



# Aplicação De Modelos De Linguagem De Grande Escala Na Otimização De Problemas De Dimensionamento De Lotes De Produção

ANDRÉ DA SILVA RIBEIRO

outubro de 2024

# **Aplicação De Modelos De Linguagem De Grande Escala Na Otimização De Problemas De Dimensionamento De Lotes De Produção**

**André Da Silva Ribeiro**

**Dissertação para obtenção do Grau de Mestre em  
Engenharia Mecânica, Área de Especialização em  
Gestão Industrial**

**Orientador: André Serra e Santos**

**Júri:**

Presidente:

António Amaral, Professor Adjunto, ISEP

Vogais:

Isabel Espiro Santo, Professora Auxiliar, UM

André Serra e Santos, Professor Adjunto Convidado, ISEP

Porto, setembro 2024



# Agradecimentos

Agradeço, primeiramente, à minha família pelo apoio incondicional e encorajamento contínuo durante toda a jornada acadêmica, sempre acreditando em mim e no meu potencial. Aos meus amigos, pela companhia, compreensão e suporte nos momentos de maior desafio.

Agradeço também aos meus professores, pelo conhecimento e dedicação que transmitiram ao longo do curso. Em especial, ao meu orientador, Professor André Serra e Santos, pela orientação rigorosa, paciência e disponibilidade ao longo deste trabalho. Sua orientação foi essencial para o desenvolvimento e conclusão desta dissertação. A todos, o meu sincero agradecimento.



# Resumo

Os problemas de dimensionamento de lotes têm como objetivo identificar as quantidades ideais de produção para atender à procura, minimizando, simultaneamente, os custos de *setup* e o custo de *stock*. Esta dissertação explora como os *LLMs* está transformou a forma como o problema de dimensionamento de lotes é abordado. O trabalho foca-se inicialmente na evolução da literatura sobre técnicas de IA aplicadas a problemas de dimensionamento de lotes, com o intuito de fornecer uma visão abrangente da área de estudo, focando depois nas plataformas de *LLMs* destacando as diversas formas de aplicação, e mais tarde, no problema de dimensionamento de lotes. Apesar do interesse crescente da comunidade científica na aplicação do ChatGPT, a sua utilização em problemas de otimização ainda não tem recebido atenção. A variabilidade intrínseca ao ChatGPT torna a engenharia de *prompt* essencial para adaptar a ferramenta às especificidades de cada problema, contudo, é frequentemente negligenciada. Esta realidade destaca a necessidade de desenvolver uma abordagem estruturada para observar, testar e avaliar a aplicabilidade do ChatGPT em problemas de dimensionamento de lotes. Desta forma foi investigada a eficácia das plataformas ChatGPT e Copilot na resolução do problema de dimensionamento de lotes. Inicialmente, ambas as plataformas apresentaram resultados insatisfatórios, com inconsistências e falta de repetibilidade. Para superar estas limitações, foi testado o modelo GPT-4 do ChatGPT, mais tarde substituído pelo GPT-4o, que demonstrou maior consistência e capacidade de captar o contexto do problema, resultando em soluções próximas do ótimo. Após várias iterações, foi desenvolvido um *prompt* otimizado que proporcionou resultados mais consistentes na resolução do problema com o ChatGPT. Este *prompt* foi posteriormente aplicado em duas amostras distintas para avaliar o desempenho da *LLMs* m diferentes cenários. Esta abordagem permitiu examinar a eficácia e consistência do ChatGPT em condições variadas, oferecendo uma análise mais robusta da sua aplicabilidade e adaptabilidade aos desafios específicos de cada amostra. Os resultados mostraram que, em ambas as amostras, o ChatGPT superou as restantes heurísticas analisadas, não apresentando diferenças estatisticamente significativas em relação ao método de Wagner-Whitin.

**Palavras-chave:** Dimensionamento de lotes; Inteligência Artificial; *LLMs*; ChatGPT; Heurísticas; Prompt



# Abstract

Lot-sizing problems aim to identify the optimal production quantities to meet demand while simultaneously minimizing setup and inventory costs. This dissertation explores how Large Language Models (LLMs) have transformed the approach to lot-sizing problems. The work initially focuses on the evolution of literature of AI techniques applied to lot-sizing problems, aiming to provide a comprehensive view of the study area, then shifts to LLM platforms, highlighting various application methods, and later delves into the lot-sizing problem itself. Despite the growing interest of the scientific community in applying ChatGPT, its use in optimization problems has yet to receive significant attention. The intrinsic variability of ChatGPT makes prompt engineering essential for adapting the tool to the specifics of each problem, however, this is often overlooked. This reality underscores the need to develop a structured approach to observe, test, and evaluate the applicability of ChatGPT in lot-sizing problems. Thus, the effectiveness of ChatGPT and Copilot platforms in solving batch sizing problems was investigated. Initially, both platforms presented unsatisfactory results, with inconsistencies and a lack of repeatability. To overcome these limitations, the GPT-4 model of ChatGPT was tested and later replaced by GPT-4o, which demonstrated greater consistency and ability to capture the context of the problem, resulting in near-optimal solutions. After several iterations, an optimized prompt was developed, which provided more consistent results in solving the problem with ChatGPT. This prompt was subsequently applied to two distinct samples to assess the performance of LLMs in different scenarios. This approach allowed examining the effectiveness and consistency of ChatGPT under varied conditions, offering a more robust analysis of its applicability and adaptability to the specific challenges of each sample. The results showed that in both samples, ChatGPT outperformed the other heuristics analyzed, with no statistically significant differences compared to the Wagner-Whitin method.

**Key-words:** Lot-Sizing; Artificial Intelligence; LLMs; ChatGPT; Heuristics; Prompt



# Índice

1. Introdução.....	1
1.1. Contextualização e enquadramento .....	1
1.2. Objetivos .....	2
1.3. Metodologia .....	2
1.4. Estrutura do documento .....	3
2. Revisão Bibliográfica .....	5
2.1. Conceitos fundamentais de dimensionamento de lotes .....	5
2.1.1. Planeamento da produção.....	5
2.1.2. Dimensionamento de lotes de Produção .....	6
2.2. Métodos de dimensionamento de lotes USILSP .....	10
2.3. Métodos Heurísticos .....	10
2.3.1. Lot for lot (LFL).....	11
2.3.2. Quantidade económica (EOQ) .....	11
2.3.3. Ordens periódicas (POQ) .....	11
2.3.4. Parts Period Balancing (PPB).....	12
2.3.5. Silver-Meal (SM).....	12
2.4. Métodos Exatos.....	12
2.4.1. Algoritmo Wagner-Whitin (WW) .....	13
2.5. Avaliação da Performance de heurísticas .....	13
2.6. Inteligência Artificial.....	15
2.6.1. Machine learning .....	17
2.6.2. Deep Learning .....	17
2.6.3. Redes Neurais .....	18
2.6.4. Algoritmos Genéticos.....	19
2.7. Plataforma ChatGPT.....	19
2.8. Análise crítica da Bibliometria.....	23
2.8.1. Futuros Temas de Investigação.....	25
2.9. Uso de Inteligência Artificial no Dimensionamento de Lotes .....	26
2.10. Conclusão Bibliográfica .....	33
3. Métodos e Aplicações .....	35
3.1. Modelo Matemático .....	35
3.1.1. SISLP .....	35
3.1.2. SILSP com lotes constantes.....	37
3.1.3. Dimensionamento e Escalonamento De Lotes .....	39
3.2. Utilização do ChatGPT e Copilot.....	42
3.2.1. Prompts.....	42

3.2.2. Desafios e problemas.....	46
3.2.3. Conhecimento no Problema discutido .....	48
3.3. Resolução do Problema através das Plataformas IA.....	53
3.4. Conclusão .....	64
4. Resultados e Discussão .....	65
4.1. Implementação do ChatGPT .....	65
4.2. Apresentação dos resultados.....	66
4.3. Análise dos Resultados.....	70
4.3.1. Estatística Descritiva da Aplicação Do ChatGPT ao Problema .....	70
4.3.2. Inferência Estatística da Aplicação Do ChatGPT ao Problema.....	73
4.4. Conclusão final dos Resultados .....	77
5. Conclusão .....	79
5.1. Conclusões finais .....	79
5.2. Trabalhos futuros .....	81
Referências.....	83
Declaração de Integridade .....	89
Anexos A.....	91

# Lista de Figuras

Figura 1 - Horizonte do planeamento [modificado dex: (Maravelias & Sung, 2009)] .....	6
Figura 2 - Sistema de Classificação das técnicas [modificado de: (Guimarães et al., 2014)] .....	9
Figura 3 - A diferença percentual média entre o custo total do cronograma e a solução de referência de 300 períodos para vários tamanhos de horizontes [fonte: (Simpson 2001)].	15
Figura 4 - Investimento corporativo global anual em inteligência artificial, por tipo [fonte:(Charlie Giattino, Edouard Mathieu, 2023)] .....	16
Figura 5 - Uma ilustração da posição do DL, comparando com ML e IA [modificado de:(Sarker 2021)] .....	18
Figura 6 - Uma arquitetura geral de a) rede superficial com uma camada oculta e b) uma rede neural profunda com múltiplas camadas ocultas [fonte:(Sarker, 2021)] .....	19
Figura 7 - Tempo para alcançar 1 milhão de utilizadores [modificado de:(Bansal, 2023)] .....	20
Figura 8 - Resultados do ChatGPT vs GPT-4 em vários exames [fonte:(OpenAI, 2023)] .....	21
Figura 9 - Inteligência artificial: desempenho em testes de conhecimento vs. computação de treino[fonte:(Charlie Giattino, Edouard Mathieu, 2023)] .....	22
Figura 10 - Publicações na área de Lot-Sizing .....	24
Figura 11 - Publicações na área de AI .....	24
Figura 12 - Áreas de estudo relacionadas com Lot-Sizing .....	24
Figura 13 - Termos mais utilizados no Dimensionamento de lotes .....	25
Figura 14 - Uma arquitetura de rede neural para dimensionamento e sequenciamento de lotes [modificado de:(Lee et al., 2001)] .....	27
Figura 15 - Uma arquitetura de rede neural de retro propagação de três camadas para dimensionamento de lotes [modificado de:(Lee et al., 2001)] .....	27
Figura 16 - Uma rede neural aplicada no sequenciamento em tempo real [fonte:(Lee et al., 2001)] .....	28
Figura 17 - Exemplo geração Imagem com MidJourney (MidJourney, 2024) .....	42
Figura 18 - Exemplos resposta clara (1) .....	43
Figura 19 - Exemplo resposta clara (2) .....	43
Figura 20 - Resposta (2) ChatGPT .....	44
Figura 21 - “Taxonomia de técnicas de engenharia de prompts em LLMs, organizadas em torno de domínios de aplicação, fornecendo uma estrutura diferenciada para personalizar prompts em diversos contextos [fonte:(Sahoo et al., 2024)] .....	45
Figura 22 - Exemplo resposta Errada .....	47
Figura 23 - Exemplo resposta certa .....	47
Figura 24 - Resposta ChatGPT 4 (Codigo no Anexo A) .....	61
Figura 25 - Resposta ChatGPT 4 (2) .....	62
Figura 26 - Gráfico Desvios Amostra 1 (ChatGPT-4o) .....	69
Figura 27 - Gráfico Desvio Amostra 1 (ChatGPT-4) .....	69
Figura 28 - Gráfico Desvios Amostra 2 .....	69
Figura 29 - Gráfico de barras das Frequências do desvio (Amostra 1) .....	71
Figura 30 - Gráfico de barras das Frequências do desvio (Amostra 2) .....	73



# Lista de Tabelas

Tabela 1 - Resumo dos desvios para métodos de dimensionamento de lotes com aumento da variância e índices de sazonalidade [modificado de:(Florim et al. 2019)].....	14
Tabela 2 - Uma nova tipologia/classificação de novas tecnologias alimentadas por IA.....	16
Tabela 3 - Exemplos de Áreas de aplicação do ChatGPT .....	22
Tabela 4 - A ilustração de como obter o padrão de pedido ideal [modificado de:(Şenyiğit and Atici 2013)].....	29
Tabela 5 - Estruturas e treinamento para ANN.....	29
Tabela 6 - As 10 principais redes para os critérios de percentagem de desvio do custo ideal [modificado de:(Şenyiğit & Atici, 2013)].....	29
Tabela 7 - Comparação de ANNBM com métodos de dimensionamento de lotes [modificado de:(Şenyiğit & Atici, 2013)] .....	30
Tabela 8 - Tabela resumo de estudos na área .....	31
Tabela 9 - Resposta ao <i>prompt</i> P1(ChatGPT e Copilot).....	48
Tabela 10 - Resposta ao <i>prompt</i> P2(ChatGPT) .....	49
Tabela 11 - Resposta ao <i>prompt</i> P2(Copilot) .....	49
Tabela 12 - Resposta ao <i>prompt</i> P3 (ChatGPT) .....	50
Tabela 13 - Resposta ao <i>prompt</i> P3(Copilot) .....	51
Tabela 14 - Resposta ao <i>prompt</i> P4(ChatGPT) .....	52
Tabela 15 - Resposta ao <i>prompt</i> P4(Copilot) .....	52
Tabela 16 - Resposta CoPilot Exemplo PP1.....	54
Tabela 17 - Resposta ChatGPT Exemplo PP1 .....	54
Tabela 18 - Resultado ChatGPT (1) .....	56
Tabela 19 - Resultado ChatGPT (2) .....	57
Tabela 20 - Resultados CoPilot (1) .....	58
Tabela 21 - Cálculo Heurística POQ.....	58
Tabela 22 - Resultado ChatGPT 4.....	60
Tabela 23 - Resultado Wagner-Within.....	60
Tabela 24 - Resultado ChatGPT 4 (2).....	61
Tabela 25 - Resultado Wagner-Within (2).....	62
Tabela 26 – Prompt cálculo dos Custos .....	63
Tabela 27 - Resultados Amostra 1.....	67
Tabela 28 - Resultados Amostra 2.....	68
Tabela 29 - Lista de Frequências Amostra 1 .....	71
Tabela 30 - Lista de Frequências Amostra 1 .....	71
Tabela 31 - Lista de Frequências Amostra 1 .....	71
Tabela 32 - Lista de Frequências Amostra 2 .....	72
Tabela 33- Lista de Frequências Amostra 2 .....	72
Tabela 34 - Lista de Frequências Amostra 2 .....	73
Tabela 35 - ANOVA da Média do Desvio Relativo da Amostra 1 .....	74
Tabela 36 - Teste de Scheffe da Medias do Desvio Relativo Amostra 1 .....	75

Tabela 37 - ANOVA da Média do Desvio Relativo da Amostra 2 .....	76
Tabela 38 - Teste de Scheffe da Medias do Desvio Relativo Amostra 2 .....	76

# Acrónimos

## Lista de Acrónimos

AI	<i>"Artificial intelligence"</i>
ACO	<i>"Ant Colony Optimization"</i>
ANN	<i>"Artificial Neural Networks"</i>
ANOVA	<i>"Analysis of Variance"</i>
BA	<i>"Bees Algorithm"</i>
CB	<i>"Cost benefit"</i>
ChatGPT	<i>"Chat Generative Pre-Trained Transformer"</i>
CLSP	<i>"Capacitated multi-item lot-sizing Problem"</i>
CSILSP	<i>"Capacitated Single Item Lot Sizing Problem"</i>
CSLP	<i>"Continuous Setup Lot Sizing Problem"</i>
DL	<i>"Deep Learning"</i>
DLSP	<i>"Discrete lot sizing and scheduling Problem"</i>
DNN	<i>"Deep Neural Networks"</i>
DRL	<i>"Deep Reinforcement learning"</i>
ELSP	<i>"Economic lot scheduling Problem"</i>
EOQ	<i>"Economic Order Quantity"</i>
EOQ	<i>"Economic Order quantity"</i>
GA	<i>"Genetic Algorithm"</i>
IA	Inteligência Artificial
LFL	<i>"Lot for Lot"</i>
LLM	<i>"Large Language Model"</i>
LSP	<i>"Lot Sizing Problem"</i>
LSTM	<i>"Long Short Term Memory"</i>
LUC	<i>"Least unit cost"</i>
ML	<i>"Machine Learning"</i>
MRP	<i>"Material Requirement Planning"</i>
MTO	<i>"Make to Order"</i>
NLP	<i>"Natural Language Processing"</i>

PDP	Plano Mestre de produção
POQ	<i>“Periodic Order Quantity”</i>
PPB	<i>“Parts Period Balancing”</i>
PPO	<i>“Proximal Policy Optimisation”</i>
QEE	Quantidade Económica de Encomenda
RLHF	<i>“Reinforcement learning with Human Feedback”</i>
SILSP	<i>“Single Item Lot Sizing Problem”</i>
SM	<i>“Silver-Meal”</i>
USILSP	<i>“Uncapacitated Single Item Lot Sizing Problem”</i>
WW	<i>“Wagner-Within problem”</i>

# Lista de Equações

EQ. 1 .....	11
EQ. 2 .....	11
EQ. 3 .....	12
EQ. 4 .....	12
EQ. 5 .....	13
EQ. 6 .....	36
EQ. 7 .....	36
EQ. 8 .....	36
EQ. 9 .....	36
EQ. 10 .....	36
EQ. 11 .....	38
EQ. 12 .....	38
EQ. 13 .....	38
EQ. 14 .....	38
EQ. 15 .....	38
EQ. 16 .....	38
EQ. 17 .....	38
EQ. 18 .....	41
EQ. 19 .....	41
EQ. 20 .....	41
EQ. 21 .....	41
EQ. 22 .....	41
EQ. 23 .....	41
EQ. 24 .....	41
EQ. 25 .....	41
EQ. 26 .....	41
EQ. 27 .....	41
EQ. 28 .....	41



# 1. Introdução

Neste capítulo é feita uma introdução à dissertação. É feito o enquadramento da mesma, indicando quais são os principais objetivos e a metodologia utilizada para abordar o tema em questão. Por fim, é apresentada a estrutura do trabalho, na qual são referidos todos os capítulos e subcapítulos do mesmo.

## 1.1. Contextualização e enquadramento

O século XXI tem testemunhado um avanço enorme na área da inteligência artificial (IA), nomeadamente no último ano, com o lançamento de plataforma ChatGPT em novembro de 2022 (OpenAI, 2022). Esta nova revolução tecnológica está a redefinir a maneira como as indústrias e as empresas operam e tomam decisões.

A otimização de problemas de dimensionamento de lotes é uma questão crucial na indústria de produção. O dimensionamento de lotes envolve a decisão de determinar a quantidade de um item que deve ser produzido durante um ciclo de produção específico, com o objetivo de minimizar os custos totais de produção e manutenção (Florim et al., 2019). No entanto, a complexidade desses problemas aumenta com o aumento das variáveis envolvidas, acabando por tornar o problema muito complexo para resolver utilizando métodos tradicionais.

Neste contexto, a Inteligência Artificial (IA) surge como uma ferramenta promissora para lidar com esses problemas. As técnicas de IA, como *Artificial Neural Networks* (ANNs), *Genetic Algorithms* (GAs) e *Deep Learning* (DL), têm demonstrado um grande potencial para resolver problemas de otimização complexos, fornecendo soluções de alta qualidade em um tempo computacional razoável.

Esta dissertação procura implementar IA na área da gestão industrial, com foco na exploração das capacidades do ChatGPT e outras formas de IA no campo do dimensionamento de lotes, com objetivo de investigar como essas ferramentas podem ser usadas para melhorar a eficácia e a eficiência da resolução do problema.

## **1.2. Objetivos**

Um dos objetivos desta dissertação é verificar as capacidades do ChatGPT na resolução de problemas de dimensionamento de lotes. Outro objetivo é verificar e explorar as técnicas de IA utilizadas no problema descrito, e depois considerar a implementação de uma dessas mesmas técnicas.

A pesquisa tem o intuito de avaliar a viabilidade e o desempenho do ChatGPT e as técnicas de IA em comparação com as metodologias de dimensionamento de lotes convencionais. O objetivo final desta dissertação é chegar a uma conclusão sobre a possibilidade de simplificação do processo de dimensionamento de lotes por meio da aplicação eficaz do ChatGPT.

## **1.3. Metodologia**

Para cumprir o propósito estipulado, este trabalho optará por adotar uma metodologia dedutiva. Essa abordagem parte da teoria como base para a formulação do problema e, em seguida, é testada com dados de instâncias de problemas reconhecidos academicamente. A escolha por uma metodologia dedutiva permitirá uma análise aprofundada do uso do ChatGPT e de outras técnicas de IA no dimensionamento de lotes, fornecendo resultados confiáveis e conclusões fundamentadas. A elaboração do trabalho seguirá a metodologia descrita nos pontos abaixo:

- Identificar conceito do trabalho;
- Revisão da literatura envolvida nos assuntos abordados;
- Demonstração e resolução dos problemas acadêmicos com métodos heurísticos e exatos;
- Estudo das capacidades do ChatGPT;
- Estudo e implementação do ChatGPT e outras plataformas na resolução de problemas de dimensionamento de lotes;
- Avaliação através de dados estatísticos da efetividade e precisão do ChatGPT na resolução de Problemas de Lotes;

## 1.4. Estrutura do documento

A presente dissertação encontra-se dividida em 3 capítulos.

O capítulo 1 - “Introdução” ao trabalho realizado, engloba o enquadramento geral do trabalho, os objetivos que foram definidos, a metodologia seguida para a sua realização, bem como a estrutura do trabalho.

O capítulo 2 - “Revisão da Literatura” são abordados os conceitos necessários à compreensão do trabalho desenvolvido, tais como: o tema *lot-sizing*, as técnicas de IA, o ChatGPT e uma revisão sobre os trabalhos publicados com a implementação de técnicas de IA. Será também feita uma análise da bibliometria, onde os conceitos discutidos no Capítulo 2 são avaliados e analisados. Neste capítulo é ainda efetuada uma conclusão e análise de temas futuros.

No final são apresentadas as Referências Bibliográficas, onde se encontram todos os artigos, publicações e outras fontes de informação utilizadas para a realização deste trabalho.



## 2. Revisão Bibliográfica

Neste capítulo é apresentada a fundamentação teórica e revisão de literatura elaborada que serviu de base ao trabalho. Esta foca-se no Dimensionamento de Lotes, dos tipos de problemas e classificação com especial foco nos métodos de solução. Na segunda metade deste capítulo é introduzido o tema Inteligência artificial (IA) e as suas subáreas, o ChatGPT, e finaliza com a apresentação de alguns casos de estudo em que foram implementadas técnicas de IA nos problemas de dimensionamento de lotes.

### 2.1. Conceitos fundamentais de dimensionamento de lotes

O dimensionamento de lotes é essencial na gestão das operações porque influencia diretamente o custo operacional de uma empresa. Envolve a decisão de quantos produtos devem ser produzidos em cada lote, com objetivo de equilibrar custos, a procura, a capacidade de produção e o *stock*. Esta decisão está intrinsecamente ligada ao Plano Mestre de Produção (PDP), pois afeta a capacidade de atender ao plano estabelecido. Além disso, impacta o *Materials Requirement Planning* (MRP), pois determina a quantidade de matérias-primas necessárias. O dimensionamento de lotes também tem implicações diretas no planeamento das operações, uma vez que influencia a programação da linha de produção. E, por fim, afeta a gestão de *stocks*, uma vez que determina o nível de produtos acabados armazenados e os custos associados. Portanto, a decisão adequada de dimensionamento de lotes desempenha um papel crucial na otimização dos processos produtivos e no equilíbrio entre eficiência e custos numa organização (Florim et al., 2019).

