters*. Os *clusters* construídos são constituídos por dados com características idênticas, para que assim seja possível que as diferenças entre os elementos de dentro de cada *cluster* seja mínima, e as diferenças entre elementos de diferentes *clusters* seja máxima. No mesmo conjunto de dados, o uso de diferentes métodos de agrupamento, pode gerar diferentes agrupamentos, portanto o método oferece a oportunidade de descobrir agrupamentos anteriormente desconhecidos (grupos e padrões) dentro dos dados [37].

2.1.3.2.1 *K-means*

Na Figura 6, abaixo ilustrada, é representado com funciona o algoritmo K-Means.

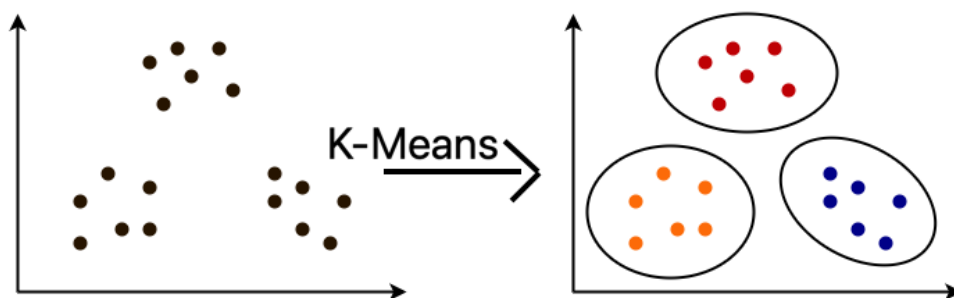


Figura 6: Exemplo do particionamento do *K-Means*.

O *K-means* é uma técnica de partição baseada em centroide ², que usa o centroide de *cluster* para representar o mesmo, normalmente considerado o centro do *cluster*. O método tem como objetivo dividir n observações em k *clusters*, cada observação pertence ao *cluster* em que a sua distância euclidiana ao centroide seja menor, concretizando o objetivo de agrupar as observações com as características mais idênticas num *cluster* [41].

²Conceito algébrico que avalia a posição da média aritmética de todas as amostras do espectro.

2.1.3.3 Algoritmos Genéticos

Os algoritmos genéticos são um método inspirado na teoria criada por *Charles Darwin*, a seleção natural. Este método segue três elementos básicos: i) a noção de aptidão de *Darwin*, visto que um indivíduo pode influenciar o futuro das gerações, ii) um "operador de acasalamento", que produz descendência para as próximas gerações, iii) "operadores genéticos", que determinam a composição genética da descendência com o material genético dos pais [35].

O processo é iniciado através da escolha aleatória de uma população de indivíduos, seguindo depois um processo iterativo, sendo que a cada interação é utilizada uma população e esta é chamada normalmente de geração. A cada interação é avaliada a aptidão de cada indivíduo, esta avaliação é baseada no valor da função objetiva do problema de otimização em causa. No indivíduo mais apto é realizada uma alteração no seu genoma, com o objetivo de formar uma nova geração de indivíduos [41].

2.2 Abordagens Preditivas

Este tipo de abordagem tem como objetivo a previsão de algo, utilizando um conhecimento obtido através de dados de histórico. Nesta secção a revisão de literatura é focada nas técnicas de ML capazes de satisfazer os requisitos deste tipo de abordagens.

2.2.1 *Machine Learning (ML)*

O Machine Learning (ML), uma aplicação de AI, é um campo de investigação com o objetivo de entender e construir modelos que através de dados que, conseguem colher conhecimento e melhorar algumas tarefas do quotidiano. Estes modelos construídos através de dados de amostra, geralmente dados de treino, conseguem realizar tarefas como previsão ou de decisão sem serem programados para o efeito. Algumas das aplicações deste métodos são usadas, na medicina, filtragem de email, previsão de datas, entre outros [41].

Anteriormente foram abordadas técnicas relacionadas a este tema, mas numa perspectiva distinta. No ML, uma das divisões dos métodos pode ser, em métodos de Classificação e de Regressão. Os métodos de Classificação, como já referido anteriormente, têm como objetivo encontrar uma classe, dentro das possibilidades, enquanto que os métodos de Regressão

têm como objetivo prever algo. Portanto, os métodos abordados neste capítulo são afiliados a métodos de Regressão.

2.2.1.1 Métodos de Regressão

Métodos de regressão é uma técnica que pode ser utilizada para duas teorias. A primeira refere que análises de regressão são normalmente utilizadas para *forecasting* e *prediction*. A segunda é que a análise de regressão pode ser utilizada em alguns casos para determinar as relações causais entre os variáveis independentes e dependentes. De referir que estes métodos mostram apenas relações entre uma variável dependente e um conjunto de dados fixos com diferentes variáveis [42]. Este tipo de métodos utiliza um conjunto de dados constituído por variáveis independentes e variáveis dependentes. As variáveis independentes são utilizadas para prever as variáveis dependentes, com o objetivo do método estimar um valor para essas variáveis dependentes [41]. Alguns dos exemplos de métodos que são abordados são Regressão Linear, *Random Forest*, *Gradient Boosting*, *Support Vector Machine (SVM)* e Redes Neurais [43].

2.2.1.1.1 Regressão Linear

A regressão linear é uma técnica de regressão paramétrica, onde a função f é calculada como uma combinação linear de características de entrada. O método permite considerar múltiplas variáveis que possam afetar a previsão da variável dependente [44]. De salientar que a abordagem alcança bons resultados quando uma função linear resulta de uma aproximação razoável da relação de dependência entre entrada e saída, ao mesmo tempo que sofre quando tal dependência é fortemente não linear [45].

2.2.1.1.2 *Random Forest*

Na Figura 7 abaixo representada é ilustrada um exemplo da estrutura do algoritmo *Random Forest*.

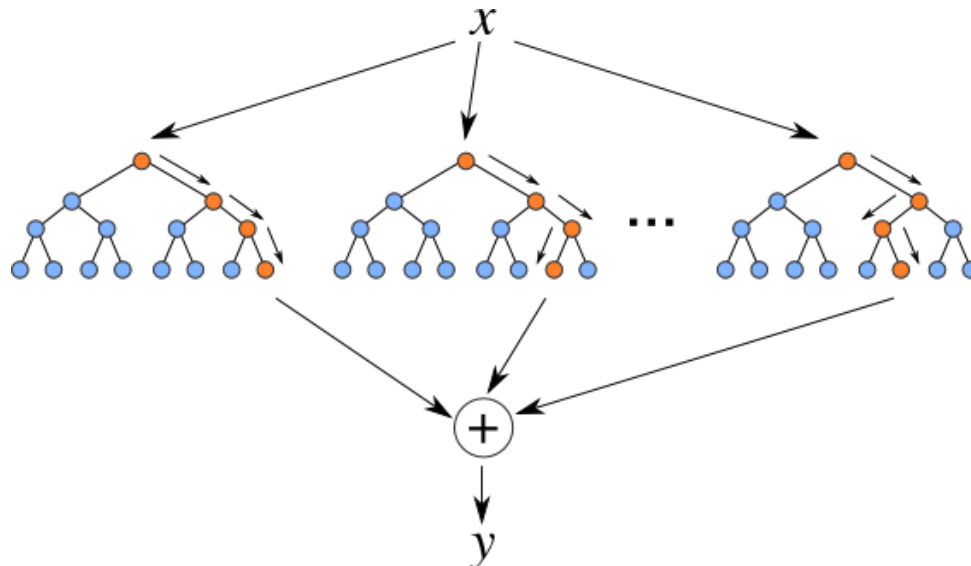


Figura 7: Estrutura Exemplo da *Random Forest*.

O método *Random Forest* é um método de Ensemble Learning³, que utiliza um conjunto de n árvores de decisão geralmente treinadas através do método de *bagging*⁴ [44]. O método pode ser realizado para tarefas de classificação ou regressão, sendo que o problema representado se trata de um problema de regressão, o resultado que será fornecido pela *Random Forest* é a previsão média ou a média das árvores individuais. A *Random Forest* fornece uma melhoria relativa às árvores de decisão, uma vez que corrigem o problema das árvores de decisão se ajustarem aos dados de treino [41].

2.2.1.1.3 *Gradient Boosting*

O *Gradient Boosting*, tal como o método anterior, é um método de *Ensemble Learning* e realiza o seu processo através da utilização de um conjunto de árvores de decisão. O método produz um modelo de previsão constituído por um conjunto de modelos de previsão fracos. O algoritmo ocorre de forma iterativa sempre com o objetivo do modelo criado, conseguir melhorar o erro do modelo construído anteriormente, com auxílio de uma função de perda [46].

³Técnica que combina um conjunto de modelos, com o objetivo de obter um melhor desempenho na previsão da solução [41].

⁴Técnica de *Ensemble Learning* utilizado para ajudar a melhorar o desempenho e precisão dos algoritmos de ML.

2.2.1.1.4 *Support Vector Machine (SVM)*

Este é um método clássico do ML, que pode ser utilizado tanto em problemas de classificação como em problemas de regressão. Com o problema de regressão em causa, o algoritmo tem como objetivo encontrar um hiperplano em um espaço n-dimensional que, classifique os pontos, normalmente chamados de vetores de suporte. Neste método a ideia principal passa sempre por minimizar o erro, individualizando um hiperplano que maximiza a margem [43].

2.2.1.1.5 *Rede Neuronal*

É um método que lida tanto com problemas de classificação como de regressão, a única diferença do método é na última camada da rede. Nos problemas de classificação pode ter n quantidade de neurônios, dependendo da quantidade de classes existentes, nos problemas de regressão apenas tem um neurônio de saída, sendo este o que dá a previsão desejada [43].

2.3 *Reflexão Crítica*

Um dos objetivos específicos deste trabalho consiste em explorar abordagens baseadas em Inteligência Artificial AI, aplicando técnicas de ML que permitam melhorar o planeamento da produção, através de análises preditivas. Neste contexto e considerando que o desafio se alicerça numa Classificação em escala pré-definida de Fornecedores/Subcontratados, os algoritmos selecionados são os que na sua génese se adequam, por definição, a este tipo de problema.

Deste modo os algoritmos a explorar no âmbito do caso de estudo são: *Random Forest*, *Gradient Boosting* e Redes Neurais (RN). A inclusão do algoritmo de RN deve-se ao facto de adicionar uma técnica que não tenha uma estrutura de análise baseada em árvores de decisão, como apresentam a *Random Forest* e a *Gradient Boosting*. Fora do domínio preditivo, o método selecionado para avaliação dos Fornecedores é o método AHP. É um método simples ideal para abordagens cujo as variáveis são fundamentalmente quantitativas num universo de dados finito e conhecido, concretamente: valores de preços, quantidades, e tempos de entrega. Com os quatro métodos selecionados será possível obter uma visão alargada sobre as diferentes perspectivas de otimização do planeamento da produção.

Capítulo 3

Conceptualização do Domínio

O conhecimento do domínio do negócio e das características da indústria, tanto a nível estrutural como a nível humano, tem um papel crucial para no desenvolvimento deste projeto. Uma vez, que através do adquirir deste conhecimento é possível primeiramente, interpretar e identificar os problemas, na fase seguinte observar o comportamento das pessoas perante os problemas, posteriormente identificar a solução para o problema, tendo em conta os processos de negócio existentes, as características do setor e os recursos, e por fim a implementação da solução no ambiente real.

Neste sentido, foi necessário realizar um estudo junto dos especialistas do domínio, com o objetivo de recolher, as característica do setor, os processos de negócio existentes e identificar e compreender a rede colaborativa envolvente.

3.1 Caracterização do Setor

A indústria do calçado é tida como um grande património de Portugal muita desta caracterização deve-se ao facto de ser uma das indústrias mais exportadoras do país, o que faz com que seja uma das que mais contribuem para a balança comercial de Portugal. Segundo a APICCAPS, no ano de 2021, o setor do calçado realizou exportações no valor de 1702 milhões USD, é o 13º país que mais exporta mundialmente e o 20º que mais produz. Já em termos de importações o valor é menor, cerca de 584 milhões USD, sendo o 34º país no mundo que mais importa [47].

O calçado português consegue ser uma grande indústria exportadora, mas também produz com muita qualidade, sendo esta qualidade reconhecida mundialmente e considerada uma das melhores do mundo, apenas ultrapassada pelo italiano. Estas características devem-se em parte ao *Know-how* das empresas produtoras, uma vez que, se trata de uma indústria já antiga que, foi adquirindo experiência ao longo do tempo, produzindo assim em grande nível de qualidade. Na maioria das empresas do setor, a sua produção é focada em *private label*¹, e alguns dos casos trabalham com apenas um cliente [48].

Apesar da qualidade do produto, as empresas ainda seguem procedimentos de trabalho muito informais, principalmente no que concerne à relação com os seus parceiros de negócio (fornecedores e subcontratados), ainda que existam processos de negócio definidos estes ainda não são seguidos por toda a cadeia produtiva. Embora os operadores tenham um grande conhecimento na questão do saber fazer, a sua baixa qualificação pode ser parte do problema anteriormente mencionado. Esta baixa qualificação não se resume apenas aos operadores, mas em alguns casos estende-se também aos administradores. Por um lado os operadores estão focados na realização da produção, e os administradores com uma visão muito reduzida tendem a não investir nas suas empresas para que, no futuro o lucro e a produtividade sejam superiores, não estando preparados para possíveis alterações às necessidades dos seus clientes [49].

Toda a *Supply Chain*, é acompanhada por características específicas como, a cada intervalo de quatro a seis meses, dá-se a mudança de coleção, o que faz com que seja necessário gerir, projetar e produzir cerca de trezentos a quatrocentos modelos diferentes para a coleção seguinte. A produção de todos os modelos é realizada de forma colaborativa, dependendo de estilistas, dos produtores de matéria-prima e produtores do produto acabado [50].

Este processo colaborativo acaba por trazer problemas de gestão, uma vez que existe uma compra de grandes quantidades de matérias-primas, devido à imensidade de modelos e variantes, conseqüentemente existe o aumento dos fornecedores de matéria-prima. Também na fase produção é frequente a subcontratação de outras empresas para determinadas etapas do processo de fabrico, ou seja, pode existir uma empresa responsável pelo molde, outra pelo corte, e a empresa inicial fica com o restante processo.

¹ Produto de Marca Própria.

Adicionalmente, o informalismo na comunicações e relações inter-empresariais, a falta de cumprimento dos processos negócio, falta de investimento ao nível dos sistemas de informação e a falta de cultura para a importância do uso dos dados no apoio à decisão, leva à inconsistência da informação, à falta de qualidade nos dados para o apoio à decisão e à inadaptação dos sistemas de informação para as técnicas otimização actuais. Sendo assim é notório a complexidade da gestão de toda a Supply Chain do setor.

3.2 Visão geral da Rede Colaborativa

3.2.1 Caracterização da Rede Colaborativa

Num contexto empresarial, como já referido, é inevitável a construção de uma rede colaborativa entre entidades. Esta rede consiste num conjunto de ligações complexas entre entidades, sejam elas internas ou externas (clientes, fornecedores, subcontratados, departamentos internos). Nestas relações circulam elementos materiais (produto acabado, matéria-prima) e imateriais (informação, dados, conhecimento, documentos), sempre de acordo com as regras estabelecidas entre as partes, com o objetivo de não colocar a subsistência das relações em causa [51]. A Figura 8 ilustrada abaixo, representa de forma sucinta como é constituída a rede colaborativa do setor e quais as relações estabelecidas.

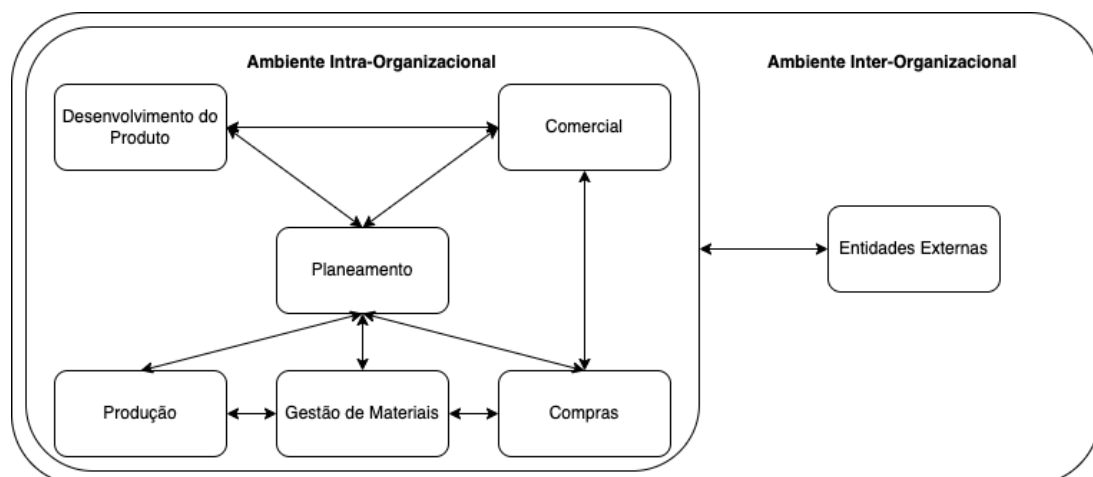


Figura 8: Rede Colaborativa do Setor do Calçado.

A rede colaborativa está dividida em dois grandes ambientes: Ambiente Intra-Organizacional e o Inter-Organizacional. O Ambiente Intra-Organizacional representa os macroprocessos internos existentes na rede colaborativa, enquanto que o Ambiente Inter-Organizacional é com-

posto pelo o Ambiente Intra-Organizacional e as Entidades Externas, intervinientes na rede. Entre estes dois ambientes existem conjuntos de relações e interações.

O Ambiente Intra-Organizacional é composto pelos seguintes macroprocessos: Desenvolvimento do Produto, cuja função é a criação de novos produtos, de acordo com inputs comerciais e as tendências da moda, que tem como objetivo realizar uma resposta rápida aos desenvolvimentos pedidos pelo cliente, de forma que seja garantida a industrialização do produto nos prazos acordados com o cliente. Este realiza uma ligação ao Comercial, onde ocorre a troca de informação relativa a comunicações realizadas pelo o cliente, com o objetivo de criar uma ficha técnica do produto.

O Comercial, tem como função a realização de toda a comunicação com os clientes, e assegurar a existência de um correto tratamento das encomendas solicitadas. Neste macroprocesso é realizada a primeira interação com as entidades externas, neste caso com os clientes, com o objetivo de receber informações relativas ao produto pedido. Uma das tarefas do Comercial passa pela comunicação das novas encomendas ao Planeamento, fornecendo as informações necessárias à atividade do Planeamento. Este tem como objetivo assegurar a produção de encomendas, de acordo com os compromissos estabelecidos, e otimizar a utilização dos recursos produtivos, sem comprometer os prazos acordados. O Planeamento executa um mapa de produção, inserindo informações sobre as encomendas a produzir, bem como quais as matérias-primas necessárias. As Compras que têm como objetivo garantir o fornecimento de materiais e componentes para a produção de encomendas com qualidade e dentro de prazos de entrega acordados, recolhe esta informação e executa a sua tarefa. Outra das interações com as entidades externas neste macroprocesso, passa pela comunicação com os fornecedores para realizar as encomendas das matérias-primas.

A Gestão de Materiais, tem como objetivo a receção dos materiais comprados, o abastecimento de materiais para a entidades subcontratadas e também a expedição do produto acabado. Este macroprocesso realiza interação constante com o Planeamento, no sentido de existir uma informação relativa ao estado da receção das matérias-primas e do envio das mesmas para os subcontratados. Por fim, a Produção tem como função a produção com qualidade e atempada do produto, e a comunicação ao Planeamento do estado da produção. De salientar, que a produção pode ser realizada internamente e/ou externamente, existindo a possibilidade da circulação de informação entre o planeamento e as entidades externas, neste caso subcontratados.

Como observado na Figura 8, o macroprocesso Planeamento está localizado ao centro, devido à sua enorme importância na rede colaborativa. Das várias interações identificadas, são destacadas algumas relevantes para o processo de Planeamento, sendo elas: a interação com o macroprocesso Comercial, uma vez que são fornecidas as encomendas e os compromissos realizados com os clientes, dando assim início ao processo de planeamento. De seguida na interação com as Compras, com o objetivo de priorizar e organizar as entregas dos materiais, mediante as datas de produção planeadas. A interação com a Gestão de Materiais com o intuito de gerir envios atempados para os produtores, evitando atrasos. A interação com a Produção realizando a passagem de informações técnicas relativas ao pedido do cliente, negociar datas para a produção e receber informação relativas ao estado da produção [52]. Com a existência de várias relações e tarefas na rede colaborativa é necessário a sistematização do processo de negócio envolvente. Na Figura 9, abaixo ilustrada, é representado o processo de negócio associado.

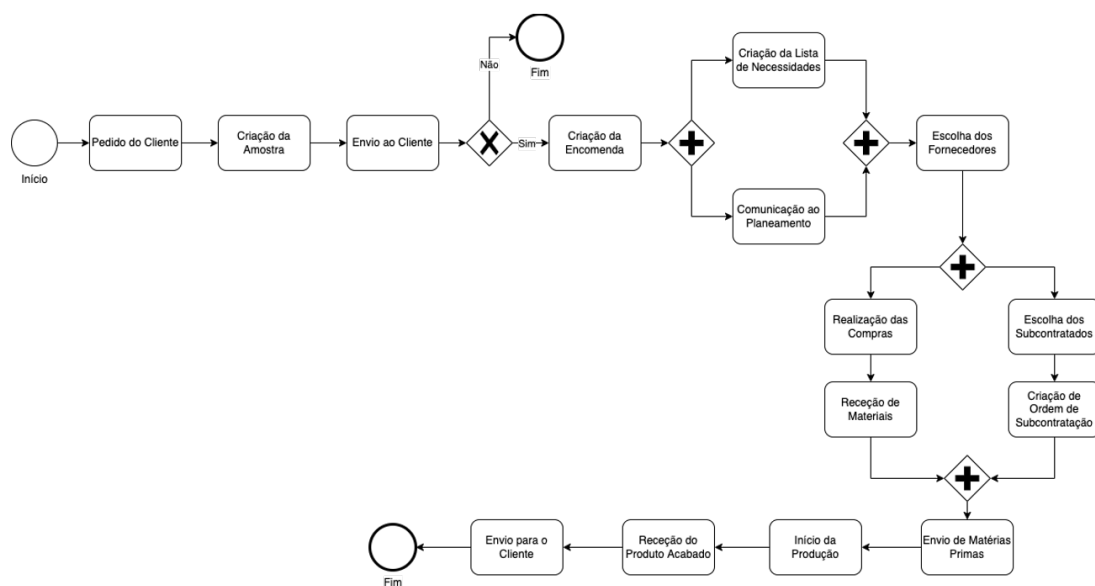


Figura 9: Processo de Negócio do Setor do Calçado.

Como observado na Figura 9, o processo é desencadeado pelo cliente, uma vez que neste setor grande parte das vezes apenas existe produção com um pedido em concreto. Nesta interação inicial, o cliente faculta informações sobre o produto que quer e quantidades, com o objetivo de receber um orçamento. Numa fase seguinte, o departamento de desenvolvimento da empresa elabora uma ficha técnica com todos os componentes necessários, enquadrados com os inputs do comercial, e também o valor de fabrico da peça. Com esta ficha técnica o departamento produz uma amostra, e envia-a juntamente com o orçamento.

Este processo pode ser repetido, não na sua totalidade, mas algumas das fases, devido a mudanças na amostra, mudanças de componentes ou negociação de preço. Ultrapassado este processo, a encomenda apenas avança com a resposta positiva do cliente, no caso de não existir o processo termina. Em caso de aceitação, é iniciado o processo de criação de encomenda, com dados específicos do produto, alguns deles são:

- Cor;
- Variantes;
- Tamanhos;
- Destinos;
- Quantidade;
- Data de Entrega;
- Preço;

Após a criação são realizadas duas tarefas em paralelo, o departamento comercial comunica ao planeamento a nova encomenda, e constrói uma lista de necessidades para fornecer às compras. O departamento de compras, com a lista das necessidades, realiza uma consulta dos fornecedores, com o objetivo de realizar uma escolha de quais utilizar. Nesta consulta é questionado ao fornecedor a possibilidade de fornecer a quantidade necessária, qual o orçamento e uma previsão de tempo de entrega.