#### 2.1.1. Planeamento da produção

O principal objetivo do planeamento da produção é desenvolver um plano baseado em previsões de procura ou pedidos de clientes que incluam os produtos a serem fabricados para um horizonte de planeamento específico. O plano deverá fornecer à empresa informação sobre quais produtos serão produzidos e em que período (Maravelias & Sung, 2009).

Segundo (Maravelias & Sung, 2009) os horizontes dos modelos de planeamento de produção são normalmente divididos em três categorias (Figura 1) : planeamento de longo prazo, médio prazo ou curto prazo.

O planeamento de longo prazo consiste na tomada de decisões estratégicas a longo prazo, como escolhas de equipamentos e produtos.

O planeamento de médio e curto prazo envolvem decisões sobre o fluxo de materiais, dimensionamento de lotes de produção e programação de produção para otimizar o desempenho geral da linha de produção.

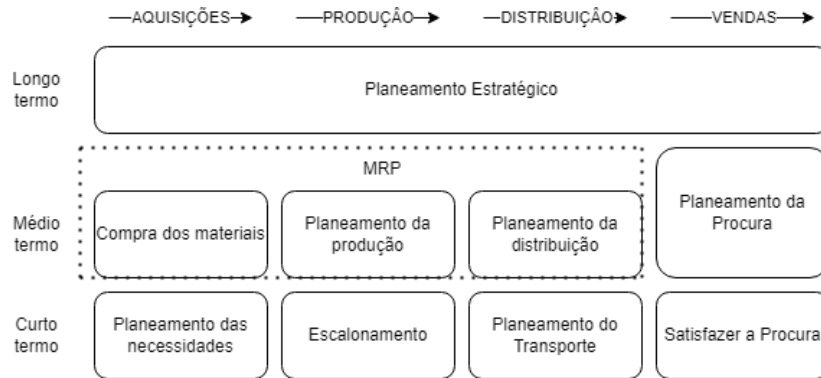


Figura 1 - Horizonte do planeamento [modificado dex: (Maravelias & Sung, 2009)]

### 2.1.2. Dimensionamento de lotes de Produção

Uma das mais reconhecidas e utilizadas técnicas de planeamento de produção é o MRP (Miclo et al., 2015, Miclo et al., 2019), que consiste em converter o (PDP) num plano para comprar ou fabricar os produtos e os seus respetivos componentes.

O dimensionamento de lotes envolve a decisão do tamanho ideal dos lotes de produção a serem produzidos ou comprados e determina quando e quanto produzir de cada produto. Segundo o trabalho de (Alves et al., 2021) esta decisão afeta diretamente:

- Redução de Custos: A decisão adequada sobre o tamanho do lote leva à redução dos custos de produção, armazenagem e transporte, resultando em maior rentabilidade.
- Eficiência Operacional: O dimensionamento de lotes contribui para o uso eficiente dos recursos da organização, como mão de obra e equipamentos.
- Satisfação do Cliente: Ao evitar falta de produtos ou excesso de *stock*, o dimensionamento de lotes melhora a capacidade de atender às demandas dos clientes de forma consistente.

Segundo (Drexel & Kimms, 1997), o estudo sobre dimensionamento de lotes começou pelo modelo da Quantidade Económica a Encomendar (QEE) ou também conhecido por *Economic Order Quantity* (EOQ). Este modelo considera o processo de produção de nível único sem restrições de capacidade, tornando-o um problema de um único produto. Estas suposições tornam o modelo muito restritivo, logo outros modelos foram desenvolvidos para preencher as falhas do mesmo.

### Classificação de problemas de dimensionamento e lotes

Dependendo das características de cada problema, ao longo dos anos foram desenvolvidos inúmeros métodos para resolver problemas de dimensionamento de lotes. Estes problemas

podem abordar o mais simples dos tipos de problemas, o *Uncapacitated Single Item Lot Sizing Problem* (USILSP), ou o mesmo tipo de problema, mas considerando restrições de capacidade, o *Capacitated Single Item Lot Sizing Problem* (CSILSP). O *Single Item Lot Sizing Problem* (SILSP) tem especial interesse devido à sua relativa simplicidade e a sua importância como subproblema para resolver problemas de lotes mais complexos (Brahimi et al., 2006).

O SILSP pode ser definido como um problema onde existe uma procura variável no tempo por um determinado produto ao longo de um horizonte de planeamento com  $T$  períodos. O objetivo é determinar os períodos em que ocorrerá a produção e as quantidades que deverão ser produzidas nesses períodos. A produção deve satisfazer a procura e minimizar os custos totais. Os custos a ter em conta são o custo de produção por unidade, o custo de *setup* e custos de posse do *stock*. O problema pode ainda ser estendido de forma a ter em conta limitações de capacidade (Brahimi et al., 2017).

Os modelos que preencheram o vazio do modelo EOQ foram o *Economic lot scheduling Problem* (ELSP) aplicado a problemas com restrições de capacidade e múltiplos artigos (Rogers, 1958, Pochet & Wolsey, 2006). O modelo de Wagner-Whitin (Wagner & Whitin, 1958) foi outro avanço nas técnicas de *Lot Sizing Problem* (LSP) pois calcula a quantidade ótima. Este método foi sendo modificado e melhorado por outros autores como (Evans, 1985, Aggarwal & Park, 1993, Wagelmans et al., 1992) e outros.

As próximas gerações de modelos combinaram abordagens com restrições de capacidade e abordagens dinâmicas. Em (Jans & Degraeve, 2007) os modelos de dimensionamento de lotes foram classificados como o único problema de dimensionamento de lote de itens sem restrições de capacidade, o problema *Capacitated multi-item lot-sizing Problem* (CLSP), o modelo de *Continuous Setup Lot Sizing Problem* (CSLP) e o modelo de dimensionamento de lote discreto e problema de escalonamento (DLSP).

Os autores deste trabalho (Brahimi et al., 2006) também mencionam os critérios de classificação a que os problemas de dimensionamento de lotes são sujeitos, tais como número de máquinas, estágios de produção (níveis), restrições de capacidade e a sua natureza (variável ou fixa) e duração dos períodos de produção. No que toca aos estágios de produção temos problema de nível único e múltiplo níveis. A duração dos períodos pode ser classificada como "*Big time Bucket*" e "*Small time Bucket*", o primeiro com períodos na ordem de horas e o segundo na ordem de alguns dias ou semanas.

### ***Big Time Bucket vs Small Time Bucket***

Em modelos de *Big time bucket*, vários itens podem ser produzidos no mesmo período, na mesma máquina; pelo contrário, em modelos *Small time bucket*, uma máquina só pode produzir um tipo de produto, por período.

No CLSP (*Big Bucket*), vários itens diferentes podem ser produzidos em cada período e há uma capacidade de produção limitada em cada. A principal diferença entre este modelo USILP é a adição das restrições de capacidade e um índice extra para identificar o item. O objetivo é o

mesmo do modelo mencionado anteriormente, minimizar os custos de produção, *stock* e custo de *setup*.

O CSLP (*Small Bucket*) usa uma nova variável de decisão que sinaliza *setups*, isto é, se a máquina precisar ser configurada para um novo artigo. Este modelo também tem uma restrição que impõe que no máximo, apenas um tipo de produto pode ser produzido em cada período.

Finalmente, o DLSP (*Small Bucket*) funciona com uma abordagem do tipo tudo ou nada para a produção. Possui estrutura semelhante ao CSLP, exceto que as restrições de capacidade se tornam uma igualdade, se houver qualquer produção em um período, nesse período deve ser utilizada a capacidade total. Existem diversas extensões destes problemas que são avaliadas por (Jans & Degraeve, 2007).

### **Fatores de Complexidade**

Os problemas de dimensionamento de lotes podem ser modelados conforme as características de um sistema de produção específico. Segundo (Karimi et al., 2003), a complexidade destes problemas está ligada diretamente às seguintes características:

- **Horizonte de Planejamento:** Pode ser finito (com procura dinâmica) ou infinito (para procura estacionária), observado de forma contínua ou discreta, dividindo os problemas em *big-bucket* ou *small-bucket*.
- **Procura:** Divide-se em estática ou dinâmica, determinística ou estocástica, independente ou dependente. Procura estática é constante, enquanto a dinâmica varia. A procura determinística tem valores conhecidos, enquanto a estocástica é baseada em probabilidades. Procura independente não depende de outros itens, mas a dependente está ligada a itens anteriores.
- **Número de Estágios de Produção:** Resulta em problemas de nível único ou multinível. Problemas de nível único envolvem uma única etapa de produção com procura independente. Multinível tem várias etapas de produção, com procura dependente.
- **Número de Produtos:** O número de produtos finais afeta a complexidade da modelagem e resolução.
- **Restrições de Capacidade ou Recursos:** Os problemas são não capacitados (sem restrições) ou capacitados (com restrições de capacidade).
- **Estrutura dos Custos e Tempos de Preparação:** Pode ser simples (independente da sequência) ou complexa (dependente da sequência), incluindo preparação entre produtos da mesma família, entre famílias e *carryover* (tempo de preparação mantido entre períodos).
- **Políticas de Atendimento da Procura:** Envolvem permitir *stock* negativo (*backlogs*), atrasos no atendimento da procura ou perdas nas vendas, influenciando custos e restrições.
- **Deterioração de Itens:** Quando há riscos de deterioração, restrições de tempo no horizonte de planejamento são consideradas.

### Integração de problemas de Dimensionamento de lotes

Segundo (Alves et al., 2021) podemos ter uma integração entre o problema de dimensionamento de lotes (problema de nível superior) e o problema de escalonamento (problema de nível inferior), o problema de nível superior leva em consideração a procura e planeamento de médio prazo, e os objetivos são atender a procura enquanto são satisfeitas as restrições dos sistemas de produção.

O objetivo em problemas de escalonamento é atribuir e sequenciar tarefas em uma ou múltiplas máquinas para minimizar o tempo máximo de conclusão de todos os trabalhos ou a soma dos atrasos dos trabalhos (Brahimi et al., 2017).

Alguns destes modelos basearam-se na integração de problemas de dimensionamento de lotes com problemas de escalonamento, como é mencionado e abordado por (Copil et al., 2017). Alguns autores (Guimarães et al., 2014) ainda propõem um sistema de classificação para este tipo técnicas representado na Figura 2.

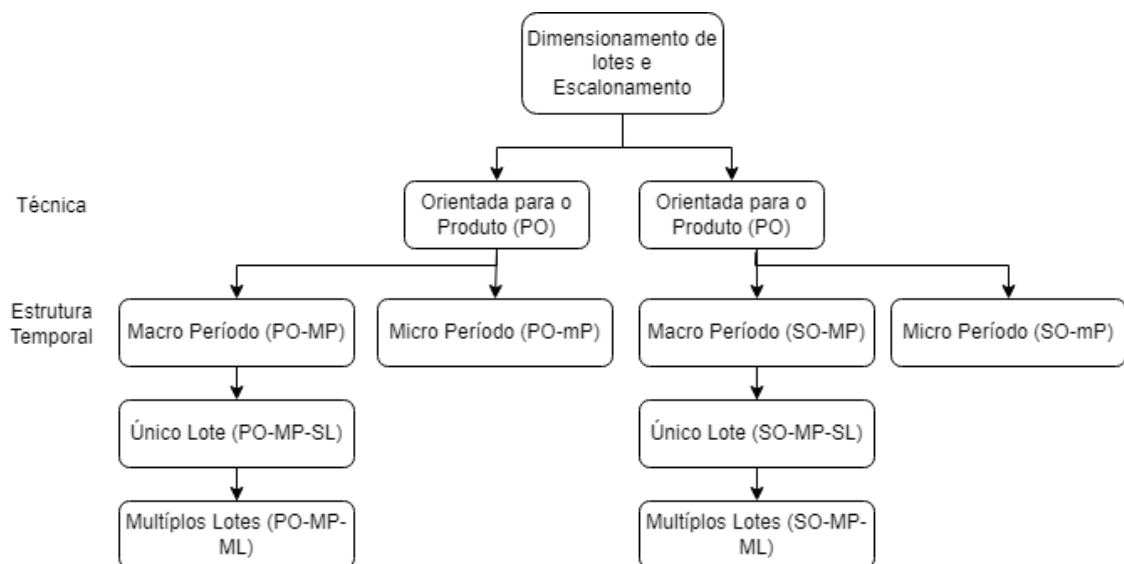


Figura 2 - Sistema de Classificação das técnicas [modificado de: (Guimarães et al., 2014)]

Modelos integrados de planeamento de produção com outros tipos de decisões nas empresas (por exemplo, distribuição, agendamento, etc.) geram lucros maiores ou reduzem custos. Tendo em conta a pesquisa de (Brahimi et al., 2017), o problema de dimensionamento de lotes pode ser integrado com outros fatores além do escalonamento, tais como:

- Localização do armazém, transporte e rotas dos veículos;
- Aceitação de pedidos e seleção de mercado;
- Seleção de fornecedores;
- Coordenação da cadeia de suprimentos.

## 2.2. Métodos de dimensionamento de lotes USILSP

As pesquisas sobre modelos de dimensionamento de lotes começaram com a fórmula clássica de dimensionamento de lotes QEE, Quantidade Económica a Encomendar, (Erlenkotter, 1990) que determina a quantidade a encomendar para um artigo individual, considerando um equilíbrio existente entre os custos de *stock* e os custos de lançamento. Este método também é aplicado na gestão das operações, sendo denominado de *Economic Order Quantity* (EOQ).

Diversos trabalhos abordam o problema de dimensionamento de lotes não capacitado, considerando a programação da produção (ou aquisição) de um único item com taxas de procura variáveis no tempo. (Brahimi et al., 2006) apresenta uma revisão sobre este problema.

Muitas contribuições estão presentes na literatura sobre todas as variantes possíveis do problema de dimensionamento de lotes, alguns autores dedicaram publicações inteiras à classificação de problemas de dimensionamento de lotes (Bahl et al., 1987); (Brahimi et al., 2006)

No entanto, de acordo com (Baciarello et al., 2013), a técnica de dimensionamento de lotes mais comumente usada em contextos industriais reais é a lote por lote (LFL), apesar de, desde a década de 1960, a pesquisa ter se concentrado em encontrar técnicas de melhor desempenho.

(Pochet & Wolsey, 2006) apresentam um modelo do problema de dimensionamento de lotes não capacitado, que inclui na função objetivo os custos de produção. Estes autores também apresentam uma formulação onde o modelo é estendido para vários produtos. Neste caso temos um modelo não capacitado, multiproduto com procura dinâmica. Como não há restrições de capacidade, o problema pode ser decomposto em  $N$  subproblemas independentes, um para cada item a ser produzido. Desta forma cada subproblema pode ser resolvido independentemente. O modelo não capacitado é um subproblema essencial para o planeamento das operações, pois é resolvido repetidamente para cada item da estrutura do produto nos modelos de dimensionamento de lotes multi artigo ou multi estágio.

## 2.3. Métodos Heurísticos

Neste capítulo serão abordados alguns métodos básicos utilizados na resolução dos problemas de dimensionamento de lotes com um único item sem restrição de capacidade.

Serão apresentados os métodos heurísticos: *Lot for Lot* (LFL), *Economic Order Quantity* (EOQ), *Silver-Meal* (SM), *Parts Period Balancing* (PPB), visto que são mais conhecidos pela sua utilização em ambientes de produção.

Métodos Heurísticos ou Heurísticas são abordagens de solução que se baseiam em regras práticas, intuição e experiência para encontrar soluções aproximadas para problemas de otimização, como o dimensionamento de lotes. Eles são projetados para encontrar boas soluções em tempo razoável, embora não garantam a solução ótima. Os métodos heurísticos são frequentemente usados quando os métodos exatos são computacionalmente intensivos ou impraticáveis (Florim et al., 2019); (Karimi et al., 2003).

### 2.3.1. Lot for lot (LFL)

Uma das heurísticas mais simples de utilizar no que toca ao dimensionamento de lotes é o (LFL). Nesta heurística a necessidade líquida do período em questão deve ser atendida pela quantidade produzida no mesmo período. O *stock* final é considerado sempre nulo. São necessárias fazer encomendas ou lotes de produção em todos os períodos, o que torna o método não muito apelativo em termos de custo-benefício, nomeadamente em situações em que os custos de *setup* são altos (Edward A. Silver et al., 1998).

### 2.3.2. Quantidade económica (EOQ)

O EOQ é um dos primeiros e mais conhecidos conceitos na teoria de gestão de *stocks*. Proposto por F. W. Harris em 1913 (Erlenkotter, 1990), este método também é conhecido como método de Wilson porque R. H. Wilson foi quem inicialmente aplicou este modelo. Este modelo também foi depois modificado e utilizado para Problemas ELSP (Rogers, 1958); (Reçus, 1964), problemas que combinam o dimensionamento e escalonamento dos lotes,

O modelo de (Harris, 1990) pressupõe um processo produtivo de nível único, sem restrições de capacidade, procura constante ao longo de um horizonte de planeamento infinito, sem descontos de quantidade e sem ruturas de *stock* (José Fernando Gonçalves, 2010).

Embora esta fórmula não seja eficiente para descrever a maioria dos sistemas de planeamento da produção, pois não considera as limitações de capacidade, ela ainda é útil nos sistemas de *stock* para determinar quantidades adquiridas de materiais.

O EOQ (EQ. 1) envolve o cálculo prévio da procura média ( $D$ ) e saber o custo de posse ( $h$ ) e custo de encomenda ( $a$ ).

$$EOQ = \sqrt{\frac{2 * D * a}{h}} \quad EQ. 1$$

### 2.3.3. Ordens periódicas (POQ)

O método (POQ) possui quantidades constantes e o intervalo de tempo em que os pedidos serão feitos é determinado através da utilização da fórmula do EOQ (EQ. 1) e do POQ (EQ. 2). O tamanho do lote deve satisfazer as necessidades até o próximo pedido (Edward A. Silver et al., 1998).

O início deste método é semelhante ao EOQ. Primeiramente é necessário calcular o EOQ e de seguida com esse valor calcular o valor de POQ, o resultado será o número de períodos para o qual uma encomenda deve satisfazer as necessidades líquidas.

$$POQ = \frac{EOQ}{D} \quad EQ. 2$$

### 2.3.4. Parts Period Balancing (PPB)

Gorham apresentou a heurística PPB em 1968, nesta o procedimento inicia-se no primeiro período do plano de planeamento e evolui em direção ao período final. Em cada período é calculado o custo de *stock* para a situação em que se encomenda o suficiente para o primeiro, segundo e terceiro período e sucessivamente até que se encontre a situação em que os custos de *stock* estejam o mais próximo possível do custo de lançamento (Edward A. Silver et al., 1998, Florim et al., 2019).

### 2.3.5. Silver-Meal (SM)

A heurística (SM) seleciona a quantidade a encomendar calculando a quantidade de reabastecimento que minimiza os custos totais por unidade de tempo. Normalmente esta heurística apresenta bons resultados, muitas vezes próxima do ótimo. A heurística dos autores Edward Silver e Harlan Meal foi desenvolvida e mais tarde modificada de forma a combater cenários onde a heurística tendia a obter piores resultados, essas situações eram (Florim et al., 2019, José Fernando Gonçalves, 2010):

- Quando a procura diminui rapidamente ao longo de vários períodos;
- Quando há vários períodos sem demanda.

Com base na (EQ. 3) onde  $T$  representa o número de períodos de tempo para os quais uma encomenda durará, irá ser calculado os custos para  $T=1$ ,  $T=2$  e assim sucessivamente até se verificar a condição da (EQ. 4), quando os custos são menores (José Fernando Gonçalves, 2010).

$$CT(T) = \frac{\text{Custo Enc.} + \text{Custo de posse até fim do período } T}{T} \quad \text{EQ. 3}$$

$$CT(T + 1) > CT(T) \quad \text{EQ. 4}$$

## 2.4. Métodos Exatos

Os métodos heurísticos são simples e eficazes, mas podem apresentar valores longe do valor ótimo. Por outro lado, os métodos exatos são abordagens matemáticas que procuram encontrar a solução ótima. Eles envolvem a exploração sistemática de todas as possíveis combinações de parâmetros, garantindo que a solução encontrada seja a melhor possível em termos de custos totais (Florim et al., 2019).

### 2.4.1. Algoritmo Wagner-Whitin (WW)

Este algoritmo foi introduzido por (Wagner & Whitin, 1958) e é um modelo dinâmico de programação que visa obter as melhores quantidades do pedido, ou seja, o tamanho do lote que minimiza os custos totais. O trabalho computacional do algoritmo é reduzido porque a solução ótima deve satisfazer as seguintes propriedades (Florim et al., 2019); (José Fernando Gonçalves, 2010):

- Um pedido só chega quando o nível de *stock* é zero;
- Existe um limite máximo para o número de períodos para qual um pedido durará.

As variáveis do algoritmo consistem em:

$C_t, N+1$  = Custo do melhor conjunto de encomendas que cobre a procura desde o início do período  $t$  até o início do período  $N+1$ .

Notar que o  $C_i, i = 0$  e que o início de um período coincide com o fim do anterior.

$E_t, k$  = Custo de uma encomenda que chega no período  $t$  e que satisfaz a procura até o início do período  $k$ .

O Custo da solução ótima é dado por  $C_t, N + 1$

$$C_{t,N+1} = \min_{t < k \leq N+1} \{E_{t,k} + C_{k,N+1}\} \quad t = N, \dots, 2, 1 \quad EQ. 5$$

O algoritmo começa no último período,  $N$ , e repete até alcançar ao período 1. Para cada período  $t$ , é selecionado o valor de  $k$  que corresponde ao menor custo total, ou seja, o número de períodos para os quais o pedido será feito no período  $t$  (José Fernando Gonçalves, 2010).

## 2.5. Avaliação da Performance de heurísticas

Este ponto desempenha um papel crucial na análise e comparação das abordagens utilizadas para o dimensionamento de lotes em diferentes cenários. A avaliação da performance é fundamental para determinar quais os métodos mais eficazes e eficientes em atender às necessidades específicas das empresas. Este subcapítulo terá como foco oferecer uma visão abrangente das vantagens e limitações de cada abordagem. Esta a avaliação será baseada maioritariamente nos trabalhos de (Florim et al., 2019, Baciarello et al., 2013, Simpson, 2001).

Na pesquisa dos autores (Florim et al., 2019) são abordados 9 heurísticas, das quais LFL, EOQ, SM e PPB já foram abordadas. Para a simulação de diversos cenários com variação de procura e de picos de sazonalidade, foram geradas combinações entre a procura segundo uma distribuição normal de  $\mu = 250$  e desvio padrão entre  $10 \leq \sigma \leq 80$  e para a sazonalidade com uma média de 250 e com os seguintes indicadores de sazonalidade [66;99;133;166;199;233]. Os resultados obtidos para cada técnica foram comparados à solução ótima obtida com o método de Wagner e estão resumidos na Tabela 1.