Com a consulta concluída são realizadas as requisiões de compra da matéria-prima. Num processo em paralelo, o planeamento com a necessidade de gerir os seus recursos produtivos e ao mesmo tempo cumprir os prazos de entrega estabelecidos com o cliente, avalia quais os subcontratados a utilizar e quais as operações que eles vão realizar. Com o plano de produção idealizado são criadas as ordens de subcontratação, sendo assim, o armazém assim que rececionar as matérias-primas, submete o envio das quantidades necessárias para os subcontratados.

Todo o processo entre a compra e a receção da matéria-prima é seguido pelo o departamento de compras, sendo que estes é que realizam qualquer comunicação com os fornecedores, como por exemplo, questionar a falta de algum material, informar a necessidade de

devolução de material com defeito. Aquando da receção o armazém associa estas entradas às requisições efetuadas, no intuito de ir abatendo a quantidade pedida, uma vez que as entregas podem ser faseadas e não na totalidade. Reunida toda a matéria-prima é então enviada para o subcontratado, para este dê início à produção. O processo seguinte dá início à transformação da matéria-prima em produto acabado os subcontratados, durante o processo fornecem informação relativa ao estado da produção, como por exemplo:

- Número de Trabalhadores Diário;
- Quantidade de peças cortadas;
- Quantidade de pele desperdiçada;
- Produção Diária;
- Tempos por operação;

Finalizada a produção é efetuado o envio para a fábrica principal, o armazém realiza a entrada e verificação do produto acabado. Este organiza os envios para o cliente, seguindo os acordos estabelecidos com o cliente, e é terminado o processo. O planeamento segue todo este processo desde que recebeu informação da nova encomenda, com o objetivo de estar a par de algum problema que surja, solucionando o mais célere possível e também para fornecer toda a informação necessária ao comercial, e este, por sua vez, vai informando o cliente.

Capítulo 4

Caso de Uso

Devido à célere mudança do mercado do setor, e aliada a esta característica a crise económico-social provocada pela pandemia e as sucessivas disrupções nas cadeias de valor em várias empresas, a necessidade das empresas melhorarem os seus processos de planeamento e a gestão da rede colaborativa, surge como um problema emergente. Nomeadamente, no que toca a questões sobre a estrutura da rede colaborativa, identificação de padrões para as interações com as entidades externas, e utilização dos dados para apoio à decisão.

Nesse sentido, neste capítulo é proposto explorar algumas técnicas para otimização dos processos, sendo estas focadas no problema da seleção dos fornecedores e na previsão temporal relativa aos processos executados pelos fornecedores. O desenvolvimento de algoritmos baseados nestas técnicas permite-nos fornecer algumas funcionalidades, nomeadamente:

- A recolha e utilização dos dados, disponíveis nos vários dispositivos físicos e virtuais (sistemas de informação, dispositivos *Internet of Things (IoT)* no apoio à decisão;
- Realizar um plano de produção baseado num algoritmo de previsão temporal;
- Identificação e alerta das falhas nas entregas;

As técnicas utilizadas para a possível solução dos problemas, estão descritas no capítulo 2. O contexto específico do caso de uso e a descrição do uso das técnicas, é descrita na secção seguinte.

4.1 Descrição do Problema

Como já referido anteriormente, o mercado cada vez está mais dinâmico e exigente para com as empresas. Isto significa que as empresas têm que se adaptar para ganharem vantagem competitiva e conseguirem os melhores negócios, para isso necessitam apostar em novas abordagens, no sentido de conseguir melhorar todo o seu processo. O projecto CNI4.0-Foot¹, conta com a participação do INESC TEC², tem como objetivo promover a investigação e o desenvolvimento de soluções baseadas em tecnologias de informação que englobem a rede colaborativa das empresas do setor do calçado, envolvendo as atividades da cadeia de valor e fornecimento de forma a consolidar e promover a inovação nos processos de negócio [53]. Neste projeto são apresentados cinco objetivos:

- O **Primeiro Objetivo**, consiste na criação de um Sistema de Geração Automática de Informação Técnica do Produto e Inventário de Materiais, com o objetivo de automatizar a recolha e geração da informação técnica de produtos e de materiais;
- O **Segundo Objetivo**, corresponde à criação de um Sistema de Avaliação de Fornecedores e Subcontratados, no sentido da seleção correcta dos fornecedores, sendo assim a função do sistema é a análise e classificação automáticas dos Fornecedores e Subcontratados;
- O **Terceiro Objetivo**, construção de um Sistema Dinâmico de Planeamento da Rede Colaborativa, com o objetivo de desenvolver algoritmos de planeamento considerando informação em tempo real, recolhida do chão de fábrica e de uma plataforma colaborativa da rede;
- O **Quarto Objetivo**, consiste no desenvolvimento de uma Sistema de Acompanhamento e Alerta para o Planeamento, no sentido de fornecer ao planeador uma previsão dos tempos de entrega das matérias-primas e a realização de alertas quando existe a falha nas entregas de matéria-prima necessária para as operações de fabrico;
- O **Quinto Objetivo**, inovação de uma nova ferramenta e conceitos de calçado "ecológico e sustentável" com redução da Pegada Ambiental, aspeto inovador da ferramenta *SMART ECO SHOES* foca-se na criação de uma solução que suporte o departamento de design/concepção

¹<https://www.inesctec.pt/pt/projetos/cni40foot>

²<https://www.inesctec.pt/>

de calçado na seleção dos materiais e desenvolvimento do produto (forma/dimensões/área, peso, sistema construtivo, entre outros) com base em conhecimento estruturado de ACV.

Estas novas ferramentas são integradas juntamente com as já existentes na Figura 10, abaixo identificada, é ilustrada a relação entre elas.

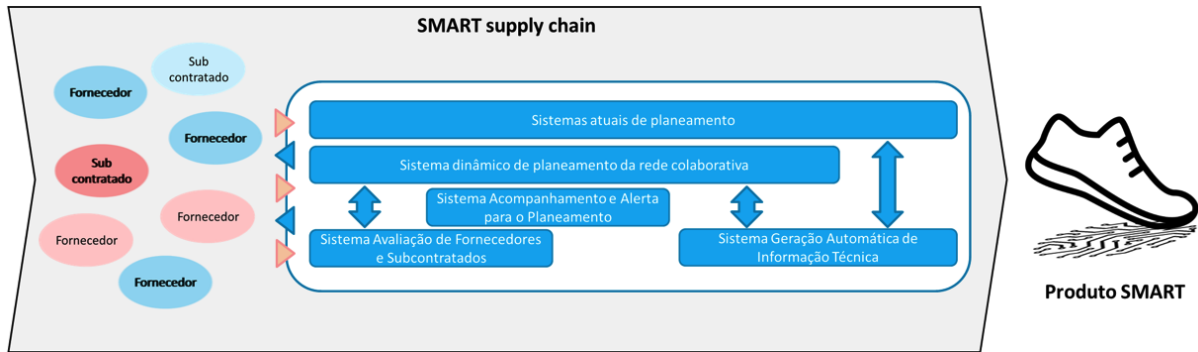


Figura 10: SMART Supply Chain

O caso de uso para a validação dos cenários propostos está relacionado com o segundo e quarto objetivo do projeto CNI4.0-Foot. Para o primeiro objetivo proposto foi criada uma arquitectura, sendo demonstrada a ligação entre os vários componentes até ao resultado final, na Figura 11, abaixo representada.

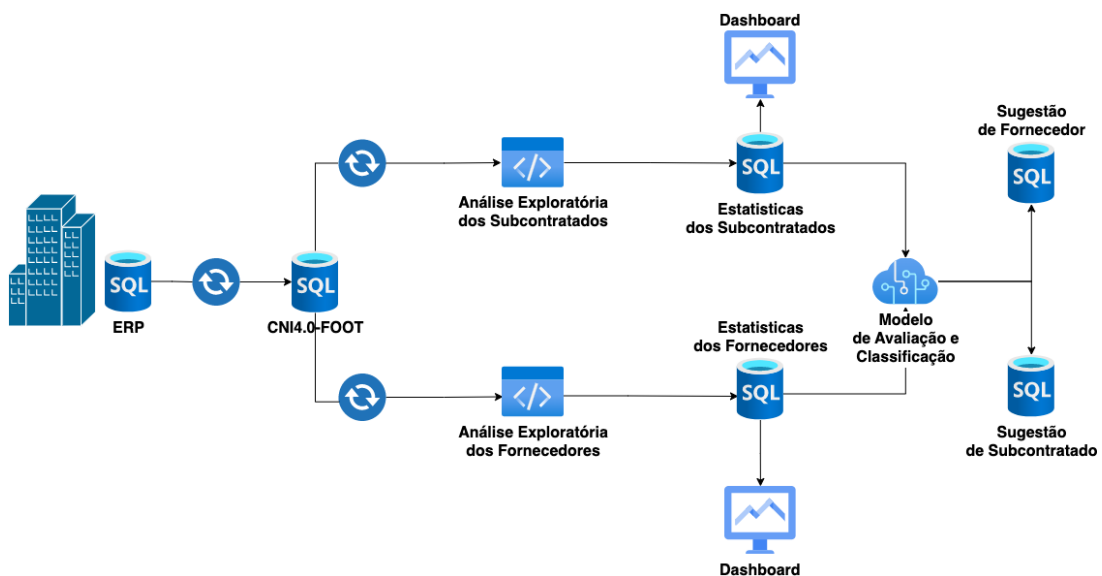


Figura 11: Arquitetura do Sistema de Avaliação de Fornecedores e Subcontratados.

A partir da Figura 11, pode-se observar todo o fluxo existente no sistema. Existe uma recolha de dados de histórico do sistema de informação associado ao projeto, numa fase seguinte, é realizada uma análise exploratória aos dados utilizando alguns métodos estatísticos. Estes dados estatísticos são disponibilizados aos utilizadores numa plataforma, com o objetivo de avaliarem a prestação dos fornecedores em determinados pontos, e por fim os dados existentes são submetidos a um algoritmo que vai classificar os fornecedores em causa, tendo em conta determinados critérios.

Para o outro objetivo, foi igualmente criada uma proposta de arquitetura, esta está ilustrada na Figura 12, abaixo representada.

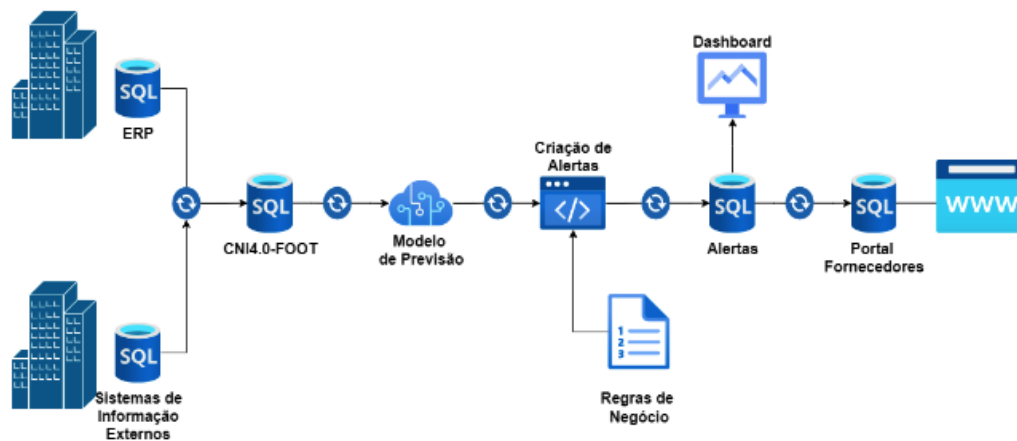


Figura 12: Arquitetura do Sistema de Acompanhamento e Alerta para o Planeamento

Na Figura 12, é possível observar os componentes integrantes do sistema e as suas ligações. Esta arquitetura é sustentada em dados, sendo assim é necessária uma recolha contínua de dados, vindos, por exemplo de dispositivos do chão da fábrica, sistemas de informação, plataformas colaborativas, entre outros. Com estes dados recolhidos, é necessário um tratamento antes da sua submissão ao algoritmo de previsão. Um exemplo do tipo de dados submetidos ao algoritmo de previsão, é ordem de fabrico, com o objetivo de se prever as datas de entregas das matérias-primas e consequentemente identificar uma possível data de início de produção.

Através destes dados de previsão são criados alertas para o planeador, com o intuito de identificar materiais que ainda não chegaram, e de acordo com a data de previsão calculada, pode ser uma matéria-prima que atrase o início da produção, levando a um atraso na entrega da encomenda. Estes alertas são criados de acordo com regras de negócio definidas pelos es-

pecialistas de domínio. Tanto os alertas como as previsões calculadas estão disponíveis no portal interno como também numa plataforma colaborativa com as entidades externas.

Importante salientar, e indo de encontro ao objetivo proposto, o projeto visa apenas o desenvolvimento até ao componente do algoritmo de previsão, a restante arquitetura será elaborada no contexto do projeto CNi4.0-Foot, sendo também um trabalho para futuro no que diz respeito ao tema abordado neste projeto.

4.2 Abordagens Desenvolvidas

Identificada a arquitetura é necessária a implementação de uma abordagem para o desenvolvimento da solução. Neste capítulo aborda em termos práticos as abordagens escolhidas e quais as tarefas realizadas até à obtenção dos resultados.

4.2.1 Avaliação de Fornecedores e Subcontratados baseada em AHP

A abordagem escolhida para desenvolvimento da solução de Avaliação dos Fornecedores foi o método *Analytic Hierarchy Process* AHP. Para a construção do método é necessária a realização de certas etapas, sendo isso ilustrado abaixo na Figura 13

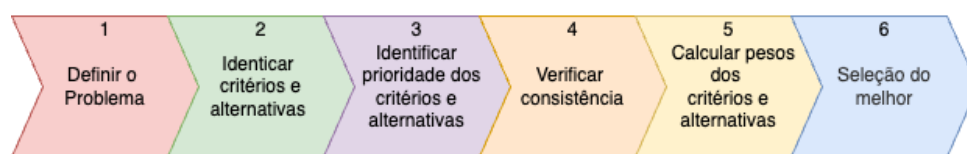


Figura 13: Processo de Construção da AHP.

Como em todos os algoritmos de decisão o processo inicia-se sempre com a definição do problema, no sentido de conhecer o domínio em causa e suas características. Numa fase seguinte, já com o conhecimento do negócio, a etapa de identificação dos critérios e das alternativas, os responsáveis nesta etapa são normalmente os especialistas do domínio, com o objetivo de identificarem numa perspectiva do que é necessário otimizar. A identificação das prioridades dos critérios e alternativas é realizada tendo por base a Escala de Saaty [54]. Esta vai de 1 a 9, com o objetivo numa comparação entre dois critérios ou alternativas, atribuir uma

importância em relação ao outro. O valor 1 tem como significado a indiferença de importância entre os critérios, enquanto o 9 atribui uma extrema importância de um critério sobre o outro [11]. Na Tabela 1 abaixo ilustrada são identificados os vários níveis de importância e uma breve explicação [55].

Intensidade de Importância	Definição	Explicação
1	Mesma Importância	As duas atividades contribuem igualmente para o objetivo.
3	Importância pequena de uma sobre a outra.	A experiência e o julgamento favorecem levemente uma atividade em relação à outra.
5	Importância grande ou essencial.	A experiência e o julgamento favorecem fortemente uma atividade em relação à outra.
7	Importância muito grande ou demonstrada.	Uma atividade é muito favorecida em relação à outra;
9	Importância absoluta.	A evidência favorece uma atividade em relação à outra com o mais alto grau de certeza.
2,4,6	Valores intermediários entre os valores adjacentes	Quando se procura uma condição de compromisso entre duas definições.
Recíprocos dos valores acima de zero	Se a atividade i recebe uma das designações diferentes acima de zero, quando comparada com a atividade j, então j tem o valor recíproco quando comparada com i.	Uma designação razoável.
Racionais	Razões resultantes da escala.	Se a consistência tiver de ser forçada para obter valores numéricos n, somente para completar a matriz .

Tabela 1: Escala de Saaty.

Realizadas as comparações entre critérios e alternativas, obtém-se uma matriz de comparação. No sentido de avaliar a qualidade e a consistência das matrizes *Saaty* são utilizados dois cálculos, abaixo ilustrados na Figura 14 e na Figura 15.

$$C.I. = \frac{\lambda_{\max} - n}{n - 1}$$

Figura 14: Índice de Consistência)

$$CR = \frac{CI}{RI}$$

Figura 15: Rácio de Consistência

De forma a matriz ser consistente, o valor obtido através do cálculo tem de inferior a 0.1, caso não seja é necessário refazer as prioridades definidas [14]. A seguinte tarefa, após o concretizar da consistência da matriz, é calcular os pesos dos critérios e alternativas inicialmente é necessário normalizar a matriz, com a divisão da importância de cada célula pela soma das importâncias da coluna. O valor do peso é obtido através da média da linha da matriz. Através da multiplicação dos pesos das alternativas em relação a cada critério e o peso de cada critério, é encontrada a prioridade da alternativa, sendo assim a alternativa com o maior valor é a selecionada.

Com base na revisão de literatura e na arquitetura desenvolvida no âmbito do projeto CNI4.0-Foot, foi possível chegar-se a uma solução final e cumprir-se os objetivos propostos para a Classificação e Avaliação de Fornecedores. Todo o trabalho desenvolvido foi orientado pelas etapas anteriormente detalhadas. O primeiro passo na utilização do método AHP é a selecção de critérios que irão apoiar o processo de tomada de decisão realizado pelo algoritmo. Para seleccionar os critérios, foi necessário intervir com peritos do domínio para compreender quais os critérios mais importantes na escolha de um fornecedor para a compra de matérias-primas, e também analisar dentro dos dados disponíveis a possibilidade de utilização desses mesmos critérios. A Figura 16 ilustrada abaixo, demonstra os critérios seleccionados.

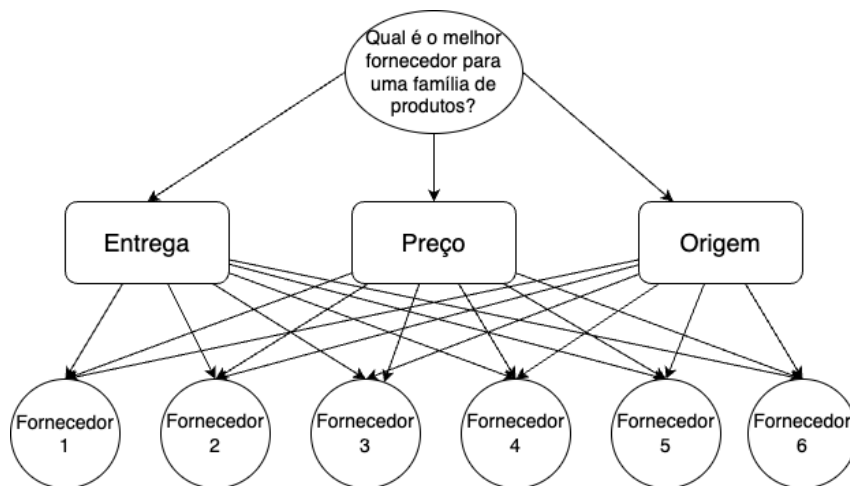


Figura 16: Estrutura da AHP baseada no problema.

Para o desenvolvimento da solução é imprescindível a recolha dos dados, a sua utilização é essencial em todas as etapas do processo. A recolha de dados é realizada diariamente, uma vez que os dados estão em constante inserção e alteração, e essas mudanças podem ser refletidas no resultado do algoritmo. Os dados recolhidos refletem todas as compras com receções dos últimos quatros anos da empresa, este período foi definido com os parceiros do projeto, no sentido de poder avaliar a performance do fornecedor nos diferentes periodos e também existir uma quantidade razoável de dados para a execução do algoritmo. Abaixo, na Tabela 2 é ilustrado o dataset utilizado pelo o algoritmo, tendo em conta os dados existentes e os critérios escolhidos.

Atributo	Descrição	Tipo
NR FORNECEDOR	Número Identificativo do Fornecedor.	<i>Integer</i>
NOME FORNECEDOR	Nome do Fornecedor	<i>String</i>
COD FAMILIA	Número Identificativo da Família de Materiais associado ao Fornecedor.	<i>Integer</i>
NOME FAMILIA	Nome da Família de Materiais associado ao Fornecedor.	<i>String</i>
TEMPO ENTREGA	Valor médio de tempo de entrega, tendo em conta Fornecedor e Família.	<i>Float</i>
PRECO	Valor médio do preço, tendo em conta Fornecedor e Família.	<i>Float</i>
PAIS	País do Fornecedor	<i>String</i>

Tabela 2: Caracterização do dataset utilizado no algoritmo AHP.

Analisando a tabela 2 e a figura 16, o objetivo do algoritmo passa pela escolha de fornecedores em relação a uma família de artigos e não a um artigo específico. Isto deve-se a uma das características do sector, uma vez que na maioria das vezes existem fornecedores únicos para o artigo, e também da informação recolhida junto dos especialistas, pois a empresa analisada no contexto do projeto, realiza a divisão dos fornecedores por família de materiais. Existiu uma análise às famílias, no sentido de observar os artigos associados e perceber se existia uma grande discrepância dos dados entre eles. Feita essa avaliação, as famílias contêm artigos apenas com pequenas diferenças sendo assim não existe uma diferença abrupta sobre as características avaliadas na AHP.

Para classificar a importância atribuída a cada critério, foi utilizada a escala de importância *Saaty*. Como já referido anteriormente, o objetivo da escala é identificar a importância de um critério em relação a outro. A tabela 3 ilustra a importância e o peso de cada critério. De notar que esta priorização foi identificada em conjunto com os especialistas do domínio.

Critério	Entrega	Preço	Origem	Peso
Entrega	1	4	8	0.69862
Preço	1/4	1	5	0.23704
Origem	1/8	1/5	1	0.06434

Tabela 3: Matriz de Importância.

Com a matriz construída, foi utilizada a linguagem *Python* e uma biblioteca dedicada ao desenvolvimento do método AHP [56]. No exerto de código 1 é ilustrada a inserção da matriz de importância no algoritmo desenvolvido.

```

1 import ahpy
2 data = "SELECT fa.cliente NR_FORNECEDOR, f.NOME FORNECEDOR, fa.familia COD_FAMILIA,
        f.DESCRICAO FAMILIA, media_temp_entrega TEMPO_ENTREGA, preco_medio PRECO, PAIS "
3 "FROM stats_fornecedores_familia as fa join Fornecedores as f on f.CLIENTE=fa.
        cliente and f.FAMILIA=fa.familia "
4 "WHERE fa.familia=' " + family + "' ORDER BY NR_FORNECEDOR, COD_FAMILIA "
5
6     response = connection.execute(data)
7     df = pd.DataFrame(response.fetchall())
8     df.columns = response.keys()
9
10    criteria_comparisons = (('Delivery', 'Price'): 4, ('Delivery', 'Origin'): 8, ('

```

```

11     Price', 'Origin'): 8}
12     criteria = ahpy.Compare('Criteria', criteria_comparisons, precision=5)
13     report = criteria.report(show=True)

```

Exemplos de Código 1: Inserção da matriz de importância no algoritmo.

Através desta biblioteca é possível também realizar o cálculo do Rácio da Consistência da matriz, como já referido anteriormente, é necessário que a matriz seja consistente para o sucesso dos resultados. O valor obtido é de 0,0956. No excerto de código 2 é ilustrado o valor obtido através do algoritmo, bem como os pesos da matriz de importância.

```

1 {
2     "name": "Criteria",
3     "global_weight": 1.0,
4     "local_weight": 1.0,
5     "target_weights": {
6         "Delivery": 0.69862,
7         "Price": 0.23704,
8         "Origin": 0.06434
9     },
10    "elements": {
11        "global_weights": {
12            "Delivery": 0.69862,
13            "Price": 0.23704,
14            "Origin": 0.06434
15        },
16        "local_weights": {
17            "Delivery": 0.69862,
18            "Price": 0.23704,
19            "Origin": 0.06434
20        },
21        "consistency_ratio": 0.09566
22    }
23 }

```

Exemplos de Código 2: Exemplo do modelo de dados do relatório do algoritmo.