Tabela 1 - Resumo dos desvios para métodos de dimensionamento de lotes com aumento da variância e índices de sazonalidade [modificado de:(Florim et al. 2019)]

Método de Dimensionamento de lotes	Erro para menores variações da procura (10%)	Erro para maiores variações da procura (80%)	Desvio de erros em relação a 10% e 80% da demanda variações	Erros para pequenos índices de sazonalidade (66)	Erros para maiores índices de sazonalidade (233)	Desvio de erros em relação a 66 e 233 sazonalidade índices
EOQ	27%	31%	4	19%	36%	17
LFL	14%	17%	3	16%	34%	18
F&C	14%	10%	-4	15%	4%	-11
MOM	5%	3%	-2	4%	6%	2
POQ	0%	2%	2	2%	4%	2
var.EOQ	0%	4%	4	2%	0.3%	-1,7
S-M	0%	2%	2	1%	17%	16
LUC	0%	4%	4	1%	11%	10
PPB	0%	2%	2	1%	1%	0

### Influência da Procura e Sazonalidade

Com base nos resultados obtidos pelos autores desta pesquisa (Florim et al., 2019), estes verificou-se que o aumento da variabilidade da procura ou sazonalidade afeta todos os métodos, aumentando o erro em relação ao ótimo. Os autores do trabalho concluíram que, dos métodos abordados no subcapítulo anterior, o EOQ e o LFL resultaram em um maior desvio em relação à solução ótima em cenários que envolvam uma grande variação na procura. Uma ocorrência que também se verificou noutros métodos, como POQ, S-M, PPB e outros. Esta correlação entre variabilidade da procura e desvio da solução ótima é abordada também por (Baciarello et al., 2013)

No que toca à sazonalidade os métodos cujos resultados obtiveram um erro maior foram o EOQ, LFL e o S-M. Neste parâmetro os autores identificaram que na técnica LFL, o erro aumentava de forma linear com o aumento do indicador de sazonalidade, ao contrário das técnicas EOQ, S-M, PPB, F&C, LUC e MOM que apresentavam irregularidade com o aumento sazonalidade.

O trabalho de (Simpson, 2001) por outro lado, avalia a influência do número de períodos no custo total de cada método e compara com a situação de *rolling horizon*, abordado também por (Bookbinder & H'ng, 1986) e (Blackburn & Millen, 1980), onde a procura é conhecida para um determinado número de períodos no futuro. O número de períodos futuros para os quais a demanda é deterministicamente conhecida é chamada de "janela de previsão". Neste trabalho Simpson avaliou 9 técnicas, onde para o número de períodos a considerar obteve o gráfico da Figura 3.

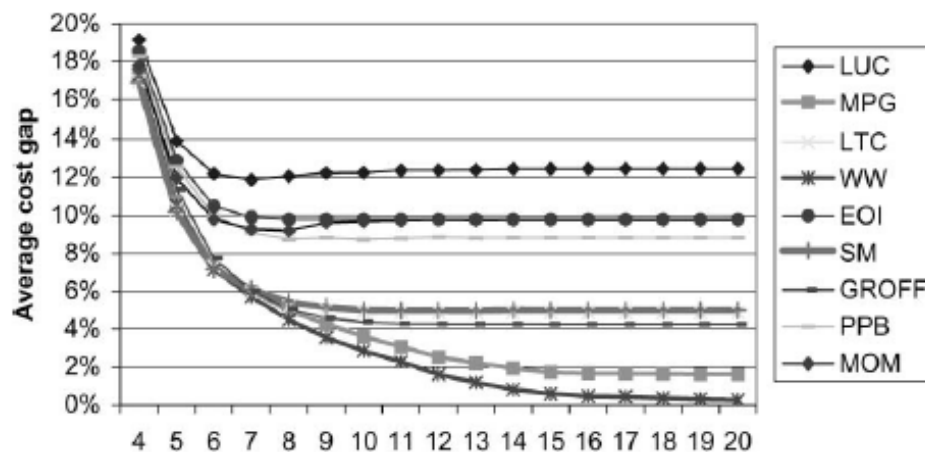


Figura 3 - A diferença percentual média entre o custo total do cronograma e a solução de referência de 300 períodos para vários tamanhos de horizontes [fonte: (Simpson 2001)]

Os autores demonstraram que todas as regras são em grande parte equivalentes e apresentam uma penalidade de custo substancial quando o horizonte de planeamento é mais curto, melhorando o desempenho de custos à medida que o horizonte se estende e para cenários com baixa variabilidade na procura e sem períodos sem encomendas, os resultados dos métodos heurísticos não variam significativamente. Também é abordada a influência dos métodos sob condições de *rolling horizon* (Simpson, 2001).

## 2.6. Inteligência Artificial

O mundo enfrenta uma nova revolução industrial, impulsionada pelos grandes avanços no ramo da inteligência artificial (IA), desde automóveis com condução autónoma a IA capazes de diagnosticar casos de Covid (Ludermir, 2021), (Li et al., 2020).

Alguns dos conceitos de IA já foram conceptualizados nos anos 50, mas apenas se tornaram possíveis de praticar e desenvolver nos últimos anos, devido ao poder computacional necessário e da disponibilidade de dados (Ludermir, 2021). O aumento do poder computacional nas últimas duas décadas foi um dos fatores que condicionou o avanço das técnicas de IA, mas já desde muito antes foi colocada a questão sobre a capacidade de um computador aprender. Na década de 1950, Alan Turing criou um teste conhecido hoje como teste de Turing. O objetivo do teste era determinar se uma máquina é capaz de exibir comportamento inteligente indistinguível de um humano. O teste consistia em um interrogador que faz perguntas a um humano e a uma máquina, sem saber qual é qual, se o interrogador não conseguir distinguir entre as respostas da máquina e as do humano, a máquina é considerada inteligente. Esta definição para determinar a inteligência de uma máquina acabou por ser muito influente e criticada por outros autores na área. (TURING, 1950). Nos últimos anos temos visto um grande crescimento deste tipo de aplicação (*chatbots*) que são cada vez mais utilizados como assistentes automáticos em universidades e lojas on-line.

Nos dias de hoje, a inteligência artificial é uma área muito complexa e abrangente, integrando diversas tecnologias que utilizadas em conjunto aumentam a eficiência de certas tarefas, análises e decisões que normalmente seriam desempenhadas por humanos (Chan-Olmsted,

2019). Dependendo do tipo de tarefa, repetitiva e não repetitiva, e modelo de atividade, comportamental e cognitiva, as tecnologias suportadas por IA assumem as classificações da Tabela 2 (Hoyer et al., 2020):

Tabela 2 - Uma nova tipologia/classificação de novas tecnologias alimentadas por IA [modificado de: (Hoyer et al., 2020)]

Tarefa	Comportamental	Cognitiva
Repetitiva	Robô (e.g., Roomba)	Assistente Virtual, <i>Chatbot</i> (e.g., Alexa)
Não Repetitiva	Robô (e.g., Vector, Cozmo)	Assistente Criativo (e.g., MuseNet)

Devido às oportunidades que a IA pode criar na redução de custos, otimização de recursos e como consequência no aumento dos lucros (Kar et al., 2019), na última década foi visto um aumento nos investimentos nesta tecnologia, devido à sua capacidade em executar uma grande variedade de tarefas. Este aumento é visível nos dados da Figura 4.

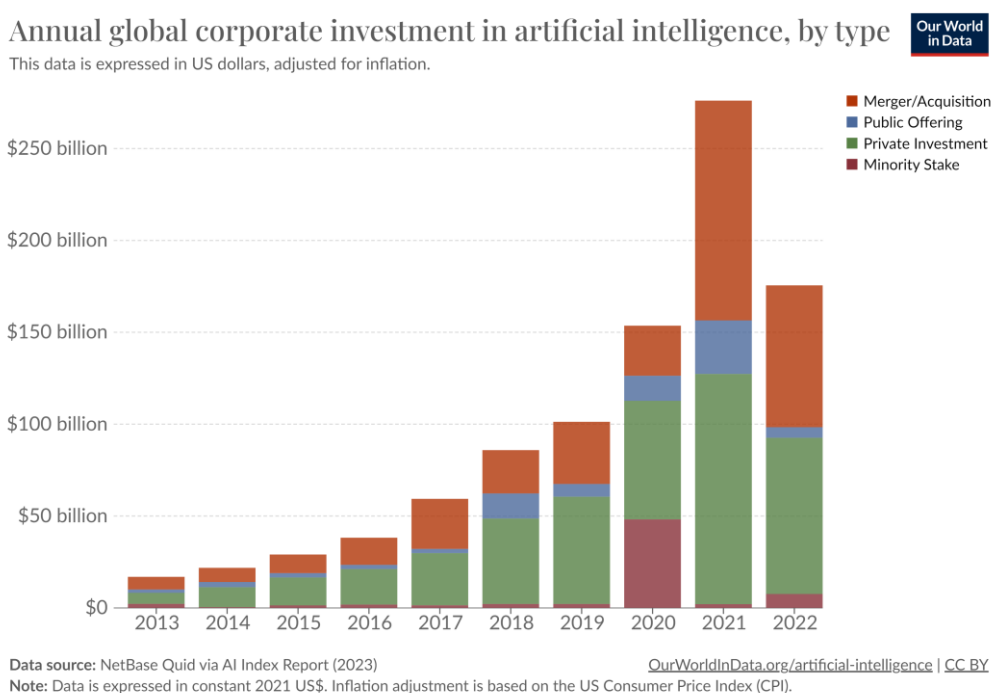


Figura 4 - Investimento corporativo global anual em inteligência artificial, por tipo [fonte:(Charlie Giattino, Edouard Mathieu, 2023)]

Com o aumento do investimento na IA, funções desempenhadas por trabalhadores serão afetadas. Um estudo em 2022 estimou que 34% das tarefas no local de trabalho eram desempenhadas por máquinas e que esse valor tendia a aumentar para 43% no ano de 2027. Este estudo também revela que das empresas inquiridas, 75% tencionam adotar IA e metade considera que a adoção de IA nas suas atividades resultara num aumento da produtividade. Por outro lado apenas 25% das empresas acreditam que a IA irá originar a perda de empregos (Wolrd Economic Forum, 2023).

### 2.6.1. Machine learning

*Machine Learning* (ML) é uma subárea da IA que consiste na capacidade de os computadores ou as máquinas aprenderem determinada tarefa sem estarem explicitamente programadas, ou seja, sem a necessidade de interação humana. ML explora o estudo e a construção de algoritmos que podem aprender com a experiência em relação a alguma tarefa e alguma medida de desempenho. Se seu desempenho em X, medido por Y, melhorar com a experiência então o programa é chamado de programa ML (K Kar, 2018, LeCun et al., 2015, Das et al., 2015).

ML divide-se em 3 áreas principais: *Reinforcement Learning*, *Supervised Learning* e *Unsupervised Learning* (Kar et al., 2019, Sheikh et al., 2023) .

Em *Supervised learning* o programa recebe dados previamente filtrados e classificados. Quanto maior a quantidade fornecida ao algoritmo, maior a facilidade e exatidão na previsão do resultado. O algoritmo é treinado com os dados de entrada e depois testado para ver se consegue aplicar corretamente os filtros a novos dados. O *unsupervised learning* supervisionado não tem etapa de treinamento e, portanto, o algoritmo precisa procurar padrões nos dados por si só. É alimentado com grandes quantidades de dados não filtrados, nos quais começa a reconhecer padrões por si próprio. No caso do *reinforcement learning* o algoritmo é treinado para seguir determinadas estratégias por meio de um sistema de *feedback* positivo e negativo. É comumente aplicado em contexto de jogos, desde o clássico PacMan a jogos de cartas como poker. O algoritmo tem como objetivo otimizar o valor da pontuação, correlacionando todos os tipos de ações e desenvolvendo uma estratégia ideal com objetivo de obter a maior pontuação possível (Kar et al., 2019, Sheikh et al., 2023).

### 2.6.2. Deep Learning

O *Deep learning* (DL) faz parte de uma família mais ampla de métodos de ML baseados na aprendizagem de representações de dados, em oposição a algoritmos específicos de tarefas. Esses algoritmos tentam aprender automaticamente vários níveis de abstração, representação e informações a partir de um grande conjunto de dados. As técnicas de DL fizeram avanços significativos em muitas áreas, incluindo reconhecimento de fala, reconhecimento de texto, leitura labial, diagnóstico auxiliado por computador, reconhecimento facial, etc (K Kar, 2018).

Atualmente, *Artificial Intelligence* (AI), *Machine Learning* (ML) e *Deep Learning* (DL) são três termos populares que às vezes são usados de forma intercambiável para descrever sistemas ou *software* que se comporta de forma inteligente. Na Figura 5 é ilustrada a posição do DP, comparando com ML e IA. O DL é uma parte de ML, bem como uma parte da IA de área ampla. Em geral, IA incorpora comportamento humano e inteligência às máquinas ou sistemas, enquanto ML é o método para aprender com dados ou experiência, que automatiza a construção de modelos analíticos. O termo “*Deep*” na metodologia de DL refere-se ao conceito de múltiplos níveis ou estágios através dos quais os dados são processados para construir um sistema baseado em dados modelo (Sarker, 2021).

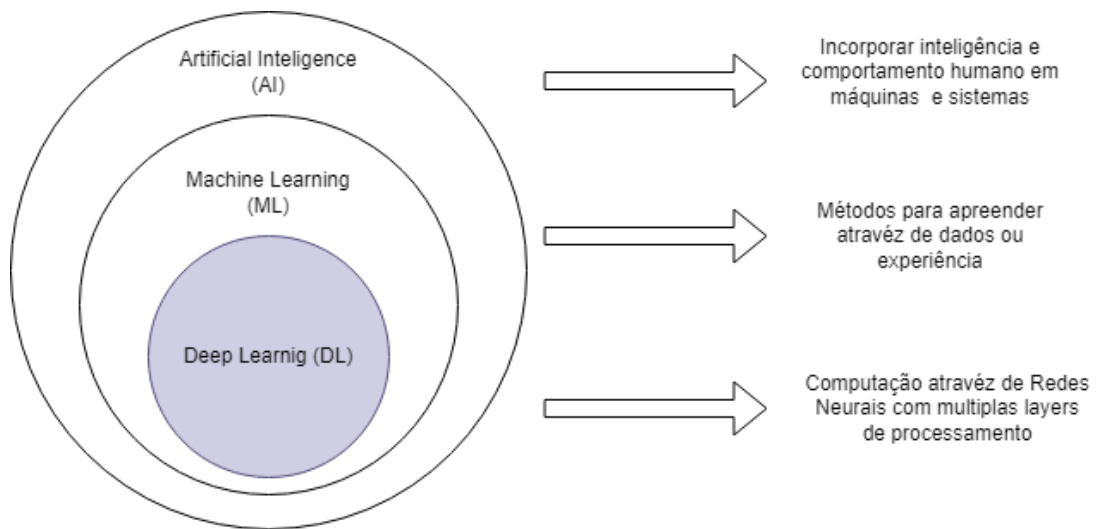


Figura 5 - Uma ilustração da posição do DL, comparando com ML e IA [modificado de:(Sarker 2021)]

### 2.6.3. Redes Neurais

Uma das técnicas de ML que tem tido sucesso em resolver muitos problemas são as *Artificial Neural Networks* (ANN). Estas são modelos matemáticos que se inspiram nas estruturas neurais biológicas e que têm a capacidade computacional adquirida por meio de aprendizagem. Uma das redes neurais mais antigas baseadas no sistema cerebral humano é o *Perceptron*, que possui uma arquitetura muito simples constituída por duas camadas: camada de entrada e camada de saída que estão diretamente conectadas entre si. Este tipo é mais adequado para classificar padrões lineares separáveis. (Ludermir, 2021, Arora et al., 2018, Sarker, 2021).

Para resolver padrões complexos existem ANNs que usam uma camada extra, denominada de *hidden layer*. Estas redes neurais são geralmente denominadas *Deep Neural Networks* (DNN). A diferença entre a estrutura de uma rede ANN comum e uma DNN é demonstrada no esquema da Figura 6 (LeCun et al., 2015, Arora et al., 2018). As DNN são caracterizadas pelo fator da sua “profundez”, que por sua vez está diretamente relacionada ao número de *hidden layers* da rede neural (Arora et al., 2018). Atualmente existe uma grande variedade de *Deep Neural Networks* que podem ser selecionadas dependendo da aplicação e do domínio do problema, algumas delas são amplamente utilizadas para detecção de imagens, processamento de linguagem natural e classificação de dados (LeCun et al., 2015, Arora et al., 2018).

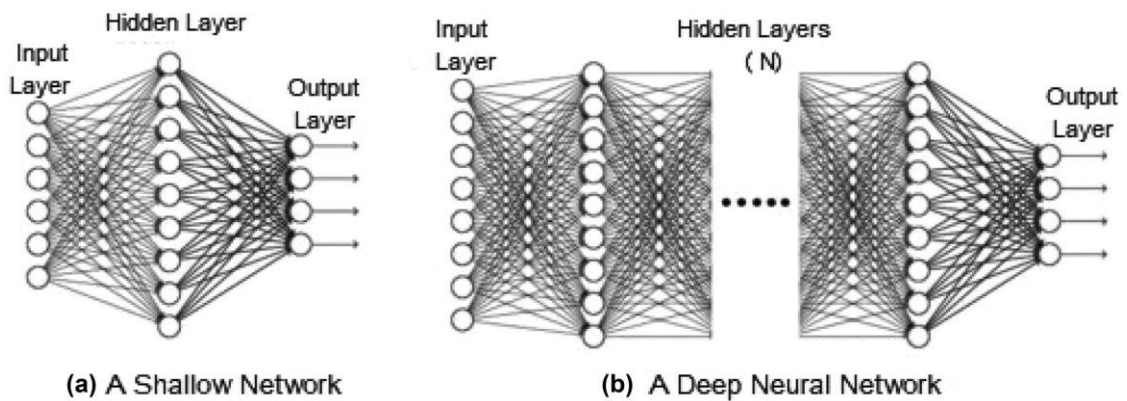


Figura 6 - Uma arquitetura geral de a) rede superficial com uma camada oculta e b) uma rede neural profunda com múltiplas camadas ocultas [fonte:(Sarker, 2021)]

#### 2.6.4. Algoritmos Genéticos

*Genetic Algorithms* (GA) são métodos eficientes e não determinísticos de pesquisa estocástica que utilizam as teorias da evolução e da seleção natural para resolver um problema dentro do espaço de soluções complexas. GA pertence á classe de *evolutionary algorithms*. Esta classe pode ser dividida em quatro tipos principais de algoritmos: GA, *Genetic Programming*, *Evolution Strategies* e *Evolutionary Programming*. Estes tipos de algoritmos foram aplicados a diversos tipos de problemas, quer para problemas de localização, *stock*, produção, programação, distribuição ou cronogramas (Pisinger & Ropke, 2007, Bhoskar et al., 2015).

Os GAs são inspirados pelo processo de evolução natural, sendo assim, ao longo de muitas gerações, as populações evoluem de acordo com o princípio da seleção natural e da “sobrevivência do mais apto”. Ao imitar esse processo e por meio de codificação adequada, os GAs usam um conjunto de soluções, a partir das quais, através do melhoramento seletivo e estratégias de recombinação, soluções cada vez melhores podem ser alcançadas (Guner Goren et al., 2010).

A principal vantagem dos GA é o fator de crescerem linearmente com o tamanho crescente do problema e, portanto, é serem mais eficiente na resolução de problemas complexos, especialmente quando comparados com métodos exatos (Bhoskar et al., 2015).

## 2.7. Plataforma ChatGPT

O ChatGPT (*Chat Generative Pre-Trained Transformer*), propriedade da OpenAI, foi lançado no dia 30 de novembro de 2022 (Chatterjee & Dethlefs, 2023), tornando-se uma sensação cultural em menos de 2 meses (Thorp, 2023), alcançando a marca de 1 milhão de utilizadores em apenas 5 dias. Outras aplicações como o Instagram, Twitter e outras presentes na Figura 7 demoraram meses e até anos (Bansal, 2023).

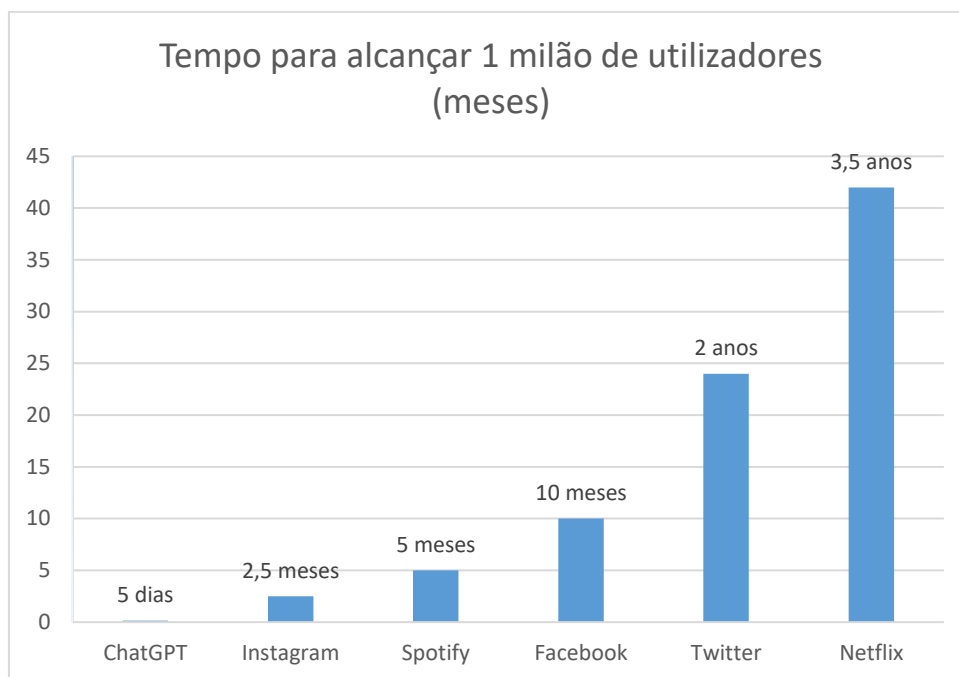


Figura 7 - Tempo para alcançar 1 milhão de utilizadores [modificado de:(Bansal, 2023)]

O ChatGPT é um *Chatbot* com IA com o poder de responder e compreender instruções escritas com linguagem natural (NLP) através de conversação (Cao et al., 2023). Este programa é um LLM (*Large Language Model*) treinado usando técnicas de ML com supervisão em conjunto com técnicas de *Reinforcement Learning From Human Feedback* (RLHF) (OpenAI, 2022). É um modelo robusto de NLP que pode compreender e criar linguagem natural para uma ampla gama de aplicações, incluindo texto, compreensão da linguagem e programas interativos (Gill & Kaur, 2023); (Bockting et al., 2023).

A OpenAI tem lançado vários LLMs (Chatterjee & Dethlefs, 2023) e de momento, novembro de 2023, o ChatGPT tem como base o modelo GPT-3.5, uma versão refinada do modelo anterior, GPT-3 (The Lancet Digital Health, 2023). O que diferencia os modelos um do outro é o facto da versão 3.5 ser capaz de apreender com *feedback* humano (Chatterjee & Dethlefs, 2023).

GPT-4 é a mais recente iteração da série GPT e é o LLM mais poderoso até o momento, oferecendo recursos e capacidades ainda mais avançados do que seus antecessores. Embora a OpenAI ainda não tenha divulgado muitos detalhes técnicos sobre o GPT-4, foi demonstrado que ele supera o ChatGPT, baseado em GPT-3.5, em várias tarefas (OpenAI, 2023), como por exemplo na resolução de diversos exames nas mais diversas áreas de estudo, Figura 8.