Existe a necessidade da criação de uma matriz de importância para cada critério, no sentido de realizar a comparação de todas as alternativas em detrimento dos critérios. Para isso, utilizando a escala de Saaty, foi desenvolvido um algoritmo para automatizar a importância de um fornecedor em relação a outro. Este automatismo foi realizado através da obtenção do conjunto de diferenças entre os valores de entradas dos fornecedores. Dentro do conjunto foi obtido o valor máximo, o mínimo, a média, o primeiro quartil e o terceiro quartil, no sentido de dividir a escala em quatro intervalos. Na Tabela 4, é ilustrada a divisão da tabela.

1	2	3	4	5	6	7	8	9
0		2		10,60		22		27
Mínimo		1ºQ		Média		3ºQ		Máximo

Tabela 4: Exemplo da divisão da escala de Saaty no critério de Entrega.

Os valores apresentados são relativos a um conjunto de fornecedores de uma determinada família, escolhida para demonstração no projeto. Os restantes exemplos são referenciados em anexo. Para automatizar o processo de comparação entre fornecedores, foi criado um algoritmo em *Python*, com a tarefa de realizar a divisão da escala de Saaty, tendo em conta o conjunto das diferenças e construir a matriz de comparação entre os vários fornecedores. No excerto de código 3, é ilustrado um exemplo desse algoritmo.

```

1 def define_values_delivery(data, comparacions):
2     data['TEMPO.ENTREGA.FLOAT'] = data['TEMPO.ENTREGA'].astype(float)
3     difValues = calc_diff(data, comparacions)
4
5     min = np.min(difValues)
6     first_quartil = np.quantile(difValues, 0.25)
7     mean = np.mean(difValues)
8     ter_quartil = np.quantile(difValues, 0.75)
9     max = np.max(difValues)
10
11     valuesDelivery= build_comparison_matrix(data, comparacions, min, first_quartil,
12     mean, ter_quartil, max)
13
14     data.drop('TEMPO.ENTREGA.FLOAT', axis=1)
15     return valuesDelivery

```

Exemplos de Código 3: Função para construção da matriz de importância no critério Entrega.

Obtida a matriz de importância dos critérios e após a normalização, é realizado o cálculo para a obtenção do peso do fornecedor em cada critério. No exemplo de código 4, é demonstrado como é realizado o cálculo com auxílio à biblioteca *ahpy* [56].

```

1 import ahpy
2
3 delivery_values = define_values_delivery(df, suppliers_pairs)
4 delivery_comparisons = dict(zip(suppliers_pairs, delivery_values))
5 print("\n", delivery_comparisons)
6
7 delivery = ahpy.Compare('Delivery', delivery_comparisons, precision=5)
8 reportDelivery = delivery.report(show=True)

```

Exemplos de Código 4: Utilização da biblioteca *ahpy* para cálculo do peso.

O resultado obtido da utilização do exemplo de código 1 e do exemplo de código 4 é demonstrado através da Tabela 5, abaixo ilustrada.

Fornecedor	30	46	256	650	698	1195	Peso
0030	1	9	3	1	3	4	0.308
0046	1/9	1	1/8	1/9	1/8	1/7	0.022
0256	1/3	8	1	1/3	1	4	0.148
0650	1	9	3	1	3	4	0.308
0698	1/3	8	1	1/3	1	4	0.148
1195	1/4	6	1/4	1/4	1/4	1	0.067

Tabela 5: Exemplo da matriz de importância do critério Entrega.

Após a construção de todas estas matrizes para cada critério, está reunido o necessário para a obtenção do resultado final. Os resultados finais da implementação do algoritmo são demonstrados no capítulo seguinte.

4.2.2 Abordagem de *Machine Learning* na Previsão do Plano de Produção

Neste objetivo proposto foi utilizado um algoritmo de *Machine Learning*, no sentido de realizar uma tarefa preditiva no apoio à construção do Plano de Produção, na Figura 17, abaixo ilustrada, está representado o processo necessário para a construção do algoritmo.

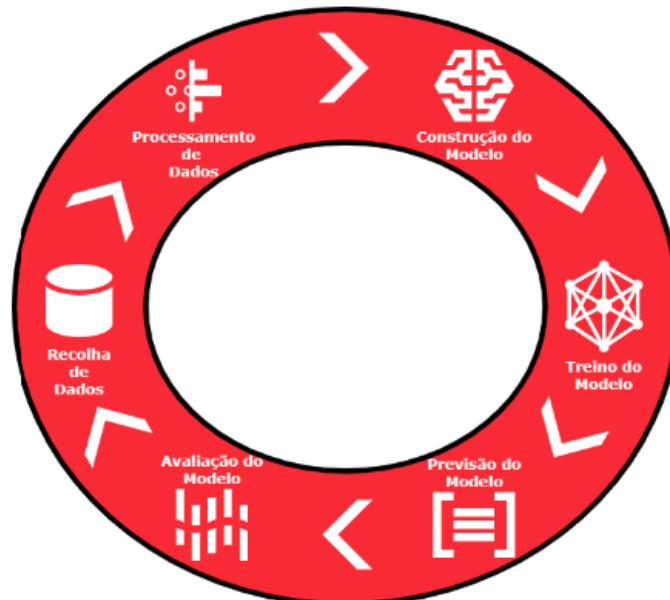


Figura 17: Processo de Construção de um algoritmo de *Machine Learning*

O primeiro passo envolve a compreensão do problema, nomeadamente:

- Qual é o alvo do problema?;
- Qual o tipo de problema? Classificação ou Regressão?;
- Que possíveis variáveis a ser utilizadas pelo modelo, no sentido de fornecer os conhecimentos necessários para obter uma boa solução.

A resposta a estas questões é auxiliada pelos especialistas do domínio. Recolhido o conjunto de respostas, é necessária a recolha dos dados das fontes existentes, no sentido de analisar, organizar e estruturar o *dataset* a ser utilizado pelo algoritmo. Esta etapa envolve uma análise profunda dos dados, com a utilização de métodos estatísticos, gráficos de distribuição de valores, entre outros. Através deste estudo do *dataset* recolhido, é possível observar a quantidade de dados existentes, número de variáveis, falhas de alguns dados, discrepâncias de alguns valores, tendo em conta os valores médios e desvios.

Com a análise, é necessário intervir nos dados e realizar a limpeza necessária nos dados, no sentido de se ficar com o *dataset* final. As tarefas de limpeza dos dados, podem ser, por

exemplo, a limpeza de valores que não façam sentido no domínio do negócio, o preenchimento de valores, tendo em conta os valores médios da variável. Outra das técnicas utilizada para esta estruturação dos dados, é a normalização dos dados com o objetivo de evitar variações abruptas nas variáveis independentes do problema, este método é normalmente identificado por *Feature Scaling* [44].

Realizadas todas as tarefas de estruturação do dataset, existe uma separação do *dataset*, que consiste em dividi-lo em dois conjuntos de dados, um de treino e um de teste. O *dataset* de treino tem como objetivo o fornecimento de conhecimento ao algoritmo o *dataset* de teste é utilizado, no sentido de avaliar a qualidade do conhecimento fornecido pelo *dataset* de treino e conseqüentemente o desempenho do modelo. [41].

Com o *dataset* repartido, é efetuado o treino do modelo e para avaliar a qualidade dos resultados previstos pelo algoritmo no *dataset* de teste são utilizadas algumas métricas. Essas métricas podem diferir dependendo do tipo de problema, para um caso de regressão são normalmente utilizados *Root Mean Squared Error (Root Mean Squared Error (RMSE))*, *Mean Squared Error (Mean Squared Error (MSE))* e *Mean Absolute Error (Mean Absolute Error (MAE))*, para o caso de classificação é utilizado *Accuracy*, *Precision* e *Recall* [41]. Para o problema em questão são utilizadas as métricas de regressão, estas são ilustradas nas figuras abaixo.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}|$$

Figura 18: *Mean Absolute Error (MAE)*

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2$$

Figura 19: *Mean Squared Error (MSE)*

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2}$$

Figura 20: *Root Mean Squared Error (RMSE)*

Implementada e avaliada a qualidade do modelo, este pode ser utilizado em ambiente real, no caso dos resultados irem de encontro ao esperado, caso contrário, é necessário realizar novamente todo o processo, uma vez que, a falta de variáveis pode não trazer conhecimento suficiente ao modelo, tratando-se de um dos problemas da qualidade do modelo. Dependendo do problema em que estão, estas tarefas podem ser cíclicas, uma vez que os dados serão incrementados ao longo do tempo, existindo a necessidade de re-treinar o modelo. O problema em causa é um exemplo disso. Os modelos abordados neste projeto são *Random Forest* [57], *Gradient Boosting* [41] e *Redes Neurais* [41].

O fluxo do trabalho desenvolvido no contexto desta abordagem, foi de encontro às tarefas abordadas anteriormente. A primeira tarefa foi a discussão do problema com os parceiros do projeto e especialistas do domínio, no sentido de reunir o objetivo da previsão e quais as possíveis variáveis necessárias para a obtenção do resultado pretendido. Para construir o conjunto de dados do problema houve a necessidade de realizar uma recolha de dados, sendo esta, realizada através de uma interação com as fontes de informação (sistemas de informação, plataformas colaborativas, máquinas industriais) de uma empresa do setor. Foram recolhidos vários dados, tais como, Ordens de Compra, Receções, Artigos, Famílias de Artigos e Fornecedores. Os dados recolhidos representam os últimos quatro anos de compras da empresa tal como a abordagem anterior, esta decisão foi tomada em conjunto com os especialistas do domínio, uma vez que, existe a necessidade agrupar uma quantidade de dados suficiente para passar conhecimento ao algoritmo e este facultar uma solução viável. Através deste conjunto de dados foram selecionadas as variáveis para a construção do dataset, indo de encontro ao já analisado com os especialistas do domínio. Na Tabela 6 é representada a estrutura do *dataset* do problema.

Atributo	Descrição	Tipo
ARTIGO	Número identificativo do Artigo comprado.	<i>Integer</i>
FORNECEDOR	Número Fornecedor do Artigo.	<i>Integer</i>
FAMILIA	Número identificativo da Família de Materiais associado ao Artigo.	<i>Integer</i>
ANO	Ano da compra do Artigo.	<i>Integer</i>
MÊS	Mês da compra do Artigo.	<i>Integer</i>
SEMANA	Semana da compra do Artigo.	<i>Integer</i>
NR COMPRAS	Número de compras ao fornecedor, até ao momento, no mês da compra.	<i>Integer</i>
QUANTIDADE	Quantidade de compra do artigo.	<i>Decimal</i>
PAIS	País de origem do artigo.	<i>Enum</i>
TEMPO ENTREGA (Variável Dependente)	Tempo de entrega da compra.	<i>Integer</i>

Tabela 6: Características do dataset utilizado no algoritmo de *Machine Learning*.

Para este problema a variável TEMPO ENTREGA é a variável dependente, visto que é o valor que vai ser previsto pelo algoritmo. O TEMPO ENTREGA, é resultado da diferença entre a data de compra e a data da última receção, uma vez que a entrega de matérias-primas pode ter várias receções, mas apenas na última está reunido o material requisitado. O ARTIGO, FORNECEDOR e FAMILIA são as variáveis identificativas do problema, ou seja, identificam o artigo comprado, qual a família de materiais que artigo pertence e qual o fornecedor que foi requisitado para a compra. As variáveis ANO, MÊS, SEMANA e NR COMPRAS, têm relevância para o sucesso da previsão, devido à sazonalidade existente na produção do calçado. Existem meses em que a ocorrência de compras de matérias-primas é elevada relativamente a outros, conseqüentemente os fornecedores têm uma elevada carga de requisições, havendo assim a possibilidade de aumentar o tempo de entrega. A seleção da quantidade da compra influencia igualmente o tempo de entrega, uma vez que, quanto maior a quantidade requerida maior será o seu tempo de produção, conseqüentemente, o tempo de entrega aumenta.

O dataset foi separado em dois, um *dataset* de treino e um *dataset* de teste. O *dataset* de treino é composto por todas as compras efetuadas dos anos de 2019, 2020 e 2021, tendo um tamanho de 34007 registos, que refletem a existência de 3424 artigos distintos. O conjunto de dados de teste é composto por todas as compras desde 2022 até à data, consistindo em 1832 registos. A utilização de algumas medidas estatísticas fornecem uma análise da distribuição dos valores da variáveis quantitativas do dataset, para isso foi utilizada a linguagem *Python* e algumas bibliotecas dedicadas ao desenvolvimento deste tipo de abordagens. No excerto de código 5, é possível observar a utilização dessas bibliotecas na obtenção da análise ao *dataset*.

```
1 import pandas as pd
2 datasetTreino = "SELECT ID ARTIGO, FAMILIA, FORNECEDOR, ..."
3                 "Where ANO<=2021"
4 datasetTeste = "SELECT ID ARTIGO, FAMILIA, FORNECEDOR, ..."
5                 "Where ANO>=2022"
6
7 responseTreino = connection.execute(datasetTreino)
8 df = pd.DataFrame(responseTreino.fetchall())
9 responseTeste = connection.execute(datasetTeste)
10 dfTeste = pd.DataFrame(responseTeste.fetchall())
11 #MAPA DE CORRELA O
12 print(df.corr())
13 #AN LISE EXPLORATRIA
```

```
14 print ( df . describe ( ) )
```

Exemplos de Código 5: Exemplo da análise exploratória do algoritmo.

Com a utilização deste exerto de código são obtidos os valores desejados relativos às variáveis quantitativas. Na Tabela 7, são ilustrados os valores.

Atributo	Média	Desvio	Min.	25%	50%	75%	Max.
TEMPO ENTREGA	20,59	21,16	0	6	14	28	149
QUANTIDADE	335,034	1217,19	0,014	18	88	282	49000

Tabela 7: Dados Estatísticos das variáveis numéricas do dataset.

Realizada uma análise exploratória do dataset, é necessário realizar uma limpeza dos dados. As operações efetuadas nos conjuntos de dados recolhidos foram a eliminação de inputs de entrada em que o tempo de entrega era superior a 150 dias. Este valor máximo foi obtido em consenso com os responsáveis do domínio empresarial, sendo que a existência destes valores, se deviam a má utilização dos sistemas de informação por parte do utilizador, e possivelmente falha de comunicação entre os intervenientes do processo. Outra tarefa, foi a eliminação de inputs com tempo de entrega negativo, ou seja, má inserção por parte do utilizador. Os valores da Tabela 7, foram obtidos após esta limpeza e já demonstram o resultado destas tarefas realizadas.

Alguns dos modelos de ML necessitam a realização da tarefa de limpeza dos *Outliers*, eliminando os valores que diferenciam drasticamente dos outros. Nas abordagens em causa apenas no algoritmo de redes neuronais pode existir a necessidade de efetuar esta remoção, nas restantes abordagens são algoritmos que lidam com a existência de *outliers*. No exerto de código 6, é ilustrada a função desenvolvida para efetuar a remoção.

```

1 import numpy as np
2 def remove_outliers(dataset) :
3     for x in ['tempo_demora']:
4         q75, q25 = np.percentile(dataset.loc[:, x], [75, 25])
5         intr_qr = q75 - q25
6
7         max = q75 + (1.5 * intr_qr)
8         min = q25 - (1.5 * intr_qr)
9
10        dataset.loc[dataset[x] < min, x] = np.nan
11        dataset.loc[dataset[x] > max, x] = np.nan
12
13    dataset = dataset.dropna(axis=0)
14    return dataset;

```

Exemplos de Código 6: Função de remoção de *outliers*.

No sentido de normalizar o *dataset* é efetuada a tarefa de *Feature Scaling*, com auxílio à biblioteca *sklearn* do *Python* é utilizado o *StandardScaler*, no excerto de código7 é possível visualizar.

```

1 from sklearn.preprocessing import StandardScaler
2 sc = StandardScaler()
3 X_train_std = sc.fit_transform(X_train)
4 X_test_std = sc.transform(X_test)

```

Exemplos de Código 7: Função de *Feature Scaling*.

Uma vez os dados caracterizados e organizados, é necessário treinar os dados com os modelos escolhidos. Para o *Random Forest* e *Gradient Boosting*, os parâmetros variáveis são o *Number Estimators* e a *Maximum Depth*, e para as Redes Neurais, *Batch Size* e as *Epochs*. A tabela 8, ilustra os vários conjuntos de valores escolhidos para preencher os diferentes parâmetros. Estes conjuntos são utilizados pela técnica de *Cross-Validation*, tendo como objetivo o treino dos modelos com todos estes parametros, identificando qual o conjunto de dados apresentados com melhor performance [41].

Modelo	<i>Nrº Estimators</i>	<i>Max Depth</i>	<i>Epochs</i>	<i>Batch Size</i>
<i>Random Forest</i>	50,100,150,200	10,15,20	—	—
<i>Gradient Boosting</i>	50,100,150,200	10,15,20	—	—
<i>Rede Neuronal</i>	—	—	10,20,50,100	10,20,50

Tabela 8: Parâmetros variáveis dos modelos.

No sentido de se avaliar a performance são utilizadas as métricas já referidas anteriormente, RMSE, MSE, MAE e também uma medida estatística o R-Quadrado (R2), tendo como objetivo medir o quão próximos os dados estão da linha da regressão. O valor do R2, varia no intervalo de 0-1, quanto maior for melhor o modelo se ajusta aos seus dados. Tendo em conta o estudo feito é dada a prioridade à métrica RMSE, que calcula o erro médio ao quadrado entre os valores observados e as previsões. A globalidade do processo, aqui descrito, no contexto do problema em causa é realizado diariamente, uma vez que existe uma inserção constante de dados nas fontes, dados esses que podem fazer com que a qualidade do modelo seja superior. Os resultados do processo aqui descrito é discutido no capítulo seguinte.

Capítulo 5

Resultados

Realizados os passos necessários nas abordagens utilizadas, neste capítulo são ilustrados os resultados. No que diz respeito à abordagem da AHP, reunidos os pesos de cada alternativa relativa a cada critério, obtem-se a matriz de importância final, abaixo ilustrada na Tabela 9.

Critério — Alternativa	Entrega	Preço	Origem	Prioridade	Class.
30	0.31 x 0.7	0.028 x 0.24	0.26 x 0.064	0.2386	2º
46	0.022 x 0.7	0.039 x 0.24	0.26 x 0.064	0.0412	6º
256	0.15 x 0.7	0.086 x 0.24	0.26 x 0.064	0.1403	4º
650	0.31 x 0.7	0.26 x 0.24	0.10 x 0.064	0.2846	1º
698	0.15 x 0.7	0.24 x 0.24	0.025 x 0.064	0.1608	3º
1195	0.067 x 0.7	0.35 x 0.24	0.09 x 0.064	0.1346	5º

Tabela 9: Exemplo da classificação final relativa a uma família de materiais.

Observando os valores obtidos, pode-se concluir que o melhor fornecedor para esta família é o 650. Observando, um pouco mais em detalhe os pesos dos fornecedores, pode-se concluir que os valores de peso nos critérios mais importantes, Entrega e Preço, são os mais elevados, reflectindo-se no valor global do fornecedor. Como já referido anteriormente o processo é realizado diariamente, os dados aqui fornecidos dão apoio no processo de seleção de fornecedores durante o processo de negócio de uma encomenda, tanto na fase de compra de matéria-prima, como escolha de um Subcontratado. Nesse sentido os dados são integrados no sistema de informação associado ao projeto, mas também disponibilizados numa plataforma colaborativa e desenvolvida no contexto do projeto, como ilustrado na Figura 21, abaixo representada.

Fornecedor	Nome do Fornecedor	Família	Nome da Família	Pontuação	Data de Criação
0698	VITOR COSTA - ARTIGOS PARA CALÇADO,	RVT	RIVETES	0.5	2022-11-16
0820	RZ - COMPONENTES PARA CALÇADO, LDA	RVT	RIVETES	0.5	2022-11-16
1290	PALMIFORRA UNIPessoal, LDA	062	MATERIAL DIVERSO	0.06	2022-11-16
0698	VITOR COSTA - ARTIGOS PARA CALÇADO,	FCH	FECHOS	0.11	2022-11-16
1397	HIJOS DE MANUEL GONZALEZ, SL	CX	CAIXAS	0.01	2022-11-16
0421	GRIFAGEM JOSE PEDRO-SOC.UNIPessoal, LDA	ME	MATERIAL DE EMBALAGEM	0.15	2022-11-16
0613	GODINHO & SANTOS, LDA.	ME	MATERIAL DE EMBALAGEM	0.14	2022-11-16
0650	MIGUEL DE ALMEIDA CRUZ, LDA.	ME	MATERIAL DE EMBALAGEM	0.13	2022-11-16
1322	PLASTICOS MACAR LDA	ME	MATERIAL DE EMBALAGEM	0.13	2022-11-16
1487	ETIMED ETIQUETAS, S.L.	ME	MATERIAL DE EMBALAGEM	0.1	2022-11-16

Figura 21: Plataforma colaborativa (Classificação da família de materiais usada como exemplo).

Apesar de ser possível obter uma classificação final, isto pode não expressar a melhor classificação possível, devido a certas características do setor, que foram sendo mencionadas ao longo do documento. A baixa quantidade de critérios é uma das características na revisão de literatura realizada a quantidade de critérios a utilizar neste tipo de abordagens é imensa, mas para a sua utilização é necessário que as fontes de recolha associadas ao projeto tenham a capacidade de os fornecer.

A falta de exigência das organizações na execução correta dos processos de negócio, faz com o preenchimento dos dados necessários não seja realizado de forma rigorosa e atempada, consequentemente a qualidade das análises e qualidade na sua utilização em abordagens deste género seja menor. Sendo assim estas características podem influenciar os dados e na sua utilização demonstrar uma análise irrealista, que faz com que os gestores possam tomar decisões erradas.

No que diz respeito à abordagem preditiva utilizando algoritmos de ML, para o seu desenvolvimento além da linguagem *Python* foi utilizado uma plataforma open-source, *H2O*¹, no sentido da possibilidade de criação de modelos de forma escalável e rápida. Com a utilização das ferramentas da plataforma, é possível obter os melhores parâmetros para os modelos utilizados, bem como uma maior planóplia de métricas de performance. Na Tabela 10 são ilustrados os resultados obtidos.

¹<https://h2o.ai/platform/ai-cloud/make/h2o/>

Modelo	Nr ^o Estimators	Max Depth	Epochs	Batch Size	RMSE	MAE	MSE	R2
Random Forest	100	10	—	—	14.98	9.73	224.37	0.092
Gradient Boosting	150	10	—	—	13,26	8.67	175.85	0.29
Rede Neuronal	—	—	10	10	15.61	9.80	243.64	0.014

Tabela 10: Precisão dos algoritmos ML testados com base em diferentes medidas de erro.

No sentido de se perceber a importância das variáveis definidas para o algoritmo, na Figura 22 é possível analisar quais as variáveis mais importantes para o algoritmo.

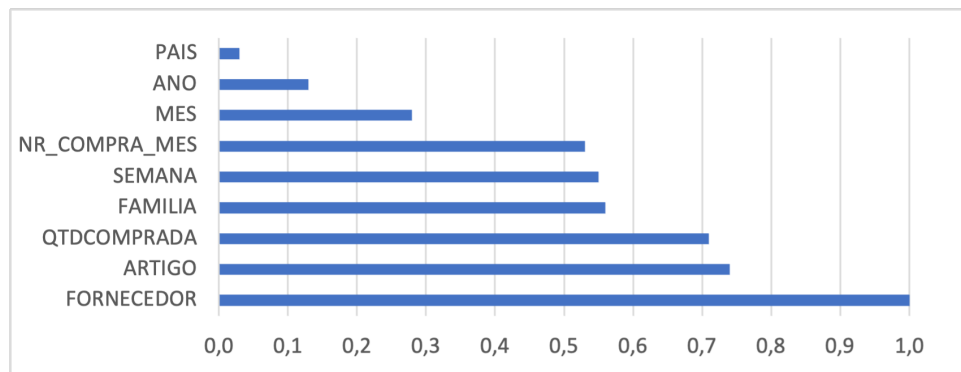


Figura 22: Importância das variáveis (Modelo com melhores resultados).