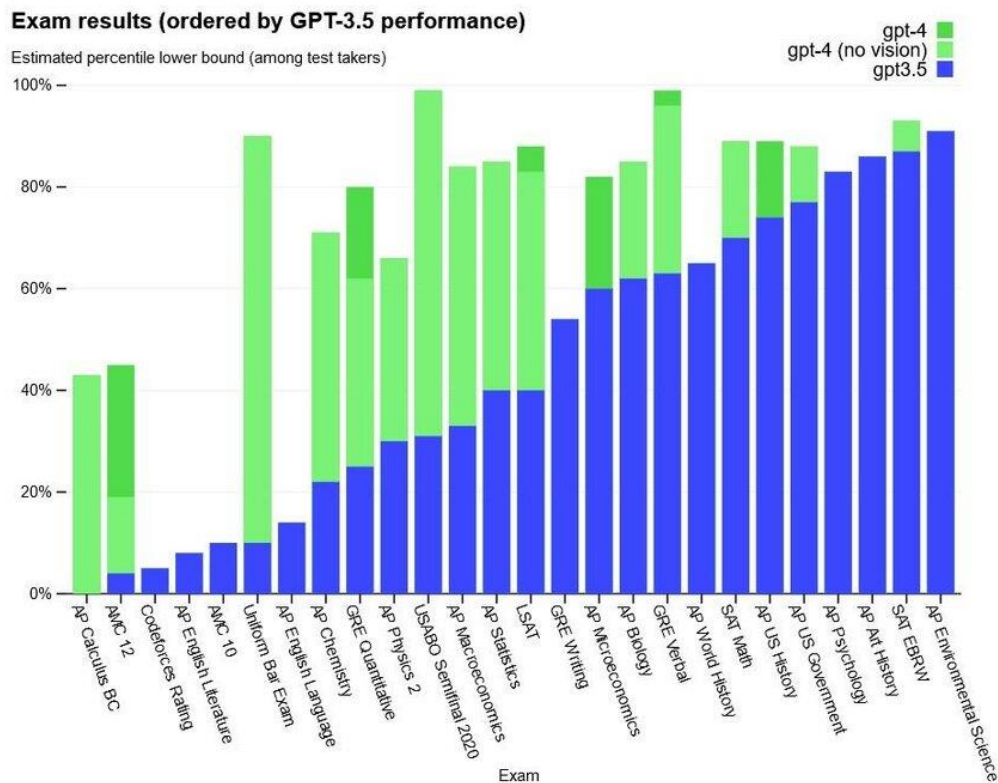


Figura 8 - Resultados do ChatGPT vs GPT-4 em vários exames [fonte:(OpenAI, 2023)]

### Limitações

O ChatGPT não é perfeito e apresenta falhas, a lista de limitações indicada pela OpenAI foca-se no fornecimento de respostas, por vezes, incorretas ou sem sentido, na sensibilidade do *chatbot* a pequenas nuances aleatórias, resultados demasiado extensos e explicativos, incapacidade de o modelo negar resposta a pedidos ambíguos que sejam prejudiciais, proteção de dados da privacidade das pesquisas dos utilizadores. O modelo revela ser discriminatório em determinadas situações devido à influência dos dados nos quais foi treinado, por vezes mostra-se muito repetitivo ao utilizar excessivamente as mesmas frases. Além disso, apresenta lacunas de conhecimento atualizado, uma vez que a versão gratuita teve o seu treinamento até 2021, recentemente alargado até janeiro de 2022 (Liu et al., 2023); (Gordijn & Have, 2023); (OpenAI, 2022); (Chatterjee & Dethlefs, 2023).

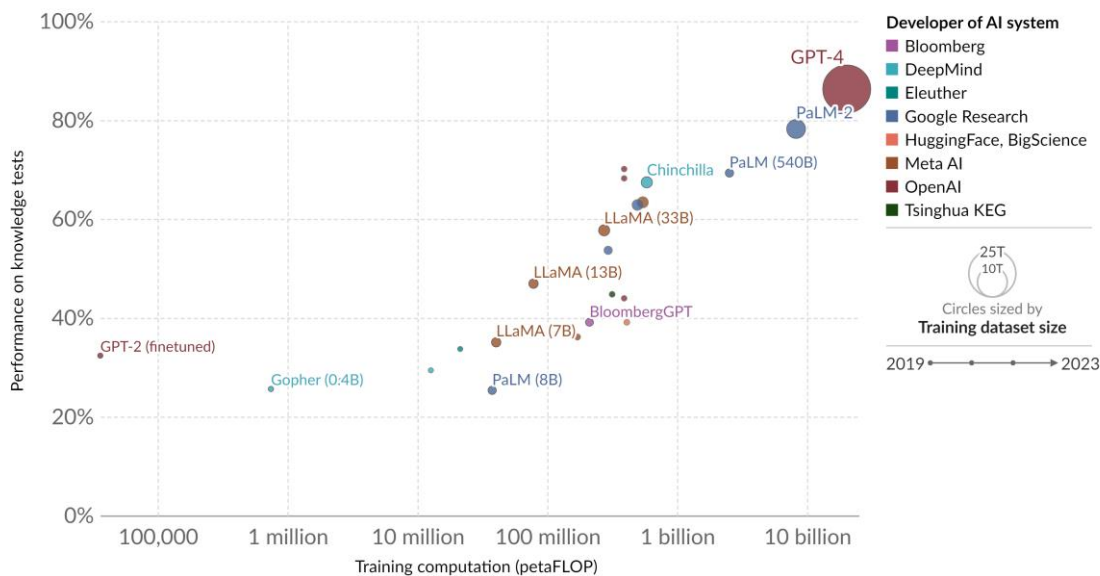
### Outras plataformas

Novas ferramentas de IA aparecem com cada vez mais frequência, por exemplo, a Google, em março de 2023, lançou o seu próprio *chatbot* chamado Bard que, na prática, é um concorrente direto ao ChatGPT. Da mesma maneira, a Microsoft anunciou este ano uma nova versão do seu motor de busca Bing, que incorpora um *ChatBot* que utiliza tecnologia GPT-4 da OpenAI (Mehdi, 2023). Na Figura 9 é mostrada uma comparação entre as diversas tecnologias de modelos de linguagem criados pelas maiores empresas de tecnologia no momento, verificando-se uma superioridade do modelo GPT-4, utilizado na versão paga do ChatGPT e no Bing (Charlie Giattino, Edouard Mathieu, 2023).

## Artificial intelligence: Performance on knowledge tests vs. training computation



Performance on knowledge tests is measured with the MMLU benchmark<sup>1</sup>. Training computation is measured in total petaFLOP, which is  $10^{15}$  floating-point operations<sup>2</sup>.



Data source: Epoch (2023)

[OurWorldInData.org/artificial-intelligence](https://OurWorldInData.org/artificial-intelligence) | CC BY

Note: The values for training computation and dataset size are estimates and come with some uncertainty, especially for models for which only minimal information has been disclosed, such as GPT-4.

Figura 9 - Inteligência artificial: desempenho em testes de conhecimento vs. computação de treino[fonte:(Charlie Giattino, Edouard Mathieu, 2023)]

### Aplicações ChatGPT

Devido à sua adaptabilidade e capacidades superiores, o ChatGPT encontrou uso em uma variedade de aplicações (Liu et al., 2023). Na Tabela 3 são mencionadas algumas maneiras pelas quais a plataforma pode revolucionar os processos, reforçar a colaboração e gerar novas ideias.

Tabela 3 - Exemplos de Áreas de aplicação do ChatGPT

Áreas	Formas de aplicação do ChatGPT	Referências
Estudos na área da Medicina	Providenciar ajuda em análise de diagnóstico de pacientes e indicações de tratamentos; Oferecer conhecimento medicinal aos pacientes de forma simples de entender.	(Lin et al., 2023) (Sallam, 2023) (Fui-Hoon Nah et al., 2023)
Finanças e Negócios	Análise de avaliações de clientes e criação de guias para desenvolvimento de produtos táticas de marketing; Criação de <i>chatbots</i> de suporte ao cliente, desde resposta a clientes, oferta de sugestões e recomendar produtos; Análise de dados de mercados e economia e oferecer recomendação de investimentos.	(Liu et al., 2023) (Fui-Hoon Nah et al., 2023) (Jain et al., 2023)

Áreas	Formas de aplicação do ChatGPT	Referências
Suporte Legal	Ajudar advogados a sumarizar e sintetizar textos legais como leis e acordos; Analisar e prever resultados com base no histórico judicial; Criação de um ChatBot baseado em ChatGPT capaz de providenciar concelhos e responder a questões comuns na área.	(Liu et al., 2023) (Bommarito & Katz, 2022)
Educação	Servir como assistente na busca de informação; Criação de conteúdo para educação; Formulação de problemas; Análise de informação.	(Liu et al., 2023) (Fui-Hoon Nah et al., 2023)
Criação de Conteúdo	Criação de anúncios a partir de descrição; Aumento da eficiência na criação de conteúdo devido a facilidade de utilização; Enriquecer escrita.	(Liu et al., 2023) Fui-Hoon Nah et al. 2023) (Jain et al., 2023)
Programação	Geração de código através de descrições ou especificações; Otimização de código; Ajuda na identificação e correção de erros no código.	(Liu et al., 2023)

## 2.8. Análise crítica da Bibliometria

Como revisão da bibliografia, foi decidido realizar uma análise á bibliometria, com a ajuda do *software* VosViewer e do site da Web Of Science para tratar dados sobre a bibliografia relacionada ao tema de dimensionamento de lotes e inteligência artificial. Foi possível observar que a partir do ano de 2017 houve um crescimento acelerado de pesquisas no âmbito da inteligência artificial, como se pode verificar analisando a Figura 11. Por outro lado, no ramo de dimensionamento de lotes, a sua pesquisa tem sido linear desde o início do século 21, como pode ser visto na Figura 10. A área de estudo com mais publicações no ramo são a engenharia e ciência de computadores, visto na Figura 12, comprovando a utilidade deste problema nas mais várias áreas de estudo.

Na Figura 13 é possível observar as palavras-chave mais utilizadas pelos autores dos artigos sobre a área de dimensionamento de lotes. A palavra mais usada é “*lot Sizing*”, sistema, formulação, heurística, políticas e outras. Essas palavras-chave são de grande relevância, pois as heurísticas são frequentemente usadas para resolver os problemas de dimensionamento dos lotes. Por fim, também é interessante destacar a palavras “*scheduling problem*”, já que, na verdade, o escalonamento é muitas vezes abordado em conjunto com o problema de “*lot Sizing*”, como nos é mostrado nos trabalhos de (Alves et al., 2021) e (Jans & Degraeve, 2007).

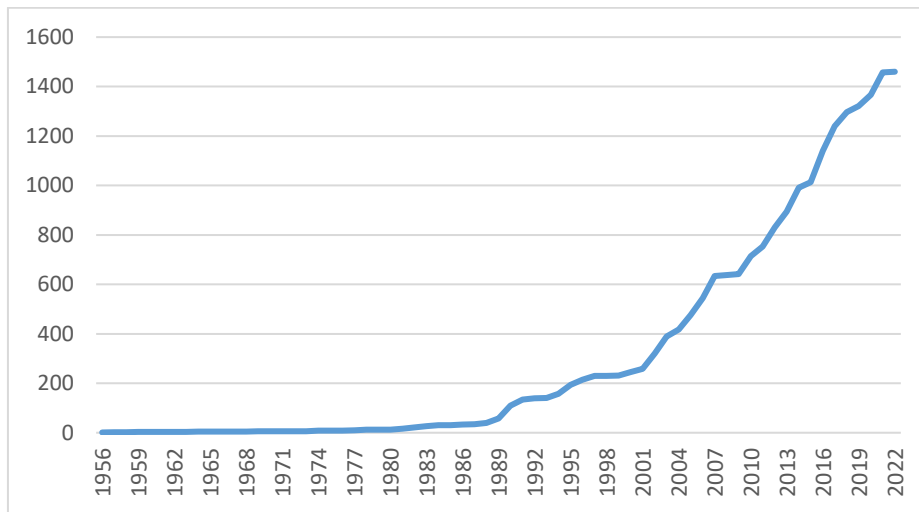


Figura 10 - Publicações na área de Lot-Sizing

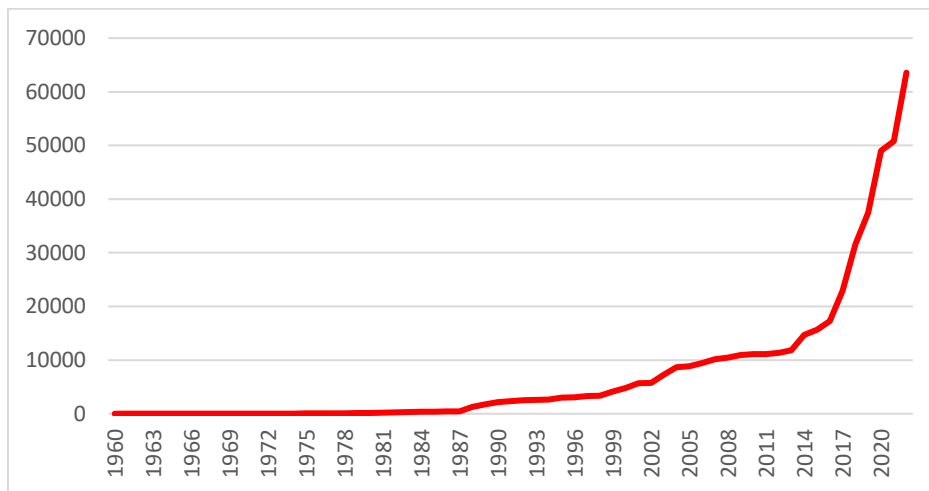


Figura 11 - Publicações na área de AI

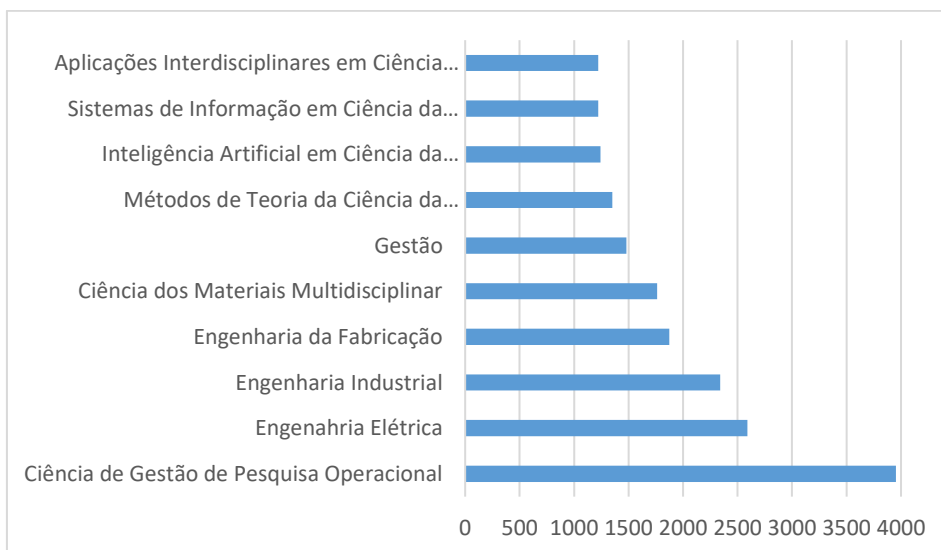


Figura 12 - Áreas de estudo relacionadas com Lot-Sizing



promete redefinir o panorama dos sistemas autónomos, mas também aborda a necessidade crítica de gestão em tempo real. Sendo assim um possível tema de investigação seria “Sistemas de Decisão e Gestão da Produção em Tempo Real”.

## 2.9. Uso de Inteligência Artificial no Dimensionamento de Lotes

O dimensionamento de lotes desempenha um papel central na otimização operacional em ambientes logísticos e de produção. Este capítulo concentra-se em explorar o impacto e utilização da Inteligência Artificial (IA) no problema de dimensionamento de lotes. Notavelmente, esta análise destaca a ausência de aplicação direta do ChatGPT ou outras aplicações semelhantes no dimensionamento de lotes devido a ausência de casos de aplicação, possivelmente devido à atualidade das mesmas, bem como o facto deste tipo de problemas não ser um dos focos da aplicação. Sendo assim a pesquisa de casos de estudo foi direcionada para outros métodos existentes de IA no LSP, como *Genetic Algorithms* (GA), *Artificial Neural Networks* (ANN) e outros algoritmos e técnicas de *Machine Learning* (ML).

**Caso1: A Multi-Neural-Network Learning for Lot Sizing and Sequencing on a Flow-Shop** (Lee et al., 2001)

O primeiro caso de estudo apresentado tem como base a pesquisa de (Lee et al., 2001). Os autores desenvolveram uma estrutura de *multi-Neural Networks*, mais concretamente, duas redes neurais, uma treinada para tomar decisões no dimensionamento de lotes e outra para o sequenciamento das operações, os seja, interligava ambos os problemas de dimensionamento de lotes e escalonamento. Os autores propõem esta alternativa em relação ao sequenciamento de trabalhos sem consideram o dimensionamento dos lotes, afirmando que esta última possa não ser a melhor.

Segundo os autores, as decisões de sequenciamento afetam as decisões de dimensionamento de lote. Devido à natureza interativa do dimensionamento de lote e sequenciamento, os cronogramas gerados por qualquer uns dos problemas é de qualidade de solução inferior ao gerado tendo em ambos os problemas.

O primeiro módulo de quatro presentes na estrutura da *multi-Neural Network* é composto por um Algoritmo Genético devido ao facto do problema de sequenciamento ser de natureza NP-hard (um tipo de problema muito complexo). O GA gera soluções ótimas ou muito próximas do ótimo para o sequenciamento e o dimensionamento. A estrutura utilizada para resolver o problema em causa e demonstrado na Figura 14:

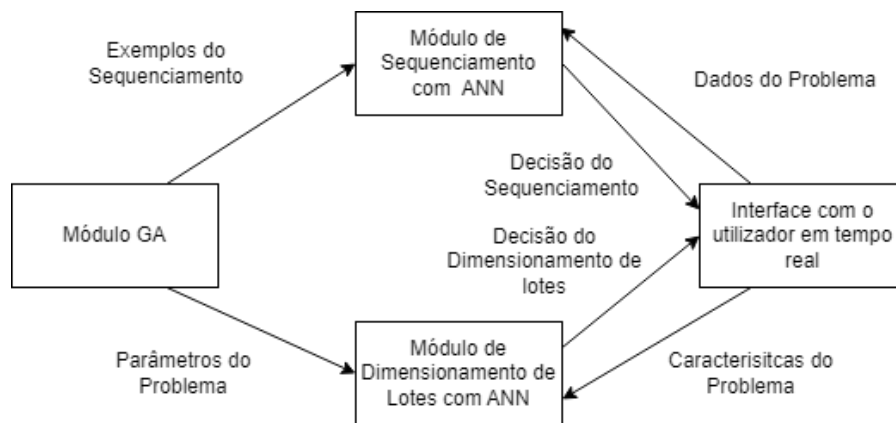


Figura 14 - Uma arquitetura de rede neural para dimensionamento e sequenciamento de lotes [modificado de:(Lee et al., 2001)]

O GA incorpora um procedimento de procura para identificar o valor ideal de dimensionamento de lote segundo certos parâmetros, simultaneamente procurando a sequência que minimiza o *makespan*. Como cada ambiente de fabrico difere, foi necessário obter um grande número de dados sob condições diferentes para treinar as ANNs. A primeira rede neural foi sujeita a quatro parâmetros de *input*: tempo de processamento médio, desvio do tempo de processamento, tempo de *setup* médio e desvio do tempo de *setup*. Sob estes parâmetros, a ANN ainda está dependente do número de máquinas, sendo necessários  $4 \cdot n$  (número de máquinas) *inputs*. De seguida o parâmetro ideal de dimensionamento de lote identificado no módulo de GA é usado como *output* alvo para cada exemplar. A Estrutura da Rede neural é demonstrada na Figura 15:

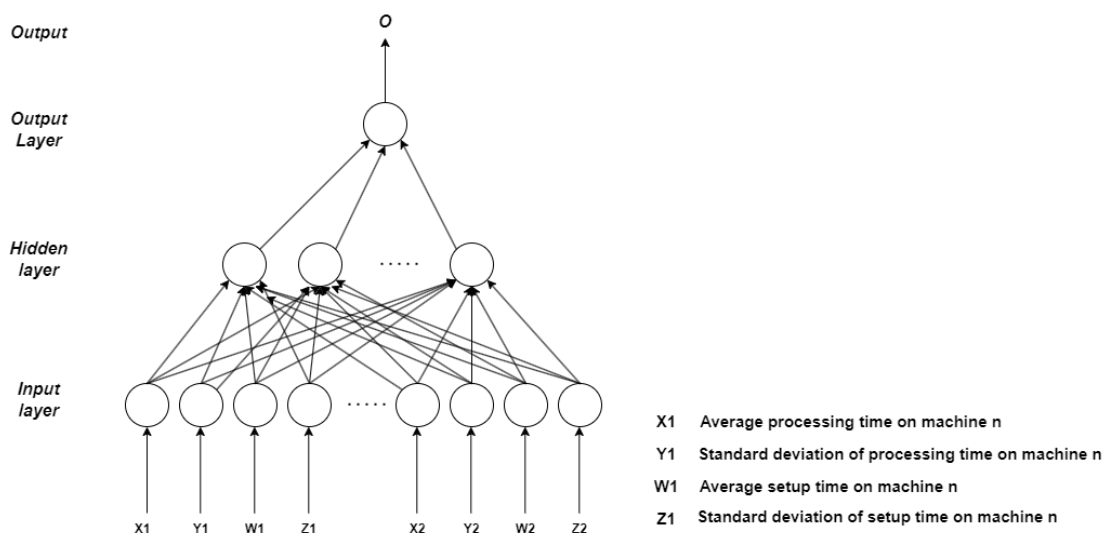


Figura 15 - Uma arquitetura de rede neural de retro propagação de três camadas para dimensionamento de lotes [modificado de:(Lee et al., 2001)]

Para o treinamento da ANN encarregue do sequenciamento, cada elemento dos *inputs* é composto por trabalhos a serem processados. Os *outputs* para os correspondentes *inputs* é a sequência de trabalhos desejada. Os *inputs*, exemplos de sequenciamento, para a ANN são gerados pelo módulo GA. De seguida os autores tiveram de implementar um método para a rede neural guardar o seu conhecimento e poder aplicar em novos casos. A técnica que resolveu

esta situação foi uma *Neural Weight Matrix*, que associa as melhores seqüências entre trabalhos. Com a matriz de pesos, a ANN de sequenciamento (Figura 16) utilizará a mesma para efetuar decisões de sequenciamento em tempo real.

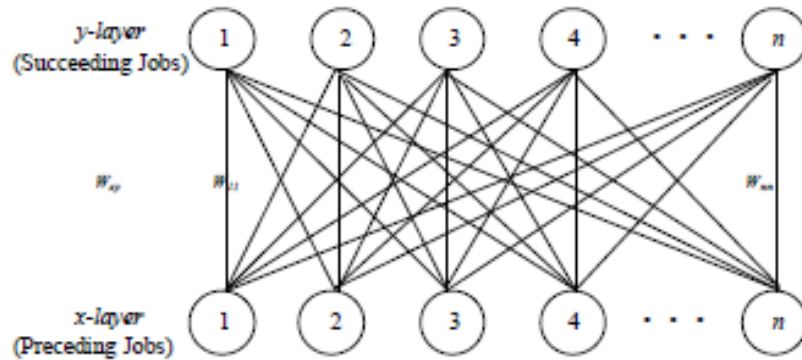


Figura 16 - Uma rede neural aplicada no sequenciamento em tempo real [fonte:(Lee et al., 2001)]

Com a aplicação desta técnica utilizando duas redes neurais, foi verificado pelos autores que a qualidade da solução foi consistentemente melhor do que os resultados duma ANN aplicada apenas ao sequenciamento. Também verificaram que com o aumento de máquinas, a qualidade da solução numa ANN com apenas sequenciamento piora com o aumento do número de máquinas, contrariamente á solução da *multi-Neural network*, que manteve a qualidade da solução.

Conclui-se assim que o estudo de Lee, Gupta e Amar (2001) propôs uma estrutura de *multi-Neural Networks* para resolver problemas de dimensionamento de lotes e sequenciamento. A estrutura inclui um Algoritmo Genético (GA) e ANNs. A primeira ANN foi treinada com vários parâmetros e a segunda rede neural foi responsável pelo sequenciamento. Uma *Neural Weight Matrix* foi usada para guardar o conhecimento da ANN e aplicá-lo em novos casos. A aplicação desta técnica resultou numa solução de melhor qualidade em comparação com uma ANN aplicada apenas ao sequenciamento, mantendo a qualidade mesmo com o aumento do número de máquinas.

**Caso 2 - Artificial neural network models for lot-sizing problem: A case study** (Şenyiğit & Atici, 2013)

Neste caso de estudo os autores desenvolveram uma nova ANN com esperança de que esta levasse a uma melhor performance em relação aos métodos mais utilizados no dimensionamento de lotes, como as heurísticas SM, EOQ, PPB e LUC. Como treinamento foram utilizados dados entre 2004 e 2009 que correspondem a artigos de *stock* de uma fábrica na Turquia. Os fatores em consideração neste estudo foram os custos de compra, de *stock* e *setup*.

Na Tabela 4 é mostrado como é fornecido á ANN a informação para cada período tendo em consideração se é efetuada uma encomenda, a procura e a quantidade que deveria de facto encomendar. Esta última é obtida pelo método de W-W e dado a ANN como output enquanto a demanda e custos foram assinalados como *inputs*.

Tabela 4 - A ilustração de como obter o padrão de pedido ideal [modificado de:(Şenyiğit and Atici 2013)]

Períodos	1	2	3	4	5	6	7	8	9	10	11	12
Procura	45	48	6	43	22	32	39	40	24	17	22	3
Quantidade da Encomenda	45	54	0	65	0	61	0	64	0	42	0	0
Padrão de Encomenda	1	1	0	1	0	1	0	1	0	1	0	0

Nesta pesquisa foram criadas e comparadas 400 ANNs utilizando diferentes estruturas e métodos de treinamento, de forma a obter a melhor possível combinação que origine melhores resultados, como se pode ver na Tabela 5.

Tabela 5 - Estruturas e treinamento para ANN

Estruturas	Métodos de treino
Multilayer perceptron	Transfer function
	Learning Rule
Self-Organizing Feature Map Network	Neighborhood Shape
	Learning Rule
	Transfer function
Radial Basis Function	Metric
	Competitive Rules
	Transfer function
	Learning Rule

Primeiramente os autores observaram se as ANN assinalavam corretamente os períodos a encomendar, ou seja, se identificavam corretamente o padrão de encomenda (identificado por 1 e 0 na Tabela 4), e de seguida compararam os custos obtidos em relação à solução ótima. Na Tabela 6 e Tabela 7 é demonstrado o desempenho das ANNs:

Tabela 6 - As 10 principais redes para os critérios de percentagem de desvio do custo ideal [modificado de:(Şenyiğit &amp; Atici, 2013)]

Ordem por classificação	Número da Rede	A percentagem de desvio do custo ideal	Estrutura da Rede	Learning rules	Transfer functions
1	220	0,130	RBF	DeltaBarDelta	LinearTanhAxon
3	180	0,147	RBF	DeltaBarDelta	LinearTanhAxon
4	190	0,150	RBF	DeltaBarDelta	SigmoidAxon
5	200	0,197	RBF	DeltaBarDelta	TanhAxon
6	202	0,202	RBF	Step	Axon
7	130	0,204	SOFM	DeltaBarDelta	SigmoidAxon
8	128	0,205	SOFM	ConjugateGradient	SigmoidAxon
9	177	0,211	RBF	Momentum	LinearTanhAxon
10	240	0,211	RBF	DeltaBarDelta	TanhAxon
11	15	0,212	MLP	QuickProp	SigmoidAxon

Tabela 7 - Comparação de ANNBM com métodos de dimensionamento de lotes [modificado de:(Şenyiğit & Atici, 2013)]

	EOQ		LTC		LUC		PPB		SM		ANNBM	
	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank
A percentagem de desvio do custo ideal	6,414	331	1,620	169	3,116	247	1,371	162	0,132	2	0,130	1
A percentagem de vezes que obteve o padrão de pedido ideal	79,744	152	63,782	2,36	55,000	288	71,474	191	91,218	4	91,410	3

Com base nos resultados a rede que originou os melhores resultados teve como base uma ANN com *Radial basis function*, uma *layer* de *input*, uma *layer* de *output* e uma *hidden layer*. Esta ANN foi o método que mais se aproximou da solução ótima, seguida pelo método de SM. Foi observado ainda que as ANN que tiveram um melhor desempenho foram as *Radial basis Function Networks* utilizando *DeltaBarDelta* como regra de aprendizagem.

De acordo como o autor, a principal razão para preferir usar uma ANN é a sua grande vantagem em conectar, dentro de uma única área, a flexibilidade da capacidade do poder de processamento de computadores e a adaptabilidade do cérebro humano.

### **Case 3 - A machine learning approach for identifying the best solution heuristic for a large scaled Capacitated Lotsizing Problem (Jens & Meyr, 2023)**

Neste estudo apresentado na conferência internacional de dimensionamento de lotes, é abordada a utilização de uma ANN na escolha da melhor heurística para o *Capacitated Lot Sizing Problem (CLSP)* e que obtenha um compromisso de qualidade de solução e tempo necessário para obter a mesma. Os tempos de cálculo das heurísticas aumentam significativamente com o tamanho do problema, logo testar todas as heurísticas disponíveis para instâncias de problemas muito grandes requer muito tempo, por isso, desenvolver um método que permita uma previsão da melhor heurística de solução para o respetivo problema sem ter de testar todas as heurísticas de solução disponíveis é o proposto por este artigo.

Para o método desenvolvido os autores escolherem utilizar uma ANN, que devido ao elevado tempo de computação requerido para problemas grandes, foi treinada em problemas mais pequenos, mas com diversas características. Desta forma o conjunto de dados de treino podem ser testados rapidamente e verificar quais as melhores heurísticas devem ser aplicadas a cada problema. O fator importante deste estudo é que se a ANN conseguir identificar padrões corretamente nos pequenos problemas, também poderá ser utilizada na previsão da melhor

heurística em problemas maiores. É de salientar que para esta fase de treinamento foram utilizadas técnicas de *supervised learning*.

A abordagem com melhores resultados para selecionar a melhor heurística dependendo das características do problema considerado é baseada numa ANN de duas *layers*. Recebe como *inputs* indicadores de performance que são previamente calculados, o que assegura a qualidade de solução para diferentes tamanhos de problemas, e como *output* a ANN fornece probabilidade de determinada heurística ser a melhor para o problema selecionado. Este método alcança uma exatidão de 78,17%, comparativamente a outros métodos de previsão que alcançaram no máximo 70,63%.

Este mesmo método baseado em ANN, quando sujeito a problemas 125% maiores, obteve uma exatidão de 78,61% na seleção da melhor heurística. É de salientar que o modelo não foi treinado em problemas desta dimensão. Além disso o tempo de processamento mesmo nestes problemas é bastante reduzido, menos de dois segundos, o que se torna mais eficiente do que testar todas as heurísticas.

### Outros casos de estudo

De forma a apresentar uma maior variedade de literatura com foco no problema de dimensionamento de lotes utilizando técnicas de inteligências artificial, foi desenvolvida a Tabela 8 que identifica o tipo de problema, a técnica utilizada, o resumo do estudo e a correspondente referência:

Tabela 8 - Tabela resumo de estudos na área

Tipo de problema	Técnicas de IA				Resumo	Referência
	ANN	DL	Algoritmos ML			
			GA	Outr.		
Dimensionamento de lotes para um único artigo	x				A ANN foi treinada usando o padrão de encomenda obtidos do algoritmo Wagner-Whitin. A ANN apresentou um desempenho melhor desempenho nos custos comparativamente as heurísticas mais comuns.	(Gaafar and Choueiki 2000)
Dimensionamento dinâmico de lotes	x	x	x		Este artigo aborda um problema de dimensionamento dinâmico de lotes (DLSP) com restrição de capacidade e estrutura de preços com desconto utilizando técnicas de GA, ANN e Deep Learning.	(Megala and Jawahar 2006)
Dimensionamento de lotes			x		O foco deste artigo está na literatura de GA para resolver problemas de dimensionamento de lotes. Primeiro, fornece uma visão geral dos avanços no campo até a data de publicação do mesmo. Em segundo lugar, apresenta ideias para pesquisas futuras, identificando lacunas na literatura.	(Guner Goren, Tunali, and Jans 2010)

Tipo de problema	Técnicas de IA				Resumo	Referência
	ANN	DL	Algoritmos ML			
			GA	Outr.		
Dimensionamento de lotes com restrição de capacidade e multi estágio	X				Esta pesquisa trata do LSP na cadeia de suprimentos resolvido através de reverse simulation metamodeling, uma metodologia nas ANN. A cadeia de abastecimento em estudo que opera em ambiente MTO. É caracterizada por planeamento de produção multiproduto, multi estágio e multi local com capacidade restrita e parâmetros estocásticos.	(Hachicha 2011)
Dimensionamento dinâmico de lotes	X			X	O artigo aplica um sistema de decisão, baseado em uma ANN e o Ant Colony Optimization modificado para resolver este problema estocástico de dimensionamento de lote dinâmico.	(Wong, Su, and Wang 2012)
Dimensionamento de lotes de nível único capacitado			X		Neste artigo é feita uma implementação de algoritmo genético para dimensionamento de lotes de nível único capacitado, utilizando modificações feitas no GA padrão são a forma como a população inicial é gerada.	(Zenon, Ahmad, and Ali 2012)
Dimensionamento de lotes para um único artigo	X		X	X	Neste artigo uma política de dimensionamento de lotes baseada no custo total mínimo e as incertezas da procura foram estudadas usando várias redes neurais artificiais treinadas com algoritmo genético (GA) e Bee Algorithm (BA). Essas abordagens combinadas foram examinadas juntamente com as heurísticas revistas de SM, LUC e a heurística CB.	(Şenyiğgit et al. 2013)
Dimensionamento de lotes capacitado multi-artigo com demanda estacionária		X		X	O artigo aborda a aplicabilidade do algoritmo Proximal Policy Optimisation (PPO) nos problemas de dimensionamento de lotes. PPO é um tipo de algoritmo que usa deep learning e Deep Reinforcement Learning (DRL).	(van Hezewijk et al. 2023)
Dimensionamento de lotes em dois estágios	X				Este estudo utiliza o fator “decision error” para treinar modelos de previsão, que depois foram utilizados como treinamento da rede neural, que depois acaba por mostrar resultados melhores no dimensionamento de lotes em relação a outros métodos abordados.	(Gong, Zhang, and Zhang 2023)

Tipo de problema	Técnicas de IA				Resumo	Referência
	ANN	DL	Algoritmos ML			
			GA	Outr.		
Dimensionamento de lotes de um único artigo capacitado	X			X	Neste artigo é proposto um método para a resolução do problema de lotes capacitado utilizando uma estrutura bidirecional Long Short term Memory (LSTM), que permite processar informações para frente e para trás no tempo para aprender soluções ótimas para o problema.	(Yilmaz and Büyüктаhtakin 2023)

## 2.10. Conclusão Bibliográfica

A revisão bibliográfica proporciona uma compreensão abrangente do contexto do dimensionamento de lotes de produção e sua interseção com a inteligência artificial (IA). Ao explorar os conceitos fundamentais do dimensionamento de lotes, ficou evidente que a otimização desse processo é crucial para minimizar custos de produção e manutenção. A revisão detalhada de métodos tradicionais, como *Lot for Lot* (LFL), Quantidade Económica (EOQ), e métodos exatos, como o Algoritmo *Wagner-Whitin* (WW), ressalta a complexidade desses problemas. A introdução de métodos heurísticos, como Ordens Periódicas (POQ) e *Silver-Meal* (SM), destaca a procura por soluções aproximadas em situações mais complexas. No entanto, a necessidade de explorar abordagens mais avançadas torna-se evidente, especialmente à luz da avaliação crítica da revisão bibliográfica.

A inclusão de métodos de IA, como *Genetic Algorithms*, *Machine Learning*, *Deep Learning*, e outros revelam a crescente promessa dessas técnicas na resolução de problemas complexos. A plataforma ChatGPT, com suas capacidades de processamento de linguagem natural, surgiu em 2022 como uma ferramenta inovadora e promissora para abordar vários tópicos e problemas, nomeadamente na programação. Contudo, com base na pesquisa feita, a sua aplicabilidade no dimensionamento de lotes ainda não foi comprovada.

Os casos de estudo e a literatura avaliada convergem para a conclusão de que a IA, nas suas diversas formas, oferece soluções inovadoras e adaptáveis para os problemas complexos no dimensionamento de lotes, desde modelos de ANN até à integração mais ampla de algoritmos avançados. Estes métodos, em geral, quando aplicados no problema de dimensionamento de lotes de produção mostram serem mais rápidos a obterem uma solução e normalmente com soluções de maior qualidade ou, com um melhor rácio entre qualidade de solução e tempo de processamento relativamente aos métodos tradicionais de lotes. Desta forma a IA está posicionada como uma ferramenta essencial para impulsionar eficiência, otimização e inovação no planeamento da produção.



## 3. Métodos e Aplicações

Neste capítulo, será apresentada a metodologia utilizada na aplicação do ChatGPT e do CoPilot, baseados nas tecnologias GPT-3.5 e GPT-4, respectivamente, ao problema proposto. O objetivo é comparar o desempenho dessas duas plataformas de inteligência artificial com as técnicas discutidas nos capítulos anteriores.

O dimensionamento de lotes é uma área crucial na otimização de processos de produção e logística, com implicações significativas para eficiência, custo e desempenho operacional. O uso de tecnologias de inteligência artificial, como o ChatGPT pode oferecer uma nova abordagem ao problema, aproveitando a capacidade do modelo de linguagem natural para compreender e gerar soluções eficazes.

No final deste capítulo, espera-se ter proporcionado uma compreensão clara da metodologia empregada na utilização do ChatGPT para resolver problemas de dimensionamento de lotes.

### 3.1. Modelo Matemático

Este capítulo é dedicado ao estudo do problema de dimensionamento de lotes de produção e às suas variações.

Este tipo de problema é um problema de otimização, envolvendo a decisão da quantidade de artigos a serem produzidos num determinado período de tempo de forma a minimizar os custos totais, tendo em consideração várias restrições, como a capacidade de produção, procura do mercado, custos de posse de *stock*, custos de *setup* e outros. O problema de dimensionamento de lotes, tal como mencionado no capítulo anterior, é subdividido em diversos tipos de problemas, dependendo em parte das restrições e fatores a que está sujeito.

#### 3.1.1. SISLP

O SISLP é comumente utilizado como um sub-problema em diversos algoritmos para problemas de dimensionamento de lotes muito complexos, que são por norma modelados para programação linear mista, mais conhecida por Mixed Integer Linear Programming (MILP) e Dynamic Programming (DP), como foi no caso do algoritmo de Wagner Within.

No *Single Item Lot Sizing Problem (SILSP)* é considerado um único artigo ou produto, deve atender a uma procura pré-estabelecida e pode ser formulado com ou sem restrições de capacidade. O problema de dimensionamento para um único item é a forma mais simples de

um problema de dimensionamento de lotes e engloba três aspetos principais: o nível de produção, *stock* disponível e decisões de *setup*

Para a formulação matemática proposta para o problema geral de SILSP consideraram-se as seguintes variáveis e parâmetros:

### Índices

$t = \{1, \dots, T\}$  Períodos de tempo;

### Parâmetros

$d_t$  Procura (número de produtos) no período  $t$   
 $p_t$  Custo de produção por unidade para o período  $t$   
 $h_t$  Custo de Posse do *stock* para o período  $t$   
 $s_t$  Custo de encomenda / Custo de *Setup* no período  $t$   
 $I_t$  *Stock* (número de produtos) no fim do período  $t$

### Variáveis de decisão

$X_t$  Quantidade de produção para o período  $t$  (variáveis inteiras)

### Variáveis auxiliares

$Y_t$  *Setup* no período  $t$  (variáveis inteiras)

$$\min Z = \sum_{t=1}^T (s_t Y_t + p_t X_t + h_t I_t) \quad \text{EQ. 6}$$

*Sujeito a:*

$$I_{t-1} + X_t = d_t + I_t \quad \forall t \in T \quad \text{EQ. 7}$$

$$X_t \leq M Y_t \quad \forall t \in T \quad \text{EQ. 8}$$

$$X_t, I_t \geq 0, I_0 = 0 = I_T \quad \text{EQ. 9}$$

$$Y_t \quad \forall t \in [0,1] \quad \text{EQ. 10}$$

É assumido que o *stock* no início e no fim do período  $T$  é zero.

A função de objetivo (6), minimiza a soma dos custos de *setup*, produção e de posse para a totalidade do horizonte de planeamento  $T$ .

A Restrição (7) é necessária para manter o balanceamento do *stock*, de forma que a produção excedente do último período ( $I_{t-1}$ ) é adicionada à produção do período atual ( $X_t$ ) e são utilizadas para satisfazer a procura ( $d_t$ ). O que sobra é guardado em *stock* para o período seguinte  $I_t$ .

A restrição (8) em conjunto com a (10) asseguram que o tempo e o custo de preparação são considerados apenas quando existe produção. Ainda referente à restrição (8), o “M” simboliza um número inteiro muito grande, como por exemplo 9999.

A restrição (9) define que a quantidade a produzir é sempre igual ou superior a zero, da mesma forma que a quantidade que é armazenada em *stock* é sempre igual ou superior a zero, não permitindo assim rupturas de *stock*.

A restrição (10) define que a variável  $Y_t$  é uma variável binária, tomando valor 1 quando é efetuada uma encomenda/*Setup* e valor 0 caso contrário.

Caso se pretenda considerar a capacidade como uma restrição, a única diferença para o modelo anterior é a restrição (8), que seria  $X_t \leq C_t \times Y_t \forall t \in T$ , sendo que  $C_t$  é a capacidade máxima de produção para um determinado período  $t$ .

A maior parte da literatura sobre o SILSP básico e suas extensões têm-se concentrado em métodos de solução exatos. Esses métodos incluem programação dinâmica, abordagens poliédricas e algoritmos de *branch-and-cut* e *branch-and-bound*. No entanto, métodos heurísticos, como heurísticas de construção simples, algoritmos de aproximação e algumas heurísticas de melhoria, também foram desenvolvidas. Estes desempenharam um papel importante no desenvolvimento de heurísticas para problemas de dimensionamento de lotes mais complexos. Alguns estudos concentram-se em classificar SILSPs com base em sua complexidade computacional (Brahimi et al., 2017).

### 3.1.2. SILSP com lotes constantes

Uma variação do SILSP é aquela com lotes de produção constantes, uma vez que simplifica significativamente o problema, tanto na sua formulação matemática, quanto na sua implementação, e contribui para a estabilidade operacional devido à produção de lotes constantes. Para a formulação matemática proposta para esta variante do SILSP consideraram-se as seguintes variáveis e parâmetros:

#### Índices

$t = \{1, \dots, T\}$  Períodos de tempo;

#### Parâmetros

$T$  Tamanho do horizonte de planeamento

$d_t$  Procura (número de produtos) no período  $t$

$p_t$	Custo de produção por unidade para o período $t$
$h_t$	Custo de Posse do <i>stock</i> para o período $t$
$s_t$	Custo de <i>Setup</i> no período $t$
$C_t$	Capacidade de produção no período $t$
$f_t$	Custo fixo por lote no período $t$
$v$	Tamanho do lote constante
$I_t$	<i>Stock</i> (número de produtos) no fim do período $t$
$Z_t$	Número de Lotes a produzir no período $t$

#### Variáveis de decisão

$X_t$	Quantidade de produção $t$ (variáveis inteiras)
-------	---

#### Variável auxiliar

$Y_t$	Encomenda/ <i>Setup</i> no período $t$ (variáveis inteiras)
-------	---

$$\min Z = \sum_{t=1}^T (s_t Y_t + p_t X_t + h_t I_t + f_t Z_t) \quad EQ. 11$$

Sujeito a (8),  
(10) e:

$$X_t \leq C_t \times Y_t \quad \forall t \in T \quad EQ. 12$$

$$X_t \leq v \times Z_t \quad \forall t \in T \quad EQ. 13$$

$$Y_t \leq Z_t \quad \forall t \in T \quad EQ. 14$$

$$Z_t \leq \frac{C_t}{v} Y_t \quad \forall t \in T \quad EQ. 15$$

$$I_0 = I_T = 0 \quad \forall t \in T \quad EQ. 16$$

$$Z_t \geq 0 \text{ e inteiro } \forall t \in T \quad EQ. 17$$

A função objetivo (11) minimiza os custos de produção, armazenamento de *stock*, e número de lotes de produção, contabilizando também os custos fixos de produção de lote.

As restrições (7) e (12) são respetivamente as restrições de balanceamento de inventário e a restrição de capacidade de produção.

A restrição (13) relaciona as variáveis de produção com o de tamanho e número dos lotes de produção, o que levará à produção total de artigos.

As restrições (14) e (15) relacionam as variáveis de Setup e do tamanho do lote. A primeira obriga a, caso haja um *setup*, obrigatoriamente existirem lotes a ser produzidos. A segunda restrição limita o número de lotes de produção.

No caso da Restrição (16) define que o inventário no início e fim do horizonte de planeamento é 0.

A restrição (17) define que o número de lotes a produzir é sempre inteiro e maior ou igual a 0. Da mesma forma a restrição (9) expressa o mesmo do modelo anterior.

Muitos autores realizaram a sua própria formulação para a resolução do SILSP com lotes de produção constantes, como (Van Vyve, 2007) que denominou o problema de SICLSP-SW, que possuía a característica de não ter custos de *setup*, desenvolvendo algoritmos de tempo polinomial. Os Problemas de dimensionamento de lotes com tamanhos de lote constantes foram relacionados também ao problema integrado de dimensionamento de lotes e roteamento de veículos, estudado por (Alp et al., 2003).

### 3.1.3. Dimensionamento e Escalonamento De Lotes

O problema geral de dimensionamento e escalonamento de lotes de produção aborda a integração do dimensionamento e programação de vários produtos em uma única máquina. Os tamanhos contínuos de lotes são determinados e programados com o objetivo de minimizar os custos de *stock* e os custos de *setup* dependentes da sequência de produção. Como o cronograma é independente de períodos de tempo predefinidos, o problema generaliza modelos conhecidos usando estruturas de tempo restritas.

O modelo proposto considera vários artigos  $j = 1, \dots, J$  para sequenciar num período de tempo definido pelo plano de horizonte constituído por macro períodos  $t = 1, \dots, T$ . Cada período  $t$  é dividido num número fixo não sobreposto de micro períodos com duração variável, onde  $S_t$  define os micro períodos  $s$  pertencentes ao macro período  $t$ . Todos os micro períodos são ordenados na sequência  $s = 1, \dots, S$ .