Observando os valores obtidos o modelo com melhores resultados, utilizado como métrica de referência o RMSE, o *Gradient Boosting* é o melhor modelo, com 13.62 RMSE e com 0,29 R2. Com a utilização do *Cross Validation*, os parâmetros variáveis que obtiveram melhores resultados foram, *Nr^o Estimators* de 150 e o *Max Depth* de 10. Com o auxílio da plataforma *H2O*, obtém-se de forma mais célere o modelo que melhor se adapta aos dados em causa, num passo seguinte com auxílio da linguagem *Python* e bibliotecas dedicadas ao tema (*sklearn*, *numpy*, *keras*, *tensorflow*) foi desenvolvido o algoritmo para execução. Esta execução, à imagem da abordagem anteriormente apresentada, é realizada de forma diária, devido à constante inserção de novos dados. Os resultados obtidos são utilizados no apoio à decisão no processo de planeamento, como já mencionado, numa fase inicial na execução das compras de matérias-primas, mas também no acompanhamento do planeamento quando existe um início de produção próximo, prevendo o tempo que uma matéria-prima demorará a chegar. Portanto, estes resultados são também integrados no sistema de informação e disponibilizados na plataforma colaborativa criada. Na Figura 23, é ilustrado um exemplo dos resultados obtidos numa execução do algoritmo.

Artigo	Fornecedor	Familia	Quantidade Encomendada	Data de Compra	Data Prevista Entrega	Tempo Previsto
DIV-SER083	RODAMA-IND. CALÇADO, LDA.	MATERIAL DIVERSO	1812	2022-02-24	2022-03-11	15
SLT-0222A	CAIRO - COMPONENTES PARA CALÇADOLD	SALTOS	146	2022-01-04	2022-01-22	18
PL-MTFL0041	RAMIRO OLIVEIRA & FILHOS, LDA	PELARIAS	60	2022-04-07	2022-04-22	15
TRF-12M	PEDRO MANUEL SOUSA SANTOS,UNIP. LDA	TARIFAS	135	2022-03-03	2022-03-11	8
TRF-8TT	PEDRO MANUEL SOUSA SANTOS,UNIP. LDA	TARIFAS	36	2022-01-05	2022-01-12	7
FIV-0551	DOMINGOS PINHO FERNANDES	FIVELAS	16	2022-04-01	2022-04-15	14
PL-VCFL0265A	CARPANESE SPA	PELARIAS	650	2022-03-07	2022-03-23	16
CL-0011	TECMACAL, S.A.	COLAS E DERIVADOS	48	2022-01-11	2022-01-17	6
TST-0645B	S.S.F.-COMP. P/ CALÇADO LDA	TESTEIRAS	36	2022-01-05	2022-01-20	15
TRF-4P328218	PEDRO MANUEL SOUSA SANTOS,UNIP. LDA	TARIFAS	1	2022-04-28	2022-05-05	7

Figura 23: Plataforma colaborativa- Exemplo de previsão do tempo de entrega.

De forma geral, os resultados não são ideais, observando os restantes modelos utilizados, os valores das métricas encontram-se bastante próximas. Algumas das razões para estes resultados devem-se, à baixa quantidade de variáveis úteis para o problema, influenciando a qualidade do conhecimento fornecido ao modelo. Este facto, corrobora algumas das análises realizadas aos sistemas de informação do setor. Tipicamente, estes são estruturados por modelos de dados antigos, falta de automatização de processos, necessitando de intervenção humana para o sucesso da mesma e inexistência de tipo de dados que distingam as diferentes interações. Neste sentido, pode-se concluir, que os sistemas de informação do setor ainda não têm estrutura para a utilização deste tipo de técnicas.

Outra das razões deve-se, à falta de consistência dos dados, refletindo a falta de automatização dos *softwares*, mas também resultado das características do setor, nomeadamente, a informalidade dos processos. A falta de comunicação e registo dos dados existentes nesses processos, dão origem à falta dos mesmos, por vezes, à passagem de informação incompleta. Num exemplo relatado pelos parceiros do setor, numa das reuniões efetuadas, existe uma certa compra de um material a um fornecedor, o mesmo quando termina a produção, comunica de forma informal à empresa, e esta também informalmente, solicita ao fornecedor a entrega da matéria-prima no data posterior. Este facto influencia drasticamente o valor real do prazo de entrega do artigo, uma vez que, durante toda esta comunicação existente entre o fornecedor e a empresa não foi registado qualquer dado relativamente ao assunto debatido. Uma forma de ultrapassar, é a introdução de mecanismos automáticos, que informem o estado da encomenda no fornecedor, obtendo o prazo exato de entrega do artigo. Estes são os tipos de atributos que podem distinguir as diferentes interações e passar uma melhor informação.

Outra das características do setor reflete-se nos dados existentes, como é o caso em que o valor do artigo com o maior número de registos é de 512 num conjunto de dados de 34007, sendo assim, o baixo número de registos por artigo e o alto número de artigos diferentes influenciam o resultado. O elevado número de diferentes produtos deve-se à sua constante evolução, e uma simples alteração na composição ou outra característica alterada, leva à criação de um novo código, mesmo que seja globalmente o mesmo que o anterior.

Capítulo 6

Conclusões

Realizado o estudo de diversas abordagens, no sentido da melhoria dos processos da *Supply Chain* do setor do calçado, foram analisadas as abordagens com maior referência na literatura, como métodos multicritério, mas também abordagens mais emergentes, como técnicas de ML. Estas abordagens visam oferecer aos intervenientes da rede, a construção de ferramentas de apoio à tomada de decisão, melhorando a análise e utilização dos dados existentes. Em termos de resultados obtidos, é possível concluir que não são os esperados, o modelo de ML para previsão dos tempos de entrega demonstra um erro indesejável. A avaliação dos fornecedores apesar de se obter uma solução, tendo em conta as características do setor e os dados que são fornecidos, existe a possibilidade de uma classificação não tão realista como o desejado. Contudo, foi possível obter um certo conjunto de conclusões.

A noção da importância dos sistemas de informação no setor é um dos principais problemas, como já referido, esta noção falha logo no topo. Os dirigentes não percebem a utilidade que este tipo de ferramentas lhes podem dar na tomada da decisão, neste sentido não exigem aos seus colaboradores a correcta realização dos processos de negócio. Esta falta de exigência leva, à informalidade nos processos de negócio e à falha de dados nucleares para a análise sendo assim, os dados existentes não transmitem qualidade nem confiança. Existe a necessidade de sensibilizar os dirigentes para a importância destas ferramentas, e também melhorar estes processos para que a qualidade dos dados melhore, e conseqüentemente um aumento da confiança das empresas na importância da análise dos dados para a sua superação.

Junto desta melhoria, existe a necessidade de investir e realizar uma reestruturação nas empresas tecnológicas associadas ao setor. Uma vez que, os *softwares* utilizados estão cada vez mais inadaptados para as novas tecnologias do mercado atual. Os modelos de dados por detrás dos *softwares* são limitativos em termos de recolha para abordagens como *Machine*

Learning, não existem dados, nem qualidade nos mesmos, para que possam fazer a diferença substancial na previsão ou classificação de algo. Neste sentido, existe a necessidade das empresas de repensarem a arquitetura dos seus *softwares*, com o objetivo de integrar e interagir com estas novas ferramentas e aumentar a planóplia de soluções para o apoio às empresas.

Contudo, com estas ferramentas criadas e associadas às melhorias propostas, foi possível alcançar os objetivos propostos. Apesar dos resultados obtidos, estas ferramentas fornecem às empresas uma fonte de apoio na tomada das várias decisões existentes, no decorrer do seu processo produtivo e também são a alavanca necessária para demonstrar a importância da utilização dos dados no crescimento das empresas.

6.1 Reflexão Crítica

Apesar de conseguido o objetivo da construção dos algoritmos no contexto do apoio à *Supply Chain*, os fracos resultados ,devido às limitações no setor, fez com que exista ainda um certo negacionismo na utilização destas ferramentas. No entanto, o desenvolvimento deste trabalho forneceu uma contribuição para demonstrar as melhorias necessárias a realizar nas empresas, e também a potencialidade deste tipo de ferramentas na ajuda à tomada de decisão. Ainda assim é importante referir que existe uma tarefa muito importante e necessária a fazer nas empresas, no sentido de lhe dar conhecimentos e ferramentas para o seu crescimento.

6.2 Trabalho Futuro

O trabalho futuro é centrado nos objetivos do projeto financiado associado a este trabalho, a primeira fase será a interação com as empresas demonstrando o trabalho desenvolvido, com o intuito de transmitir as potencialidades e utilidades. Numa fase seguinte, o processo de melhoria dos processos empresariais e em paralelo a reestruturação das empresas tecnológicas para a adaptação a estas novas ferramentas. Por fim, a integração das ferramentas construídas com os sistemas de informação associados.

Durante este processo, serão realizadas melhorias às ferramentas já desenvolvidas, em termos do algoritmo de classificação associando novas variáveis na tomada de decisão, no sentido de melhorar a classificação. Com a utilização do algoritmo de Previsão, será construído um sistema para criação de alertas, com o objetivo de alertar os intervenientes da rede para situações de riscos, como previsão de entrega atrasada de alguma matéria-prima.

Bibliografia

- [1] L. P. L. d. S. Mendes, *ANÁLISE DOS MÉTODOS DE SELEÇÃO DE FORNECEDORES*. PhD thesis, Instituto Superior de Engenharia do Porto, 2013.
- [2] L. d. v. d. W. Boer, *operations Research in support of purchasing. Design of a toolbox for supplier selection*. PhD thesis, University of Twente, Enschede, The Netherlands., 1998.
- [3] S. M. A. Cunha, “Seleção de Fornecedores – Estudo de Caso,” tech. rep., Universidade de Aveiro, 2008.
- [4] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design science in information systems research,” *MIS Q.*, vol. 28, p. 75–105, mar 2004.
- [5] J. Chai, J. N. Liu, and E. W. Ngai, “Application of decision-making techniques in supplier selection: A systematic review of literature,” *Expert systems with applications*, vol. 40, no. 10, pp. 3872–3885, 2013.
- [6] H. Dhiman and D. Deb, *Multi-criteria Decision-Making: An Overview*, pp. 19–36. 01 2020.
- [7] M. Aruldoss, T. M. Lakshmi, and V. P. Venkatesan, “A survey on multi criteria decision making methods and its applications,” *American Journal of Information Systems*, vol. 1, no. 1, pp. 31–43, 2013.
- [8] P. Agarwal, M. Sahai, V. Mishra, M. Bag, and V. Singh, “A review of multi-criteria decision making techniques for supplier evaluation and selection,” *International journal of industrial engineering computations*, vol. 2, no. 4, pp. 801–810, 2011.
- [9] H. Karunathilake, E. Bakhtavar, G. Chhipi-Shrestha, H. R. Mian, K. Hewage, and R. Saadiq, “Decision making for risk management: a multi-criteria perspective,” in *Methods in chemical process safety*, vol. 4, pp. 239–287, Elsevier, 2020.
- [10] S. J. Jansen, “The multi-attribute utility method,” in *The Measurement and Analysis of Housing Preference and Choice*, pp. 101–125, Springer, Dordrecht, 2011.

- [11] R. W. Saaty, "The analytic hierarchy process—what it is and how it is used," *Mathematical Modelling*, vol. 9, pp. 161–176, Jan. 1987.
- [12] R. Hruška, P. Průša, and D. Babić, "The use of AHP method for selection of supplier," *Transport*, vol. 29, pp. 195–203, Apr. 2014. Publisher: Taylor & Francis eprint: <https://doi.org/10.3846/16484142.2014.930928>.
- [13] "Criteria in AHP: A Systematic Review of Literature | Elsevier Enhanced Reader."
- [14] T. L. Saaty, *Theory and applications of the analytic network process: decision making with benefits, opportunities, costs, and risks*. RWS publications, 2005.
- [15] F. T. S. Chan and H. K. Chan, "An AHP model for selection of suppliers in the fast changing fashion market," *The International Journal of Advanced Manufacturing Technology*, vol. 51, pp. 1195–1207, Dec. 2010.
- [16] S. H. Amin and J. Razmi, "An integrated fuzzy model for supplier management: A case study of isp selection and evaluation," *Expert systems with applications*, vol. 36, no. 4, pp. 8639–8648, 2009.
- [17] L. De Boer, L. van der Wegen, and J. Telgen, "Outranking methods in support of supplier selection," *European Journal of Purchasing & Supply Management*, vol. 4, no. 2-3, pp. 109–118, 1998.
- [18] B. Roy, "The outranking approach and the foundations of electre methods," in *Readings in Multiple Criteria Decision Aid* (C. A. Bana e Costa, ed.), (Berlin, Heidelberg), pp. 155–183, Springer Berlin Heidelberg, 1990.
- [19] L. Benyoucef, H. Ding, and X. Xie, "Supplier selection problem : Selection criteria and methods," 01 2003.
- [20] F. e. G. H. Akram, Muhammad e Ilyas, "Tomada de decisão em grupo multicritério com base no método electre i em informações fuzzy pitagóricas."
- [21] S. Zhaoxu and H. Min, "Multi-criteria decision making based on promethee method," in *2010 international conference on computing, control and industrial engineering*, vol. 1, pp. 416–418, IEEE, 2010.
- [22] B. Vahdani, A. H. K. Jabbari, V. Roshanaei, and M. Zandieh, "Extension of the electre method for decision-making problems with interval weights and data," *The International Journal of Advanced Manufacturing Technology*, vol. 50, no. 5, pp. 793–800, 2010.