No modelo geral, o estado de *setup* da máquina muda no início ou fim dos micro períodos. Desta maneira a mudança de *setup* da máquina é uma das variáveis de decisão, em conjunto com a duração do micro período, que é expressa pela quantidade produzida nesse micro período. O número de micro períodos  $S_t$  dentro dos macro períodos  $t$  é fixado previamente de forma a ser possível formular em MIP (Mixed Integer programming), pois define um limite para o número de possíveis mudanças de *setup* dentro dos macro períodos  $t$ . O lote consiste numa sequência de micro períodos atribuídos ao mesmo artigo e podem continuar em diferentes macro períodos. Como consequência do número fixo de micro períodos, dentro do lote podem existir micro períodos de pausa, com produção igual a zero. Para a formulação matemática proposta para este tipo de problema de LSP consideraram-se as seguintes variáveis e parâmetros:

## Índices

$t = \{1, \dots, T\}$  Macro Períodos  $t$ ;

$j = \{1, \dots, J\}$  Artigos  $j$ ;

$s = \{1, \dots, S\}$  Micro Períodos  $s$

## Parâmetros

$T$  Tamanho do horizonte de planeamento

$S_t$  Conjunto de micro períodos  $s$  pertencentes ao macro período  $t$

$K_t$  Capacidade (tempo) produção para o macro período  $t$

$a_j$  Tempo de produção para artigo  $j$

$m_j$  Lote mínimo de artigo  $j$

$h_j$  Custos de *stock* para artigo  $j$  (por unidade e macro período)

$s_{ij}$  Custos de troca de *setup* de artigo  $i$  para  $j$

$d_{jt}$  Procura do artigo  $j$  no macro período  $t$

$I_{j0}$  *Stock* inicial do artigo  $j$

$Y_{j0}$  Igual a 1 se a máquina estiver com *setup* para artigo  $j$  no plano de horizonte

## Variáveis de decisão

$I_{jt}$  *Stock* do artigo  $j$  no fim do macro período  $t$  (variáveis inteiras)

$X_{js}$  Produção do artigo  $j$  no micro período  $s$  (variáveis inteiras)

$Y_{js}$  Estado de *Setup*: 1 se a máquina estiver preparada para artigo  $j$  no micro período  $s$  (variáveis inteiras)

$z_{ijs}$  Igual a 1 se ocorrer uma mudança de *setup* na máquina do artigo  $i$  para  $j$ . (variáveis inteiras)

$$\min Z = \sum_{j,t} h_j I_{jt} + \sum_{i,j,s} s_{ij} z_{ijs} \quad \text{EQ. 18}$$

Sujeito a:

$$I_{jt} = I_{j,t-1} + \sum_{s \in S_t} h_j I_{jt} - d_{jt} \quad \forall t, j \quad \text{EQ. 19}$$

$$\sum_{j,s \in S_t} a_j x_{js} \leq K_t \quad \forall t \quad \text{EQ. 20}$$

$$x_{js} \leq \frac{K_t}{a_j} y_{js} \quad \forall s, t \quad \text{EQ. 21}$$

$$x_{js} \geq m_j (y_{js} - y_{j,s-1}) \quad \forall s, j \quad \text{EQ. 22}$$

$$\sum_j y_{js} = 1 \quad \forall s \quad \text{EQ. 23}$$

$$z_{ijs} \geq y_{i,s-1} + y_{j,s-1} \quad \forall s, i, j \quad \text{EQ. 24}$$

$$I_{jt} \geq 0 \quad \text{EQ. 25}$$

$$x_{js} \geq 0 \quad \text{EQ. 26}$$

$$y_{js} \in [0,1] \quad \text{EQ. 27}$$

$$z_{ijs} \in [0,1] \quad \text{EQ. 28}$$

A função objetivo (18) expressa os custos de *stock* e a mudança de *setup*, procurando minimizar os mesmos.

A restrição (19) e  $I_{jt} \geq 0$  para todos  $j$  e  $t$  garantem que a procura seja atendida e sem atrasos.

A restrição de capacidade é dada por (20).

As restrições (21) e (23) expressam que a produção só pode ocorrer se a máquina estiver com *setup* para o respetivo artigo e um e apenas um estado de *setup* é definido em cada micro período. Como a duração do micro período  $s$  é determinado pelo consumo de capacidade da produção quantidade  $x_{js}$  (onde  $j$  é o artigo cujo indicador de estado  $y_{js}$  é definido como 1), são permitidos micro períodos com comprimento zero.

A restrição (22) permite tamanhos de lote mínimos, que são introduzidos para evitar mudanças do estado de *setup* sem mudanças do produto.

A restrição (24) interliga o estado de *setup* e indicadores de mudança de artigo. A Variável  $z_{ijs}$  terá como valor 1 ou 0 em qualquer solução ótima.

## 3.2. Utilização do ChatGPT e Copilot

### 3.2.1. Prompts

“*Prompting*” refere-se ao processo de fornecer informações ou instruções a um modelo de IA para gerar o resultado pretendido. Este *input* pode assumir diversas formas, dependendo do tipo de IA que se trata. O tipo e *input* mais conhecido é o texto, mas também podem ser imagens e servem como um guia para a IA produzir uma resposta ou ação específica.

Nos Modelos de IA baseados em imagens o *input* envolve o fornecimento de uma imagem inicial ou uma descrição da imagem desejada, ou seja, o *output*. Por exemplo, se pretender que a IA gere uma imagem de um pôr do sol sobre o oceano, pode fornecer um *prompt* como: "Uma imagem de um oceano calmo com um pôr do sol vibrante ao fundo". A IA então tentará gerar uma imagem que corresponda à descrição fornecida no *prompt*. Dependendo de qual IA utilizar para a geração de imagens, a ordem dos *prompts* e peso de cada um pode ser alterada, levando a diferentes imagens com a mesma instrução.

Na figura Figura 17 é demonstrado o resultado obtido para o *prompt* “Linda paisagem de anime, uma casinha à beira do lago numa floresta, céu roxo e azul com nuvens ao pôr do sol, flores na relva perto da água, árvores ao redor da casa, estilo de arte anime, fundo detalhado, colorido, cores vibrantes, pintura pastel, iluminação suave, cenário lindo, em alta resolução.”

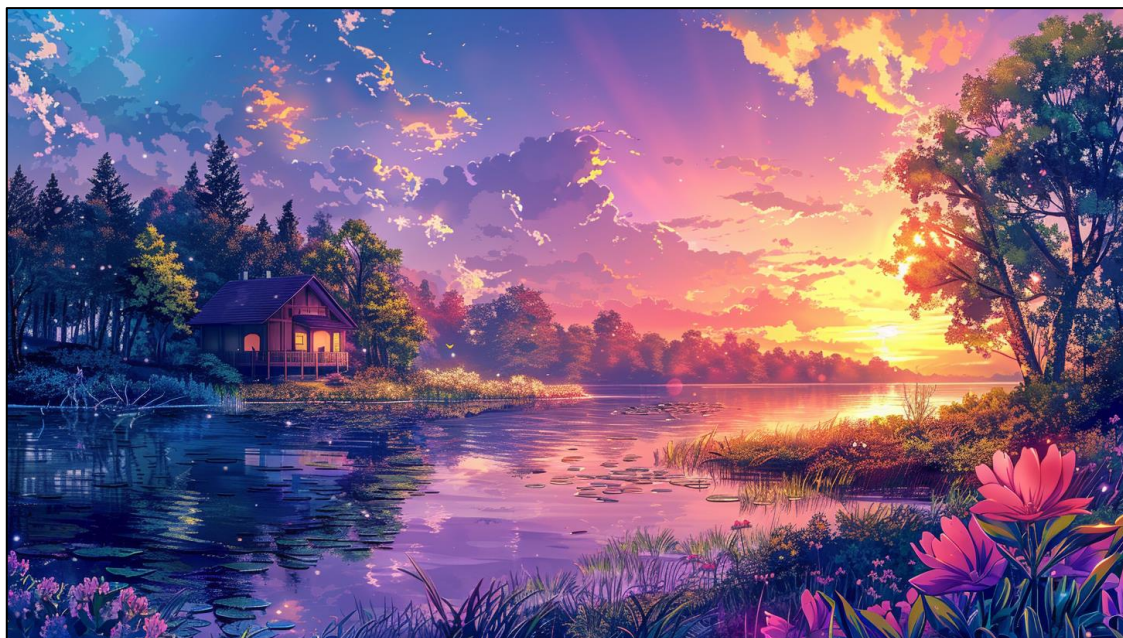


Figura 17 - Exemplo geração Imagem com MidJourney (MidJourney, 2024)

Nestes modelos de IA também podem existir “*Negative prompts*”, que funcionam como uma ferramenta para excluir características da imagem, como por exemplo se desejássemos uma floresta sem animais, como *negative prompt* colocaríamos “Animais”.

Apesar dos bons resultados para o exemplo anterior, quando a IA tem de representar características humanas como dedos, esta tecnologia ainda apresenta dificuldades. Para combater problemas como este, muitos programadores desenvolvem os seus próprios modelos para a IA para obter melhores resultados em características e objetos específicos, utilizando conjuntos de imagens onde essa característica está devidamente representada como exemplo para a IA seguir.

No caso de modelos de IA baseados em texto o “*prompt*” geralmente envolve o fornecimento de uma frase ou instrução inicial que o modelo usa como contexto para gerar uma resposta ou até mesmo gerar novo conteúdo que não foi especificado na instrução, como nos exemplos seguintes: onde na Figura 18 e Figura 19 é demonstrada a resposta do *ChatGPT* a uma pergunta clara, e Figura 20 quando não é provido de instruções muito esclarecedoras. Nesta última o *ChatGPT* revela a sua capacidade de perceção e contextualização ao perceber que o “Era uma vez” é geralmente utilizado no início de muitos contos, levando-o a criar uma pequena história de um conto de fantasia como resposta ao “*input*” inicial (Figura 20).

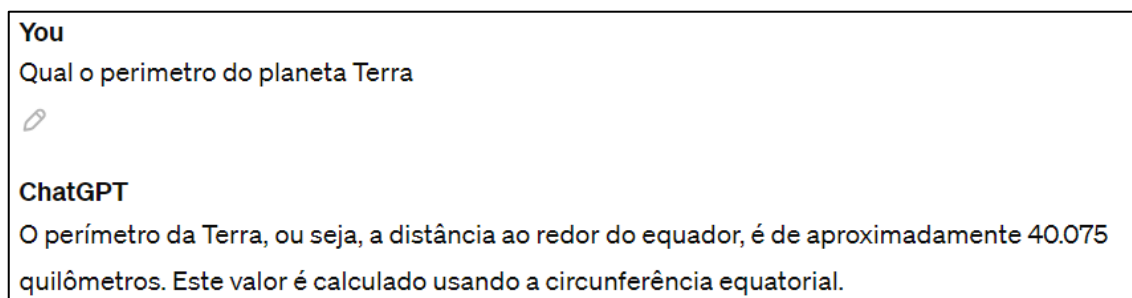


Figura 18 - Exemplos resposta clara (1)

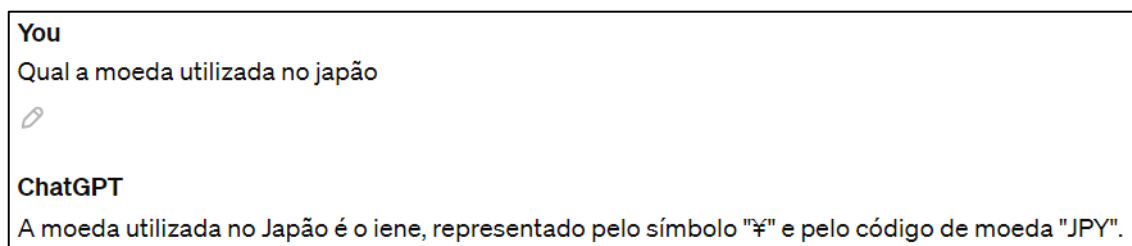


Figura 19 - Exemplo resposta clara (2)

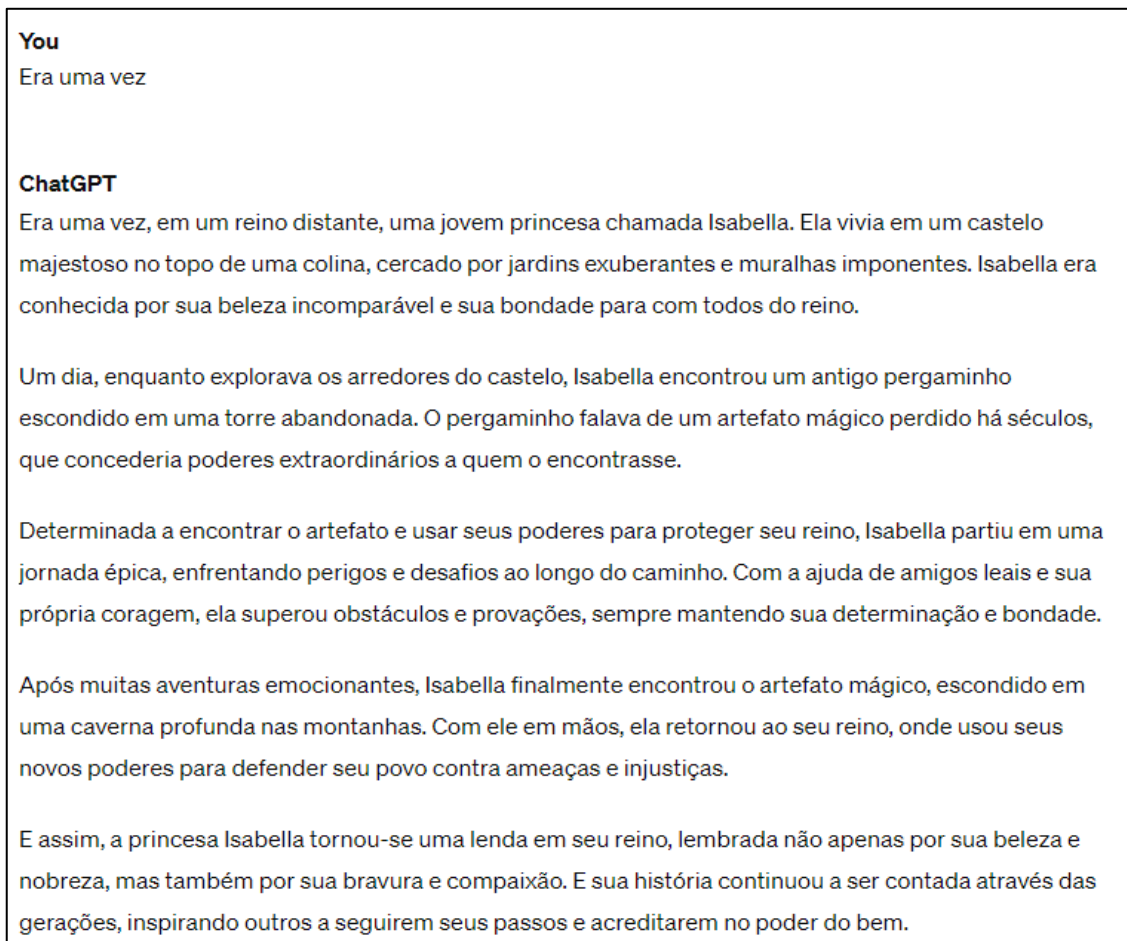


Figura 20 - Resposta (2) ChatGPT

Ao fornecer *prompts* claros e bem formulados, os utilizadores podem influenciar significativamente o resultado gerado pela IA, garantindo que as respostas atendam às suas necessidades e expectativas. Além disso, *prompts* bem elaborados ajudam os modelos a compreender melhor o contexto das perguntas, resultando em respostas mais relevantes e precisas. Assim, a capacidade de criar *prompts* eficazes é uma habilidade valiosa ao utilizar estas plataformas, pois impacta diretamente a qualidade e a utilidade das respostas geradas. A própria OpenAI fornece no seu site algumas dicas para melhorar a qualidade dos *prompts* (Open AI, 2024), mas o trabalho de Sahoo e outros autores (Sahoo et al., 2024) fornece um ponto de vista e compreensão sobre muitos dos fatores associados ao *prompting* (Figura 21).

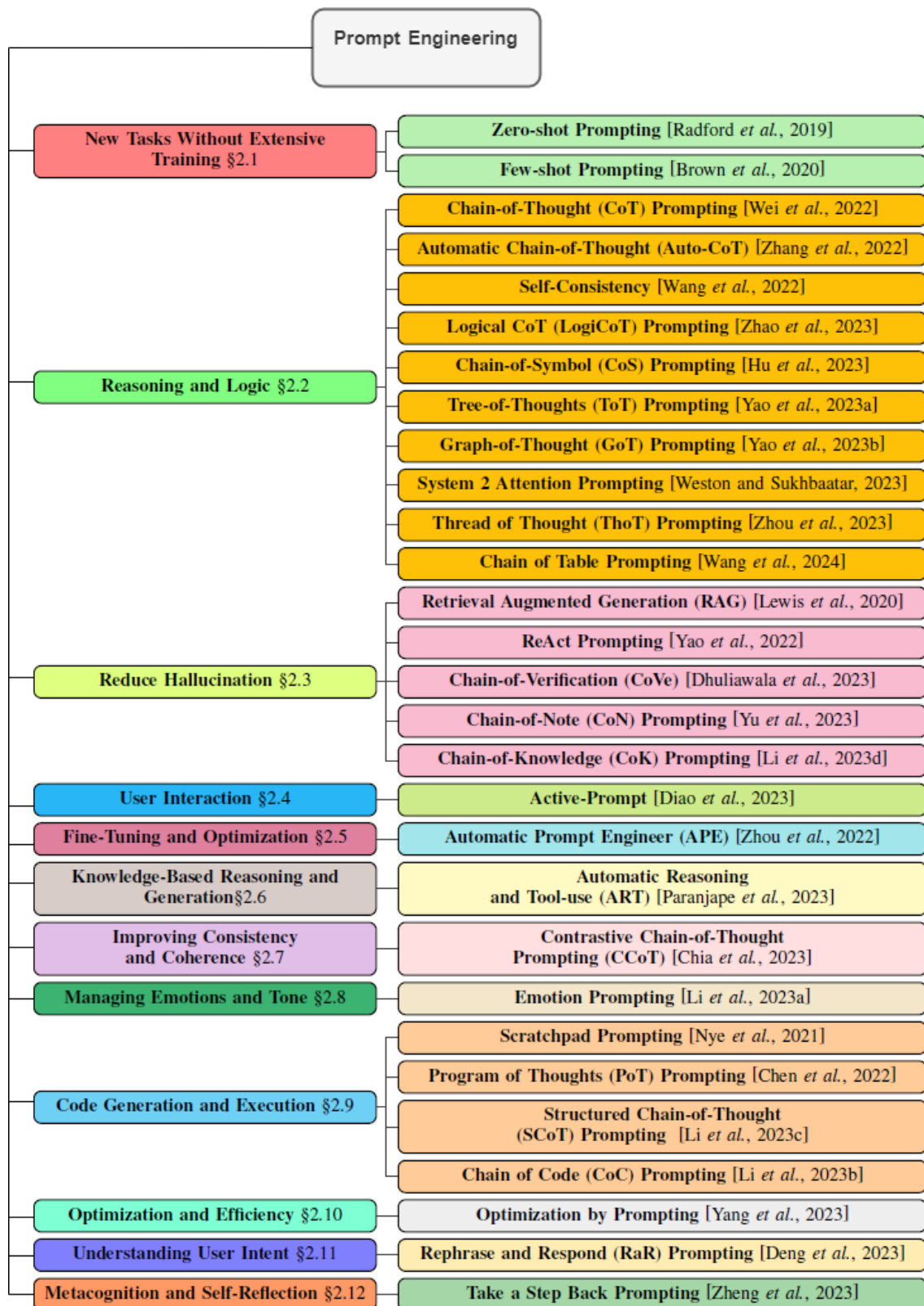


Figura 21 - “Taxonomia de técnicas de engenharia de prompts em LLMs, organizadas em torno de domínios de aplicação, fornecendo uma estrutura diferenciada para personalizar prompts em diversos contextos [fonte:(Sahoo et al., 2024)]

No domínio da inteligência artificial, a “Prompt Engineering” tem-se tornado uma força transformadora, liberando o vasto potencial de LLMs. O Trabalho de Sahoo (Sahoo et al., 2024) categoriza sistematicamente técnicas de *prompting*. A análise abrange aplicações, modelos e conjuntos de dados, esclarecendo os pontos fortes e limitações de cada abordagem. Os autores notaram que apesar de sucessos notáveis, os desafios persistem, incluindo preconceitos, imprecisões factuais e lacunas de interoperabilidade, havendo a necessidade de investigação adicional e estratégias de mitigação.

Como abordado anteriormente, *prompting* é um aspeto crucial do trabalho com modelos de IA porque permite que os humanos orientem o comportamento do modelo e o orientem para a produção de resultados que sejam relevantes ou úteis para uma tarefa ou aplicação específica. A qualidade das respostas do ChatGPT ou Copilot neste presente trabalho dependeram essencialmente da qualidade e especificidade do *prompt* fornecido, bem como das capacidades e limitações do modelo de IA utilizado.

### **3.2.2. Desafios e problemas**

Os autores (Wang et al., 2023) estudaram a eficácia do ChatGPT na geração de consultas booleanas de alta qualidade para procura literária. Eles verificaram que para dois *prompts* exatamente iguais, o ChatGPT gera resultados diferentes, indicando suas limitações em consistência e estabilidade. De certa forma esta característica não é apenas um erro mais sim uma característica implementada em *Large Language Models* (LLMs) como ChatGPT-4. Estes modelos funcionam usando uma enorme ANN para processar grandes quantidades de dados de texto. O modelo é treinado através dum elevado número de dados, através da internet ou uma coleção de livros, e aprende a reconhecer padrões nos dados. Esta aprendizagem permite que o modelo gere respostas coerentes e contextualmente relevantes para novos *prompts* de texto. Quando um utilizador insere um *prompt*, o modelo de linguagem usa sua ANN treinada para gerar uma resposta. O modelo não simplesmente fornece respostas pré-escritas, mas gera uma resposta instantaneamente, baseando-se nos padrões aprendidos e no contexto do *prompt*. É por isso que modelos como a série GPT podem gerar respostas semelhantes às humanas que parecem adaptadas ao *prompt* específico. Apesar de serem bastante precisos, por vezes as respostas para um mesmo *prompt* podem ter diferentes respostas. Isto ocorre porque o modelo é estocástico, ou seja, gera respostas baseadas em uma distribuição de probabilidade. Quando um utilizador insere um *prompt*, o modelo gera múltiplas respostas possíveis e classifica-as por probabilidade. Isso significa que mesmo pequenas alterações no *prompt* podem levar a respostas diferentes. Além disso, os LLMs também podem ser influenciados pelos próprios dados da fase de treinamento. Por exemplo, se o modelo for treinado em texto que contenha linguagem tendenciosa ou estereótipos, poderá gerar respostas que reflitam esses preconceitos.

Outro desafio substancialmente prevalente reside na ocorrência de alucinações. Os LLMs são concebidos para oferecer respostas estocásticas baseadas nos dados em que foram treinados, podendo, contudo, exibir uma propensão a uma prontidão excessiva ao fornecer respostas. Em vez de reconhecerem a incerteza e optarem por uma resposta de "não sei", frequentemente

forneem respostas incorretas, apresentando-as como verdades factuais. No estudo conduzido por Alkaissi e McFarlane, (Alkaissi & McFarlane, 2023), os autores constataram, que os dados gerados pelo ChatGPT constituem um conjunto de informações autênticas e inteiramente fabricadas. O foco da pesquisa dos autores concentrou-se no domínio da medicina, suscitando inquietações em relação à integridade e precisão dos modelos linguísticos de grande escala. Um exemplo ilustrativo pode ser observado na Figura 22 na qual é apresentada uma questão de cálculo ao ChatGPT, resultando em uma resposta incorreta. O resultado correto seria uma soma de 1135 e média de 92,92.

**You**  
Which is the average demand for the following ordering pattern: for period 1 the demand is 10; for period 2 the demand is 10; for period 3 the demand is 15; for period 4 the demand is 20; for period 5 the demand is 70; for period 6 the demand is 180; for period 7 the demand is 250, for period 8 the demand is 280; for period 9 the demand is 230; for period 10 the demand is 40; for period 11 the demand is 0; and for period 12 the demand is 10.

**ChatGPT**  
To calculate the average demand for the given ordering pattern, you sum up all the demands for each period and then divide by the total number of periods.

The sum of demands for all periods:  
 $10 + 10 + 15 + 20 + 70 + 180 + 250 + 280 + 230 + 40 + 0 + 10 = 1035$

Now, divide this total demand by the number of periods:  
Average demand =  $\frac{1035}{12} = 86.25$

So, the average demand for this ordering pattern is 86.25 units.