- [23] Y.-H. Chen, T.-C. Wang, and C.-Y. Wu, "Strategic decisions using the fuzzy promethee for is outsourcing," *Expert Systems with Applications*, vol. 38, no. 10, pp. 13216–13222, 2011.
- [24] M. Zeleny, *Multiple criteria decision making Kyoto 1975*, vol. 123. Springer Science & Business Media, 2012.
- [25] I. Emovon, R. A. Norman, J. M. Alan, and K. Pazouki, "An integrated multicriteria decision making methodology using compromise solution methods for prioritising risk of marine machinery systems," *Ocean Engineering*, vol. 105, pp. 92–103, 2015.
- [26] B. O'Neill, "Chapter 3 - euclidean geometry," in *Elementary Differential Geometry (Second Edition)* (B. O'Neill, ed.), pp. 100–129, Boston: Academic Press, second edition ed., 2006.
- [27] S. Opricovic and G.-H. Tzeng, "Compromise solution by mcdm methods: A comparative analysis of vikor and topsis," *European journal of operational research*, vol. 156, no. 2, pp. 445–455, 2004.
- [28] J. A. Crispim and J. Pinho de Sousa, "Partner selection in virtual enterprises: a multi-criteria decision support approach," *International Journal of Production Research*, vol. 47, no. 17, pp. 4791–4812, 2009.
- [29] S.-Y. Chou and Y.-H. Chang, "A decision support system for supplier selection based on a strategy-aligned fuzzy smart approach," *Expert systems with applications*, vol. 34, no. 4, pp. 2241–2253, 2008.
- [30] M. Yazdi, F. Khan, R. Abbassi, and R. Rusli, "Improved dematel methodology for effective safety management decision-making," *Safety science*, vol. 127, p. 104705, 2020.
- [31] C. A. Weber, *A decision support system using multicriteria techniques for vendor selection*. PhD thesis, University Microfilms International, Ann Arbor, 1991.
- [32] L. de Boer, E. Labro, and P. Morlacchi, "A review of methods supporting supplier selection," *European Journal of Purchasing & Supply Management*, vol. 7, pp. 75–89, June 2001.
- [33] C. Weber, J. Current, and A. Desai, "Non-cooperative negotiation strategies for vendor selection.."
- [34] J. Razmi, H. Rafiei, and M. Hashemi, "Designing a decision support system to evaluate and select suppliers using fuzzy analytic network process," *Computers & Industrial Engineering*, vol. 57, no. 4, pp. 1282–1290, 2009.

- [35] J. Branke, J. Branke, K. Deb, K. Miettinen, and R. Slowiński, *Multiobjective Optimization: Interactive and Evolutionary Approaches*. Genetic algorithms and evolutionary computation, Springer, 2008.
- [36] S. M. Lee and D. L. Olson, "Goal programming," in *Multicriteria decision making*, pp. 203–235, Springer, 1999.
- [37] J. Chai and E. W. Ngai, "Decision-making techniques in supplier selection: Recent accomplishments and what lies ahead," *Expert Systems with Applications*, vol. 140, p. 112903, 2020.
- [38] S. Cohen, "Chapter 2 - the basics of machine learning: strategies and techniques," in *Artificial Intelligence and Deep Learning in Pathology* (S. Cohen, ed.), pp. 13–40, Elsevier, 2021.
- [39] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, *et al.*, "Top 10 algorithms in data mining," *Knowledge and information systems*, vol. 14, no. 1, pp. 1–37, 2008.
- [40] K. De Jong, "Learning with genetic algorithms: An overview," *Machine learning*, vol. 3, no. 2, pp. 121–138, 1988.
- [41] A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow*. Canada: O'Reilly Media, Inc., 2nd ed., 2019.
- [42] D. Maulud and A. M. Abdulazeez, "A review on linear regression comprehensive in machine learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 4, pp. 140–147, 2020.
- [43] R. Carbonneau, K. Laframboise, and R. Vahidov, "Application of machine learning techniques for supply chain demand forecasting," *European Journal of Operational Research*, vol. 184, no. 3, pp. 1140–1154, 2008.
- [44] L. Lingitz, V. Gallina, F. Ansari, D. Gyulai, A. Pfeiffer, W. Sihn, and L. Monostori, "Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer," *Procedia Cirp*, vol. 72, pp. 1051–1056, 2018.
- [45] M. B. de Oliveira, G. Zucchi, M. Lippi, D. F. Cordeiro, N. R. da Silva, and M. Iori, "Lead time forecasting with machine learning techniques for a pharmaceutical supply chain.," in *ICEIS (1)*, pp. 634–641, 2021.

- [46] S. Islam and S. H. Amin, “Prediction of probable backorder scenarios in the supply chain using distributed random forest and gradient boosting machine learning techniques,” *Journal of Big Data*, vol. 7, no. 1, pp. 1–22, 2020.
- [47] “O calçado no mundo panorama estatístico 2021,” tech. rep., APICCAPS, 2021.
- [48] D. J. d. S. Catanho, *Indústria Portuguesa de Calçado: Alteração do Modelo de Negócio para as PME’s*. PhD thesis, Faculdade de Economia da Universidade do Porto, Porto, Sept. 2014.
- [49] L. M. Amaral, “O CONCEITO DE REINDUSTRIALIZAÇÃO, INDÚSTRIA 4.0 E A POLÍTICA INDUSTRIAL PARA O SÉCULO XXI- O caso Português,” May 2016.
- [50] A. Shamsuzzoha, T. Kankaanpaa, L. M. Carneiro, R. Almeida, A. Chiodi, and R. Fornasiero, “Dynamic and collaborative business networks in the fashion industry,” *International Journal of Computer Integrated Manufacturing*, vol. 26, no. 1-2, pp. 125–139, 2013.
- [51] M. J. d. S. P. d. Lima, *Redes colaborativas na gestão da cadeia de abastecimento: o sector vitivinícola na península de Setúbal*. PhD thesis, Escola Superior de Ciências Empresariais, 2013.
- [52] P. Gomes and R. Rebelo, “Diagnóstico as is e oportunidades de melhoria,” tech. rep., INESC TEC, 2013.
- [53] A. Correia Alves and R. Rebelo, “Cni4.0-foot - gestão de operações na indústria do calçado no paradigma i4.0,” tech. rep., INESC TEC, 2022.
- [54] T. L. Saaty, *Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process*. RWS Publications. Google-Books-ID: wct10TlbbIUC.
- [55] T. L. Saaty, “What is the Analytic Hierarchy Process?,” in *Mathematical Models for Decision Support* (G. Mitra, H. J. Greenberg, F. A. Lootsma, M. J. Rijkaert, and H. J. Zimmermann, eds.), NATO ASI Series, (Berlin, Heidelberg), pp. 109–121, Springer, 1988.
- [56] P. S. Foundation, “A python implementation of the analytic hierarchy process,” 2022.
- [57] Y. Zhang and A. Haghani, “A gradient boosting method to improve travel time prediction,” *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 308–324, Sept. 2015.

Apêndice A

Anexos

A.1 Anexo A - Definição das Matrizes de Importância

```
1 def define_classification_by_family(connection, family):
2     data= "SELECT fa.cliente NR.CLIENTE, f.NOME CLIENTE ..."
3
4     criteria_comparisons = {'Delivery', 'Price': 4, ('Delivery', 'Origin'): 8, ('
5     Price', 'Origin'): 5}
6     criteria = ahpy.Compare('Criteria', criteria_comparisons, precision=5)
7
8     suppliers = create_suppliers(df)
9     if (len(suppliers) > 1):
10         suppliers_pairs = list(itertools.combinations(suppliers, 2))
11         price_values = define_values_price(df, suppliers_pairs)
12         price_comparisons = dict(zip(suppliers_pairs, price_values))
13         delivery_values = define_values_delivery(df, suppliers_pairs)
14         delivery_comparisons = dict(zip(suppliers_pairs, delivery_values))
15         origin_values = define_values_origin(df, suppliers_pairs)
16         origin_comparisons = dict(zip(suppliers_pairs, origin_values))
17
18         price = ahpy.Compare('Price', price_comparisons, precision=5)
19         reportPrice = price.report(show=True)
20         delivery = ahpy.Compare('Delivery', delivery_comparisons, precision=5)
21         reportDelivery = delivery.report(show=True)
22         origin = ahpy.Compare('Origin', origin_comparisons, precision=5)
23         reportOrigin = origin.report(show=True)
24         criteria.add_children([price, delivery, origin])
25     report = criteria.report(show=True)
```

Exemplos de Código 1: Função para construção das matrizes de importância.

A.2 Anexo B - Automatização de atribuição de importância

1	2	3	4	5	6	7	8	9
PT		ES		FR		BG		GR
Mínimo		1ºQ		Média		3ºQ		Máximo

Tabela 11: Exemplo da divisão da escala no critério Origem (Família de Artigo TER).

1	2	3	4	5	6	7	8	9
0.24		1.05		2.24		3.45		4.37
Mínimo		1ºQ		Média		3ºQ		Máximo

Tabela 12: Exemplo da divisão da escala no critério Preço (Família de Artigo TER).

A.3 Anexo C - Análise de Dados dos Fornecedores

Estadística Fornecedor-Familia x Estadística de Fornecedores x

Estadística dos Fornecedores

Colunas

Procurar

Fornecedor	Média Tempo Entrega	Max Tempo Entrega	Min Tempo Entrega	Desvio Tempo Entrega	Média Tempo Atraso	Média Tempo Atempada	Preço Médio
0011	1	10	0	2	3	-1	4.31
0013	1	9	0	2	2	-1	2.95
0019	13	107	1	19	15	-6	13.34
0025	25	139	0	21	19	-12	0.59
0028	6	9	3	3	3	-3	3.35
0030	10	150	0	13	12	-5	0.42
0042	16	32	1	16	16	-15	0.2
0044	16	95	0	17	18	-8	13.47
0046	18	128	0	16	15	-9	0.18
0047	0	1	0	0	1	0	0.02

Linhas por página: 10 1-10 de 204

Figura 24: Análise dos Dados dos Fornecedores

Histograma dos Fornecedores

Colunas

Procurar

Fornecedor	0 a 10 DIAS	10 a 20 DIAS	20 a 30 DIAS	30 a 40 DIAS	40 a 50 DIAS	50 a 60 DIAS	60 a 70 DIAS	70 a 80 DIAS	80 a 90 DIAS	90 a 100 DIAS	100 a 110 DIAS	110 a 12
0011	26	1	0	0	0	0	0	0	0	0	0	0
0013	82	0	0	0	0	0	0	0	0	0	0	0
0019	16	8	3	0	0	0	0	0	0	0	1	0
0025	954	949	815	473	335	190	152	56	29	19	15	12
0028	2	0	0	0	0	0	0	0	0	0	0	0
0030	303	95	31	17	7	1	0	0	0	0	0	0
0042	1	0	0	1	0	0	0	0	0	0	0	0
0044	26	19	4	4	4	0	1	0	0	1	0	0
0046	1284	934	755	298	185	99	35	19	11	5	10	4
0047	12	0	0	0	0	0	0	0	0	0	0	0

Figura 25: Histograma de tempo de entrega dos Fornecedores