Figura 22 - Exemplo resposta Errada

Apesar da plataforma por vezes cometer erros que nos aparentam ser “estúpidos” como foi o caso da Figura 22, numa segunda resposta para o mesmo *prompt*, o cálculo já foi efetuado corretamente (Figura 23), revelando o já abordado, fator de aleatoriedade da plataforma.

**You**  
Which is the average demand for the following ordering pattern: for period 1 the demand is 10; for period 2 the demand is 10; for period 3 the demand is 15; for period 4 the demand is 20; for period 5 the demand is 70; for period 6 the demand is 180; for period 7 the demand is 250, for period 8 the demand is 280; for period 9 the demand is 230; for period 10 the demand is 40; for period 11 the demand is 0; and for period 12 the demand is 10.

**ChatGPT**  
To calculate the average demand for the given ordering pattern, you can sum up all the demands and then divide by the total number of periods.

Average demand =  $(10 + 10 + 15 + 20 + 70 + 180 + 250 + 280 + 230 + 40 + 0 + 10) / 12$

Average demand =  $1115 / 12 \approx 92.92$

So, the average demand for this ordering pattern is approximately 92.92 units per period.

Figura 23 - Exemplo resposta certa

### 3.2.3. Conhecimento no Problema discutido

Nesta Secção será demonstrado e verificado o conhecimento de duas plataformas de IA, ChatGPT e *Copilot*, a primeira alimentada pelo modelo GPT 3.5 e a segunda pelo GPT 4. Para poder comparar as respostas de ambas as plataformas, será utilizado o mesmo *prompt*. Cada resposta será também avaliada quanto á veracidade da resposta fornecida.

De modo ao modelo não efetuar nenhuma recolha de informação prévia e poder assim responder á questões colocadas sem conhecimento prévio, cada questão colocada foi feita numa nova conversa, com exceção de indicação do contrário. Deste modo serão avaliados os resultados em relação ao conhecimento no problema em geral e conhecimento em técnicas específicas. Na tabela Tabela 9 é identificado o *prompt* e resultados em ambas as plataformas:

Tabela 9 - Resposta ao *prompt* P1(ChatGPT e Copilot)

ChatGPT	CoPilot
P1: "You are an expert in industrial management. Explain to me in a concise matter what is the lot-sizing problem?"	
The lot-sizing problem is a key challenge in industrial management concerning the determination of the optimal quantity to produce or order at a given time, considering factors like production costs, inventory holding costs, and demand fluctuations. It aims to strike a balance between minimizing production costs and minimizing inventory holding costs, ensuring efficient production and distribution processes.	The lot-sizing problem is a common issue in production and inventory management. It involves determining the optimal quantity of inventory to order or produce while minimizing total costs. These costs can include production costs, holding costs, and setup costs. Here's a brief explanation: Production Costs: These are the costs associated with producing each unit of inventory. They usually decrease as the quantity increases due to economies of scale. Holding Costs: These are the costs associated with storing inventory over time. They usually increase with the quantity of inventory held. Setup Costs: These are the fixed costs associated with starting a production run. They are independent of the quantity produced. The challenge is to find the balance between these costs. If you order or produce too much, you'll have high holding costs. If you order or produce too little, you'll have high setup costs due to frequent production runs. The lot-sizing problem aims to find the "sweet spot" that minimizes the total cost. This is a fundamental problem in operations research and supply chain management. Various mathematical models and algorithms have been developed to solve it.

No caso deste primeiro *prompt*, concebido para avaliar a capacidade de ambas as plataformas em explicar adequadamente o problema em questão, ambas se demonstraram corretas nas suas respostas, embora se destaque a elaboração mais pormenorizada por parte do Copilot. Enquanto o ChatGPT ofereceu uma explicação do problema e suas variáveis, o Copilot foi além, fornecendo uma descrição minuciosa de cada uma dessas variáveis. Em linhas gerais, ambas as abordagens se revelaram verossímeis na sua contextualização e análise do problema apresentado, ainda que o nível de detalhe fornecido pelo Copilot tenha sido mais abrangente e informativo. O segundo *prompt* e respetiva resposta para ambas as plataformas é demonstrado na Tabela 10 e Tabela 11.

Tabela 10 - Resposta ao *prompt* P2(ChatGPT)

P2: "How can I determine the lot size and production plan?"
<b>ChatGPT</b>
<p>Determining the lot size and production plan involves several steps:</p> <p><u>Demand Forecasting</u>: Estimate future demand for your product based on historical data, market trends, and other relevant factors.</p> <p><u>Cost Analysis</u>: Calculate the costs associated with production, setup, inventory holding, and any other relevant expenses.</p> <p><u>Inventory Policies</u>: Choose an inventory policy that aligns with your business goals, such as the Economic Order Quantity (EOQ), which minimizes total inventory costs.</p> <p><u>Lot Sizing Techniques</u>: Select a lot sizing technique that suits your production environment and constraints. Common techniques include the EOQ model, Wagner-Whitin algorithm, and dynamic programming.</p> <p><u>Constraints Consideration</u>: Consider any constraints that may affect your production plan, such as capacity limitations, lead times, and minimum order quantities.</p> <p><u>Optimization</u>: Use mathematical optimization methods or software tools to find the optimal lot size and production plan that minimizes total costs while satisfying demand and constraints.</p> <p><u>Sensitivity Analysis</u>: Conduct sensitivity analysis to assess the impact of changes in key parameters, such as demand variability or production costs, on the optimal solution.</p> <p><u>Continuous Improvement</u>: Regularly review and adjust your lot sizing and production plan based on actual performance, market changes, and evolving business needs to achieve continuous improvement.</p>

Tabela 11 - Resposta ao *prompt* P2(Copilot)

P2: "How can I determine the lot size and production plan?"
<b>Copilot</b>
<p>Determining the lot size and production plan involves solving the lot-sizing problem, which is a type of optimization problem. Here are the general steps:</p> <p><u>Identify Costs</u>: Determine the specific costs associated with your production process. This includes setup costs, holding costs, and production costs.</p>

P2: "How can I determine the lot size and production plan?"
<b>Copilot</b>
<p><u>Create a Demand Forecast:</u> Estimate the quantity of the product that will be demanded over a certain period. This could be based on historical sales data, market research, or other forecasting methods.</p> <p><u>Formulate the Problem:</u> Use the costs and demand forecast to formulate the lot-sizing problem as a mathematical model. This typically involves creating an objective function that represents the total cost, which you want to minimize. The function will be subject to constraints such as meeting the demand and not exceeding the production capacity.</p> <p><u>Solve the Problem:</u> Use an optimization algorithm to solve the problem. This could be a simple method like the Economic Order Quantity (EOQ) model for certain cases, or more complex methods like dynamic programming or mixed-integer programming for more complex scenarios.</p> <p><u>Implement the Solution:</u> Use the solution to determine your production plan and lot size. Monitor the results and adjust as necessary.</p> <p>Remember, the specifics can vary greatly depending on the details of your production process and business model. It's often helpful to work with a supply chain analyst or operations research specialist when tackling these kinds of problems. Also, there are software solutions available that can help with lot-sizing and production planning. Please note that this is a simplified explanation. The actual process can be quite complex and may require a deep understanding of your production process and market conditions. It's always a good idea to consult with a professional or use specialized software.</p>

O segundo *prompt* foi imediatamente subsequente ao P1, proporcionando um contexto adequado para a resposta dos modelos. A análise das respostas geradas por cada modelo revelou uma notável semelhança, com ambos delineando um procedimento "passo a passo" para a obtenção do lote de produção. Ambos os modelos ilustram o emprego da técnica QEP em problemas de menor complexidade, enquanto sugerem algoritmos de otimização para desafios mais complexos. É digno de nota que ambas as plataformas se apresentaram corretas nas suas respostas, evidenciando uma compreensão sólida do tema em questão.

O terceiro *prompt* P3 está representado na Tabela 12 e Tabela 13.

Tabela 12 - Resposta ao *prompt* P3 (ChatGPT)

P3: "You are an expert in industrial management. Explain to me in a concise matter what heuristic techniques there are to solve the lot-sizing problem?"
<b>ChatGPT</b>
<p>In industrial management, the lot-sizing problem involves determining the optimal quantity to produce or order over a specific time horizon. Heuristic techniques offer efficient approaches to solve this complex problem. Some common heuristic techniques for lot-sizing include:</p> <p><u>Fixed-Order Quantity (FOQ):</u> This heuristic involves ordering a fixed quantity whenever inventory reaches a reorder point.</p> <p><u>Fixed-Time Period (FTP):</u> Orders are placed at regular intervals regardless of inventory levels, ensuring a constant replenishment schedule.</p>

P3: "You are an expert in industrial management. Explain to me in a concise matter what heuristic techniques there are to solve the lot-sizing problem?"
<b>ChatGPT</b>
<p><u>Wagner-Whitin Algorithm</u>: A dynamic programming-based heuristic that minimizes total costs by considering setup and inventory holding costs over multiple periods.</p> <p><u>Silver-Meal Heuristic</u>: This method balances setup and holding costs by determining the time to reorder based on the incremental cost per period.</p> <p><u>Least Unit Cost (LUC)</u>: Orders are placed to minimize unit cost, considering factors such as setup, holding, and ordering costs.</p> <p><u>Part-Period Balancing (PPB)</u>: This heuristic aims to balance costs by ordering enough to cover demand for a certain portion of the planning horizon.</p> <p>These techniques offer practical and computationally efficient ways to address the lot-sizing problem in various industrial settings. However, it's essential to assess their suitability based on specific constraints, such as production capacities, demand variability, and cost structures.</p>

Tabela 13 - Resposta ao *prompt* P3(Copilot)

P3: "You are an expert in industrial management. Explain to me in a concise matter what heuristic techniques there are to solve the lot-sizing problem?"
<b>CoPilot</b>
<p>Sure, I'd be happy to explain. Heuristic techniques are often used to solve complex problems such as the lot-sizing problem in industrial management. Here are some commonly used heuristic techniques:</p> <p><u>Silver-Meal Heuristic</u>: This technique is based on the principle of minimizing the average cost per period over a certain horizon. It starts with a lot size covering demand for one period and then incrementally adds one period at a time until the average cost per period increases.</p> <p><u>Part-Period Balancing (PPB) Heuristic</u>: This method balances the setup cost and the holding cost for a part of the period. It calculates the cost for each additional period and stops when the setup cost is less than the holding cost for the additional periods.</p> <p><u>Least Unit Cost (LUC) Heuristic</u>: This method calculates the cost per unit for each period and selects the period with the least unit cost.</p> <p><u>Wagner-Whitin Algorithm</u>: Although not a heuristic, it's worth mentioning as it provides an optimal solution to the lot-sizing problem. However, it can be computationally intensive for large problems.</p> <p>Remember, the choice of heuristic can depend on the specific characteristics of your lot-sizing problem. Each heuristic has its strengths and weaknesses, and may perform better in certain scenarios. It's often beneficial to try multiple heuristics and compare their performance.</p>

No terceiro *prompt*, as respostas divergem ligeiramente, contudo, ambas falham ao omitir o QEP como uma das técnicas, notadamente por ser reconhecida como uma das mais elementares. Por outro lado, o ChatGPT demonstra-se mais abrangente ao identificar duas técnicas adicionais.

O quarto e último *prompt* para esta secção tem os resultados na Tabela 14 e Tabela 15.

Tabela 14 - Resposta ao *prompt* P4(ChatGPT)

P4: “You are an expert in industrial management. For the Lot-Sizing problem, Explain to me in a detailed and simple matter how the SM heuristic works”
<b>ChatGPT</b>
<p>The lot-sizing problem is about determining the best way to allocate production over time to minimize costs while satisfying demand. The SM heuristic, which stands for Silver-Meal heuristic, is a simple and effective method to solve this problem.</p> <p><u>Here's how it works: Step 1: Calculate the Silver-Meal Ratio:</u> The Silver-Meal Ratio (SMR) is calculated by dividing the cost of producing one additional unit by the holding cost per unit per period. In mathematical terms, it's <math>\frac{c_{t+1} + h_t}{c_{t+1}}</math>, where <math>c_{t+1}</math> is the cost of producing one more unit in the next period, and <math>h_t</math> is the holding cost per unit in the current period.</p> <p><u>Step 2: Identify the Minimum Silver-Meal Ratio:</u> You calculate the SMR for each period. Then, you identify the period with the minimum SMR.</p> <p><u>Step 3: Determine the Production Schedule:</u> Starting from the first period, you produce units until the period with the minimum SMR is reached. Then, you stop production. Repeat this process, starting from the period following the one where production stopped, until all demand is satisfied.</p> <p><u>Step 4: Adjust the Production Schedule (if needed):</u> If demand is not met after following the steps above, you adjust the production schedule. Usually, this involves redistributing the surplus production among the periods with the lowest SMR.</p> <p>The rationale behind the Silver-Meal heuristic is to prioritize production in periods where the cost of producing additional units is closest to the holding cost. This helps to balance the trade-off between production and inventory holding costs, aiming for an overall cost-efficient solution. However, it's worth noting that while the Silver-Meal heuristic is simple and quick to implement, it may not always produce the optimal solution. In some cases, more complex algorithms or mathematical optimization techniques might be necessary for better results.</p>

Tabela 15 - Resposta ao *prompt* P4(Copilot)

P4: “You are an expert in industrial management. For the Lot-Sizing problem, Explain to me in a detailed and simple matter how the SM heuristic works”
<b>Copilot</b>
<p>Sure, I'd be happy to explain the Silver-Meal (SM) heuristic for the Lot-Sizing problem. The Silver-Meal heuristic is a simple and effective method used in production and inventory management to determine the optimal quantity to order or produce in each period to minimize total cost.</p> <p><b>Initialization:</b> Start with the first period and set it as the current period.</p> <p><b>Calculation:</b> Calculate the average cost per period if the next period is added to the current production run. This is done by dividing the total cost of the production run (including setup and holding costs) by the number of periods in the run.</p> <p><b>Comparison:</b> Compare this average cost with the cost of producing in the next period alone.</p>

P4: “You are an expert in industrial management. For the Lot-Sizing problem, Explain to me in a detailed and simple matter how the SM heuristic works”
<b>Copilot</b>
<p>Decision: If the average cost is less than or equal to the cost of producing in the next period alone, add the next period to the current production run and go back to step 2. If the average cost is greater, do not add the next period to the current production run. Instead, start a new production run beginning with the next period and go back to step 2.</p> <p>Termination: Continue this process until all periods have been evaluated. The key idea behind the Silver-Meal heuristic is to extend the production run as long as the average cost per period is decreasing or remains the same. Once the average cost starts to increase, it’s time to start a new production run.</p> <p>Remember, while the Silver-Meal heuristic is simple and often effective, it does not guarantee an optimal solution for all lot-sizing problems. Other methods may be more appropriate depending on the specific characteristics of the problem.</p>

No quarto *prompt*, destacam-se tanto o ChatGPT quanto o CoPilot, mostrando capacidade de entender e explicar a heurística SM. Do ponto de vista da explicação que melhor informa o utilizador de modo a este aplicar a heurística, a que mais se aproxima do método abordado em capítulos anteriores é o CoPilot, maioritariamente quando menciona o fator de decisão desta heurística. Em conclusão, ambas as plataformas possuem alguma compreensão na área de dimensionamento de lotes, apesar de fornecerem informação por vezes pouco esclarecedora ou errada para alguém que não tenha conhecimento técnico na área, dificultado a sua utilização como “Ajudante” na tomada e decisões de tamanho e ordens de fabrico.

### 3.3. Resolução do Problema através das Plataformas IA

Originalmente, o propósito deste estudo consistiu na obtenção de respostas acerca da quantidade e do período de encomenda, sem priorizar a qualidade da solução em termos de otimização de resultados. Sendo assim, em ambas as plataformas, enfrentaram-se obstáculos para obter respostas precisas quanto à quantidade e aos períodos de encomenda a partir de um único *prompt* ou mesmo múltiplos *prompts*. Apesar de inicialmente terem sido feitas questões em português, optou-se rapidamente pela transição para o inglês, com a expectativa de melhorar os resultados.

Numa versão inicial, com o intuito de explorar as capacidades das plataformas de inteligência artificial, adotou-se uma estratégia mais simplificada, fornecendo instruções à IA para realizar os cálculos de acordo com a heurística lot for lot. As tabelas seguintes apresentam os resultados obtidos pelas duas plataformas de inteligência artificial, ChatGPT e CoPilot. Ambas determinaram corretamente as ordens de fabricação, no entanto, os custos calculados apresentaram-se incorretos, o que pode ser parcialmente atribuído à ausência de instruções específicas sobre o cálculo de custos nos *prompts*, bem como à variedade de técnicas de cálculo de custos empregadas no dimensionamento de lotes.

Na Tabela 16 e Tabela 17 é demonstrado os resultados para a resolução do problema de lotes segundo a heurística Lot-For-Lot.

Tabela 16 - Resposta CoPilot Exemplo PP1

PP1: "Solve this lot problem using the lot for lot heuristic: ordering cost is €300 the stock cost per unit is €2 The demand (D) is as follows: for period 1 the demand is 10; for period 2 the demand is 10; for period 3 the demand is 15; for period 4 the demand is 20; for period 5 the demand is 70; for period 6 the demand is 180; for period 7 the demand is 250, for period 8 the demand is 280; for period 9 the demand is 230; for period 10 the demand is 40; for period 11 the demand is 0; and for period 12 the demand is 10."

Copilot												
The Lot-for-Lot (LFL) heuristic is a simple and commonly used inventory policy. It orders exactly what is needed for each period based on the demand. This means that the order quantity equals the demand for each period. Given your demand (D) for each period, the order quantities using the LFL heuristic would be:												
Period	1	2	3	4	5	6	7	8	9	10	11	12
Demand (D)	10	10	15	20	70	180	250	280	230	40	0	10
Order Quantity	10	10	15	20	70	180	250	280	230	40	0	10

The total cost would be the sum of the ordering cost and the stock cost per unit for each period. The ordering cost is €300 per order and the stock cost per unit is €2. So, the total cost would be:  $Total\ Cost = \sum_{i=1}^{12} (Ordering\ Cost + Stock\ Cost \times Demand_i)$  Substituting the given values:  $Total\ Cost = \sum_{i=1}^{12} (\€300 + \€2 \times Demand_i)$  This will give you the total cost of the inventory policy for the given demand using the Lot-for-Lot heuristic. Please note that this heuristic does not take into account any constraints such as capacity or minimum order quantity. It also assumes that the demand is known and constant for each period. If the demand varies, a different inventory policy might be more appropriate.

Tabela 17 - Resposta ChatGPT Exemplo PP1

PP1: "Solve this lot problem using the lot for lot heuristic: ordering cost is €300 the stock cost per unit is €2 The demand (D) is as follows: for period 1 the demand is 10; for period 2 the demand is 10; for period 3 the demand is 15; for period 4 the demand is 20; for period 5 the demand is 70; for period 6 the demand is 180; for period 7 the demand is 250, for period 8 the demand is 280; for period 9 the demand is 230; for period 10 the demand is 40; for period 11 the demand is 0; and for period 12 the demand is 10."

ChatGPT												
To solve this problem using the lot-for-lot heuristic, we need to calculate the order quantity for each period based on the demand and the given costs. The lot-for-lot heuristic implies that the order quantity is equal to the demand for each period. Therefore, we'll simply order exactly what is needed to meet the demand for each period. Let's calculate the total cost for each period and then sum them up to find the total cost over all periods.												

PP1: "Solve this lot problem using the lot for lot heuristic: ordering cost is €300 the stock cost per unit is €2 The demand (D) is as follows: for period 1 the demand is 10; for period 2 the demand is 10; for period 3 the demand is 15; for period 4 the demand is 20; for period 5 the demand is 70; for period 6 the demand is 180; for period 7 the demand is 250, for period 8 the demand is 280; for period 9 the demand is 230; for period 10 the demand is 40; for period 11 the demand is 0; and for period 12 the demand is 10."

#### ChatGPT

Period	1	2	3	4	5	6	7	8	9	10	11	12
Demand (D)	10	10	15	20	70	180	250	280	230	40	0	10
Order Quantity	10	10	15	20	70	180	250	280	230	40	0	10

Now, summing up the total costs over all periods: Total cost = €320 + €320 + €330 + €340 + €440 + €660 + €800 + €860 + €760 + €380 + €0 + €320 = €5770 So, the total cost using the lot-for-lot heuristic is €5770.

Na fase subsequente, o objetivo foi ajustar a inteligência artificial para que fornecesse valores de produção que estivessem o mais próximo possível do ideal, especificando quando e quais quantidades produzir. Para alcançar isso, introduziu-se no *prompt* a instrução para utilizar o algoritmo de Wagner-Whitin, de forma a efetuar o cálculo das quantidades ótimas. A fim de apresentar os resultados de maneira mais clara, a partir deste ponto serão exibidas somente as quantidades e ordens e custos fornecidos pelas inteligências artificiais, salvo em casos onde seja mencionado o contrário. Na Tabela 18, Tabela 19 e Tabela 20 serão demonstrados os resultados para as plataformas ChatGPT (tabela 18 e 19) e Copilot (tabela 20).

*Prompt:* "Solve this batch problem for me with Wagner Within: the setup cost is €300 the stock cost per unit is €2 per period The demand (D) is as follows: for period 1 the demand is 10; for period 2 the demand is 10; for period 3 the demand is 15; for period 4 the demand is 20; for period 5 the demand is 70; for period 6 the demand is 180; for period 7 the demand is 250, for period 8 the demand is 280; for period 9 the demand is 230; for period 10 the demand is 40; for period 11 the demand is 0; and for period 12 the demand is 10. Your objective is to satisfy the demand in all periods by placing orders in quantities that you will define. By producing more quantity, you will be able to satisfy more than one period and thus place fewer orders. On the other hand, the cost of stock will increase as you have more units in storage. What is the optimal quantity to produce and in which periods?"

Tabela 18 - Resultado ChatGPT (1)

ChatGPT												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	10	10	15	20	70	180	250	270	230	40	0	10
<b>OF a Planear</b>	10	0	35	0	70	180	250	270	230	40	0	10
<b>Stock Inicial</b>	10	0	25	10	60	170	240	270	220	30	-10	0
<b>Stock Final</b>	0	-10	10	-10	-10	-10	-10	-10	-10	-10	-10	-10
<b>Stock Médio</b>	5	-5	18	0	25	80	115	130	105	10	-10	-5

<b>Custo de Lanç.</b>	300 €	<b>a</b>	300 €	<b>h</b>	2 €	<b>D</b>	92,9
<b>Custo de Posse</b>	2 €						
<b>Custos totais Stock</b>	935 €						
<b>Custos Totais Lanç.</b>	2700 €						
<b>Custos Totais</b>	3 635 €						

Como ilustrado pela Tabela 18, o ChatGPT, ao responder ao *prompt*, aplicou o método lot for lot, apesar de mal-executado resultando num custo de 3.635 euros, o que não seria real pois não foram considerados custos de rotura. Em contraste, o CoPilot, para o mesmo *prompt*, não produziu resultados quantitativos, fornecendo apenas uma explicação insuficiente sobre como os resultados poderiam ser obtidos manualmente utilizando o método de Wagner-Whitin. Procurado melhorar os resultados, adotou-se uma nova estratégia: a utilização de dois *prompts* em vez de apenas um. Desta forma, elaborou-se um primeiro *prompt* que introduz o tipo de problema a ser analisado, enquanto o segundo especifica os valores das variáveis envolvidas. Esta abordagem em duas fases visa facilitar a compreensão da informação e obtenção de respostas mais precisas e detalhadas pelas plataformas de inteligência artificial. Os respetivos *prompts* são fornecidos de seguida, bem como a Tabela 19 com os resultados.

*1º Prompt:* "I need your help in solving production batch sizing problems"

*Resposta ChatGPT:* ...

*2º Prompt:* "The cost of holding stock is €2/unit/period. The setup cost is €300. The demand is for Period 1: 10 units Period 2: 10 units Period 3: 15 units Period 4: 20 units Period 5: 70 units Period 6: 180 units Period 7: 250 units Period 8: 280 units Period 9: 230 units Period 10: 40 units Period 11: 0 units Period 12: 10 units Make sure you satisfy the demand, satisfy the demand and use stock storage to avoid having to resort to too many setups, which would increase cost."

Tabela 19 - Resultado ChatGPT (2)

ChatGPT												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	10	10	15	20	70	180	250	270	230	40	0	10
<b>OF a Planear</b>	31	31	39	44	87	155	178	188	175	56	0	31
<b>Stock Inicial</b>	31	52	81	110	177	262	260	198	103	-71	-111	-80
<b>Stock Final</b>	21	42	66	90	107	82	10	-72	-127	-111	-111	-90
<b>Stock Médio</b>	26	47	74	100	142	172	135	63	-12	-91	-111	-85

<b>Custo de Lanç.</b>	300 €
<b>Custo de Posse</b>	2 €
<b>Custos totais Stock</b>	919 €
<b>Custos Totais Lanç.</b>	3300 €
<b>Custos Totais</b>	4 219 €

<b>a</b>	300 €	<b>h</b>	2 €	<b>D</b>	92,1
----------	-------	----------	-----	----------	------

Com a implementação deste conjunto de *prompts*, observou-se que as quantidades fornecidas na Tabela 19 pelo ChatGPT resultaram em rutura de *stock* nos últimos quatro períodos. Por outro lado, o CoPilot forneceu um código Python destinado a calcular a quantidade ótima. No entanto, este código não funcionou corretamente, mesmo após testes e comunicação dos erros à inteligência artificial para correções, situação que evidencia a necessidade de melhorias na formulação dos *prompts* fornecidas às plataformas, de modo a garantir que os algoritmos possam interpretar e executar as tarefas com maior precisão.

Devido à diversidade de respostas geradas para o mesmo *prompt*, observou-se que tanto o ChatGPT quanto o CoPilot, em algumas ocasiões, em vez de resolverem diretamente o problema apresentado pelo utilizador, fornecem código Python para a solução do problema. Contudo, esses códigos, quando executados, raramente funcionam e, nos casos em que operam corretamente, os valores gerados são no máximo equivalentes àqueles produzidos pela técnica lot-for-lot, o que representa uma qualidade de solução substancialmente inferior em comparação à ótima. Ademais, após várias iterações, foi notado que frequentemente o ChatGPT emprega a técnica QEP, porém falha em aplicá-la adequadamente, resultando em valores de produção irrazoáveis.

O melhor resultado até o momento foi alcançado utilizando o CoPilot, IA da Microsoft, na Tabela 20. Embora seja o melhor resultado obtido, ele apenas se aproxima do desempenho da heurística POQ, Tabela 21. Este resultado é apresentado na Tabela 20:

Tabela 20 - Resultados CoPilot (1)

CoPilot												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	10	10	15	20	70	180	250	270	230	40	0	10
<b>OF a Planear</b>	20	0	35	0	250	0	530	0	270	0	0	10
<b>Stock Inicial</b>	20	10	35	20	250	180	530	280	280	50	10	20
<b>Stock Final</b>	10	0	20	0	180	0	280	10	50	10	10	10
<b>Stock Médio</b>	15	5	28	10	215	90	405	145	165	30	10	15

<b>Custo de Lanç.</b>	300 €
<b>Custo de Posse</b>	2 €
<b>Custos totais Stock</b>	2265 €
<b>Custos Totais Lanç.</b>	1800 €
<b>Custos Totais</b>	4065 €

<b>a</b>	300 €	<b>h</b>	2 €	<b>D</b>	92,1
----------	-------	----------	-----	----------	------

O respetivo cálculo pela heurística POQ é apresentado na tabela seguinte:

Tabela 21 - Cálculo Heurística POQ

DIMENSIONAMENTO DE LOTES - POQ												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	10	10	15	20	70	180	250	270	230	40	0	10
<b>OF a Planear</b>	20	0	35	0	250	0	520	0	270	0	10	0
<b>Stock Inicial</b>	20	10	35	20	250	180	520	270	270	40	10	10
<b>Stock Final</b>	10	0	20	0	180	0	270	0	40	0	10	0
<b>Stock Médio</b>	15	5	28	10	215	90	395	135	155	20	10	5

<b>Custo de Lanç.</b>	300 €
<b>Custo de Posse</b>	2 €
<b>Custos totais Stock</b>	1800 €
<b>Custos Totais Lanç.</b>	2165 €
<b>Custos Totais</b>	3965 €

<b>a</b>	300 €	<b>h</b>	2 €	<b>D</b>	92,1
----------	-------	----------	-----	----------	------

## ChatGPT 4

Devido à irregularidade de respostas, qualidade de solução das mesmas e variação de resposta para os mesmos *prompts*, recorreu-se ao ChatGPT 4 para resolver o problema proposto. O ChatGPT 4 é a versão paga do ChatGPT e o modelo mais potente e inteligente no momento.

Esta versão do modelo GPT oferece capacidades melhoradas de compreensão e geração de texto, sendo concebida para interpretar de forma mais eficaz as nuances e complexidades dos *prompts* dos utilizadores. Além disso, o ChatGPT 4 foi submetido a um treino extensivo em uma diversidade maior de dados, possibilitando uma resposta mais precisa e contextualizada aos desafios apresentados, o que se espera que contribua para uma melhoria substancial na resolução de problemas complexos e na precisão dos cálculos relacionados com as exigências do projeto.

Ao utilizar o mesmo *prompt* que anteriormente foi utilizado com o ChatGPT 3.5 e o CoPilot, observou-se de imediato uma diferença significativa na inteligência e eficácia do modelo ChatGPT 4. Em resposta ao *prompt* fornecido, este modelo gerou um código Python (Anexo A) para auxiliar no processamento, o qual não só funcionou corretamente à primeira tentativa, mas também forneceu um resultado ótimo. Vale ressaltar que a utilização de código Python como suporte na resolução de problemas não é uma novidade exclusiva deste modelo, no entanto, a eficiência e a precisão com que o ChatGPT 4 executa e aplica esse código distinguem-no claramente dos seus predecessores. Esta capacidade de gerar soluções eficazes e precisas, sem necessidade de ajustes ou correções adicionais, sublinha o avanço tecnológico e a superioridade deste modelo. Na Tabela 22 é mostrado o resultado para o seguinte *prompt* utilizado:

**Prompt:** "Solve this batch problem for me :12 time periods the cost of the order is €300 the cost of ownership per unit is €2 The demand (D) is as follows: for period 1 the demand is 10; for period 2 the demand is 10; for period 3 the demand is 15; for period 4 the demand is 20; for period 5 the demand is 70; for period 6 the demand is 180; for period 7 the demand is 250, for period 8 the demand is 280; for period 9 the demand is 230; for period 10 the demand is 40; for period 11 the demand is 0; and for period 12 the demand is 10. Your goal is to satisfy the demand in all periods to place orders for the details you want. By ordering more quantity, you can be satisfied for more than one period and thus place fewer orders, on the other hand the cost of ownership will increase as there are more units in storage. What is the optimal quantity to order and in what periods."

Tabela 22 - Resultado ChatGPT 4

ChatGPT 4.0												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	10	10	15	20	70	180	250	280	230	40	0	10
<b>OF a Planear</b>	55	0	0	0	70	180	250	280	280	0	0	0
<b>Stock Inicial</b>	55	45	35	20	70	180	250	280	280	50	10	10
<b>Stock Final</b>	45	35	20	0	0	0	0	0	50	10	10	0
<b>Stock Médio</b>	50	40	28	10	35	90	125	140	165	30	10	5

<b>Custo de Lanç.</b>	300 €
<b>Custo de Posse</b>	2 €
<b>Custos totais Stock</b>	1455 €
<b>Custos Totais Lanç.</b>	1800 €
<b>Custos Totais</b>	3255 €

<b>a</b>	300 €	<b>h</b>	2 €	<b>D</b>	92,9
----------	-------	----------	-----	----------	------

Na Tabela 23 é o apresentado o resultado ótimo para comparação:

Tabela 23 - Resultado Wagner-Within

DIMENSIONAMENTO DE LOTES - WW												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	10	10	15	20	70	180	250	280	230	40	0	10
<b>OF a Planear</b>	55	0	0	0	70	180	250	280	280	0	0	0
<b>Stock Inicial</b>	55	45	35	20	70	180	250	280	280	50	10	10
<b>Stock Final</b>	45	35	20	0	0	0	0	0	50	10	10	0
<b>Stock Médio</b>	50	40	28	10	35	90	125	140	165	30	10	5

<b>Custo de Lanç.</b>	300 €
<b>Custo de Posse</b>	2 €
<b>Custos totais Stock</b>	1455 €
<b>Custos Totais Lanç.</b>	1800 €
<b>Custos Totais</b>	3255 €

<b>a</b>	300 €	<b>h</b>	2 €	<b>D</b>	92,9
----------	-------	----------	-----	----------	------

Além de fornecer a resposta em formato de texto como é mostrado na Figura 24, o modelo ainda fornece o código, que funciona sem falhas (Anexo A), mesmo ao modificar ligeiramente para calcular novas demandas e custos.

The optimal strategy for placing orders to minimize the total costs, given the demands and costs, results in a total cost of €2140. The ordering strategy is as follows:

1. Place an order of 55 units at the beginning of period 1 to cover the demand for periods 1 through 5.
2. Place an order of 70 units at the beginning of period 5 to cover the demand for period 5.
3. Place an order of 180 units at the beginning of period 6 to cover the demand for period 6.
4. Place an order of 250 units at the beginning of period 7 to cover the demand for period 7.
5. Place an order of 280 units at the beginning of period 8 to cover the demand for period 8.
6. Place an order of 280 units at the beginning of period 9 to cover the demand for period 9.

This strategy ensures that the demands are met for each period while optimizing the total cost, taking into account both ordering and holding costs. [-]

Figura 24 - Resposta ChatGPT 4 (Codigo no Anexo A)

A Tabela 24 mostra a aplicação do ChatGPT 4 noutra problema de dimensionamento de lotes através de um segundo *prompt* de seguida ao apresentado anteriormente, identificando as modificações nos custos e procura para o novo problema. É também apresentada a resposta do ChatGPT na Figura 25, bem como o resultado ótimo na Tabela 25.

Tabela 24 - Resultado ChatGPT 4 (2)

ChatGPT 4.0												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	30	60	10	120	160	140	80	50	120	160	220	50
<b>OF a Planear</b>	100	0	0	120	160	140	130	0	120	160	270	0
<b>Stock Inicial</b>	100	70	10	120	160	140	130	50	120	160	270	50
<b>Stock Final</b>	70	10	0	0	0	0	50	0	0	0	50	0
<b>Stock Médio</b>	85	40	5	60	80	70	90	25	60	80	160	25

<b>Custo de Lanç.</b>	5400 €	<b>a</b>	5400 €	<b>h</b>	40 €	<b>D</b>	100
<b>Custo de Posse</b>	50 €						
<b>Custos totais Stock</b>	36400 €						
<b>Custos Totais Lanç.</b>	37800 €						
<b>Custos Totais</b>	74200 €						

Tabela 25 - Resultado Wagner-Within (2)


DIMENSIONAMENTO DE LOTES - WW												
	1	2	3	4	5	6	7	8	9	10	11	12
<b>Necessd. Líq.</b>	30	60	10	120	160	140	80	50	120	160	220	50
<b>OF a Planear</b>	100	0	0	120	160	140	130	0	120	160	270	0
<b>Stock Inicial</b>	100	70	10	120	160	140	130	50	120	160	270	50
<b>Stock Final</b>	70	10	0	0	0	0	50	0	0	0	50	0
<b>Stock Médio</b>	85	40	5	60	80	70	90	25	60	80	160	25

<b>Custo de Lanç.</b>	5400 €
<b>Custo de Posse</b>	50 €
<b>Custos totais Stock</b>	36400 €
<b>Custos Totais Lanç.</b>	37800 €
<b>Custos Totais</b>	74200 €

<b>a</b>	5400 €	<b>h</b>	40 €	<b>D</b>	100
----------	--------	----------	------	----------	-----

 For the updated demands and costs, the optimal strategy to minimize the total costs results in a total cost of €52,200. The ordering strategy is as follows:

1. Place an order of 100 units at the beginning of period 1 to cover the demand for periods 1 through 3.
2. Place an order of 120 units at the beginning of period 4 to cover the demand for period 4.
3. Place an order of 160 units at the beginning of period 5 to cover the demand for period 5.
4. Place an order of 140 units at the beginning of period 6 to cover the demand for period 6.
5. Place an order of 130 units at the beginning of period 7 to cover the demand for periods 7 and 8.
6. Place an order of 120 units at the beginning of period 9 to cover the demand for period 9.
7. Place an order of 160 units at the beginning of period 10 to cover the demand for period 10.
8. Place an order of 270 units at the beginning of period 11 to cover the demand for periods 11 and 12.

This ordering schedule ensures that the demands are met for each period, factoring in the higher order and holding costs to optimize the total cost. [\[↩\]](#)

Figura 25 - Resposta ChatGPT 4 (2)

Novamente o ChatGPT 4 fornece o resultado ótimo, apesar de ainda não calcular os custos da forma correta, ou pelo menos da forma pretendida. Outro fator a realçar é que apesar de fornecer resultados muito satisfatórios, neste instante, foi identificado que para o *prompt* previamente usado os resultados variavam, por vezes não fornecendo as ordens de produção, deixando possibilidade para otimizar mais o *prompt*.

Sendo assim a próxima e última fase será tentar obter um *prompt* mais estável, que origine respostas sempre com as ordens de produção, com as quantidades ótimas, e que ainda forneça o custo correto.

### **Prompt para calcular os Custos**

Visto o ChatGPT não calcular os custos corretamente, formulou-se um novo *prompt*. Assim, foi fornecido um *prompt* muito semelhante as anteriores, mas com instruções específicas para aplicar a fórmula da EQ.5. Esta Equação apresenta a forma de calcular o custo do algoritmo de Wagner-Whitin. Com isso, esperava-se que o ChatGPT fornecesse tanto o resultado ótimo quanto o custo exato. A EQ. 5 foi previamente abordada no capítulo 2:

$$C_{t,N+1} = \min_{t < k \leq N+1} \{E_{t,k} + C_{k,N+1}\} \quad t = N, \dots, 2, 1$$

O *prompt* utilizado é apresentado na Tabela 26.

Tabela 26 – Prompt cálculo dos Custos

<p><i>“Solve this batch problem for me: 12 time periods the cost of the order is €5400 the cost of stock per unit is €40 per time period. The demand (D) is as follows: for period 1 the demand is 30; for period 2 the demand is 60; for period 3 the demand is 10; for period 4 the demand is 120; for period 5 the demand is 160; for period 6 the demand is 140; for period 7 the demand is 80, for period 8 the demand is 50; for period 9 the demand is 120; for period 10 the demand is 160; for period 11 the demand is 220; and for period 12 the demand is 50. Your goal is to satisfy the demand in all periods to place orders for the details you want. By ordering more quantity, you can be satisfied for more than one period and thus place fewer orders, on the other hand the cost of holding the stock will increase as there are more units in storage. What is the optimal quantity to order and in what periods. Be sure to use dynamic programming to achieve optimal quantities. Use the provided formula to calculate costs.”</i></p>
<b>ChatGPT 4.0</b>
<ul style="list-style-type: none"> <li>• Total minimum cost: €50,200.</li> <li>• Ordering Points and Quantities:             <ul style="list-style-type: none"> <li>• Order at Period 1 for Periods 1 to 3: Order 100 units.</li> <li>• Order at Period 4 for Period 4: Order 120 units.</li> <li>• Order at Period 5 for Period 5: Order 160 units.</li> <li>• Order at Period 6 for Periods 6 to 8: Order 270 units.</li> <li>• Order at Period 9 for Period 9: Order 120 units.</li> <li>• Order at Period 10 for Period 10: Order 160 units.</li> <li>• Order at Period 11 for Periods 11 to 12: Order 270 units.</li> </ul> </li> </ul>

Os dados utilizados provinham do exemplo da Tabela 24, e como é possível analisar na Tabela 26, o ChatGPT voltou a indicar corretamente o padrão de ordens de fabrico, mas voltou a não conseguir calcular corretamente o custo.

Apesar de múltiplas tentativas, os custos de armazenamento foram consistentemente calculados erradamente, possivelmente devido a dificuldades na implementação correta da lógica para calcular o *stock* excedente após cada período. A incapacidade de acumular precisamente os custos de armazenamento para o *stock* restante de cada período, sugere um problema na formulação do problema por parte do ChatGPT. Sendo assim apenas será avaliada a capacidade do ChatGPT fornecer a quantidade de produção ótima.

### 3.4. Conclusão

Neste capítulo foi explorada a utilização das plataformas ChatGPT e Copilot na resolução do problema de dimensionamento de lotes, com o objetivo de avaliar a eficácia das mesmas. Inicialmente, os resultados obtidos foram insatisfatórios em ambas as plataformas, com desempenho inconsistente e uma falta de repetibilidade na solução dos problemas, exceto em alguns casos pontuais. Essa inconsistência demonstrou as limitações dos modelos em abordar de forma robusta e repetida o problema em estudo.

Para superar estas limitações, foram realizadas novas tentativas utilizando o modelo GPT-4 do ChatGPT. Esta abordagem mostrou-se mais promissora, apresentando resultados com maior potencial e consistência na resolução dos exercícios propostos. O GPT-4 conseguiu capturar melhor o contexto e os requisitos do problema, resultando em soluções com a quantidade ótima com uma certa frequência.

Dada a superioridade dos resultados obtidos com o GPT-4, foi decidido que esta versão do ChatGPT será a ferramenta escolhida para a realização dos cálculos e análises no Capítulo 4, dedicado aos resultados e discussões.

Assim, o Capítulo 3 conclui-se com a definição clara de que o ChatGPT 4 oferece uma abordagem mais robusta e eficaz para o problema em estudo, estabelecendo uma base sólida para as análises e discussões que se seguirão no Capítulo 4. Esta decisão aponta para o potencial contínuo do GPT-4 em apoiar soluções avançadas e inovadoras para problemas de otimização, promovendo um avanço significativo na utilização de IA na área de dimensionamento de lotes.

## 4. Resultados e Discussão

Este capítulo será dedicado à aplicação prática e análise dos resultados obtidos pelo uso do *prompt* desenvolvido no Capítulo 3. Este *prompt*, concebido para obter as ordens de fabrico ótimas para um problema de dimensionamento de lotes, será aplicado a duas amostras, cada uma com 30 instâncias que foram geradas seguindo uma distribuição uniforme.

Posteriormente, os resultados serão submetidos a uma avaliação detalhada utilizando *software* estatístico avançado, mais concretamente o SPSS (Statistical Package for the Social Science). Tais ferramentas permitirão não apenas a quantificação dos resultados, mas também uma discussão profunda sobre as tendências observadas, a significância dos resultados e as possíveis implicações práticas ou teóricas derivadas deste estudo.

### 4.1. Implementação do ChatGPT

Para a criação da primeira amostra, optou-se por utilizar uma distribuição uniforme ( $x \in [100,150]$ ). Foram considerados custos de *stock* de 2€/unidade/por período e um custo de encomenda de 300€. O objetivo da primeira amostra foi representar um produto com baixa variação nas necessidades líquidas e baixo custo de *stock*.

Para a segunda amostra, decidiu-se utilizar uma distribuição uniforme ( $x \in [20,150]$ ). Foram considerados custos de *stock* de 20€/unidade/por período e um custo de encomenda de 2000€. O objetivo da segunda amostra foi representar um produto com alta variação nas necessidades líquidas e alto custo de armazenamento.

Dessa maneira, será possível avaliar a performance do ChatGPT em diferentes cenários.

Relativamente ao *prompt* utilizado, este foi sofrendo alterações enquanto se faziam os testes da performance, sendo utilizado para todas as instâncias em ambas as amostras o seguinte:

*“Solve this batch problem for me, it consists of 12 time periods, the cost of order is 2000€*

*The stock cost per unit is 20€ per period.*

*The demand (D) is as follows: for period 1 the demand is 103; for period 2 the demand is 61; for period 3 the demand is 43; for period 4 the demand is 34; for period 5 the demand is 60; for period 6 the demand is 95; for period 7 the demand is 128, for period 8 the demand is 97; for period 9 the demand is 42; for period 10 the demand is 118; for period 11 the demand is 56; and for period 12 the demand is 99. You must satisfy the demand in all periods and optimize the ordering plan. By ordering more quantity, you can satisfy more than one period and place fewer*

orders. You Must Calculate optimal quantity and for each period using the Wagner within method. Recalculate once to make sure solution provided is the optimal. At the end present the ordering plan with quantities for each period."

## **4.2. Apresentação dos resultados**

Para avaliar a qualidade das soluções geradas pelo ChatGPT, foi decidido resolver todas as instâncias de cada amostra utilizando quatro métodos distintos: o Método da Quantidade Económica de Produção (QEP), o Método de Silver-Meal (SM), o Parts Period Balancing (PPB) e o Método de Wagner-Whitin, este último servirá como controlo e para efeitos de cálculo do desvio do valor ótimo por cada técnica. Os resultados para ambas as amostras são apresentados na Tabela 27 e Tabela 28. Para facilitar a compreensão dos dados, foram também desenvolvidos os gráficos nas Figura 26, Figura 27 e Figura 28.

É relevante destacar que os cálculos realizados pelo ChatGPT foram obtidos utilizando o modelo mais recente, o ChatGPT-4o, que oferece maior capacidade analítica e habilidade para resolver cálculos complexos em comparação ao seu antecessor, o ChatGPT-4. Para avaliar a superioridade do modelo mais recente, também foram incluídos os cálculos da amostra 1 com o uso ChatGPT-4, permitindo uma comparação direta entre os dois modelos.

Dada a variabilidade das respostas inerentes ao uso do ChatGPT como ferramenta para o cálculo dos lotes, foram realizadas até quatro tentativas por instância. Este procedimento foi necessário nos casos em que o ChatGPT não fornecia os dados de quantidade ou quando os valores gerados eram inválidos, resultando em custos desproporcionais e inviáveis.

Tabela 27 - Resultados Amostra 1

	Ints.	Desvio ChatGPT-4	Desvio ChatGPT-4o	Desvio WW	Desvio SM	Desvio PPB	Desvio QEP
Amostra 1	1	0,00%	5,79%	0,00%	5,79%	1,33%	33,83%
	2	0,00%	0,00%	0,00%	1,94%	4,79%	25,99%
	3	3,96%	3,96%	0,00%	0,00%	3,13%	33,82%
	4	0,00%	0,00%	0,00%	0,00%	2,03%	28,68%
	5	4,29%	4,29%	0,00%	0,00%	6,16%	49,72%
	6	0,00%	0,00%	0,00%	0,00%	2,60%	21,80%
	7	0,00%	1,63%	0,00%	1,25%	2,41%	33,02%
	8	6,98%	0,00%	0,00%	0,34%	1,68%	26,73%
	9	0,00%	0,00%	0,00%	0,00%	4,85%	35,79%
	10	6,43%	0,00%	0,00%	0,42%	1,17%	34,01%
	11	0,00%	0,00%	0,00%	0,00%	4,24%	30,80%
	12	3,73%	0,00%	0,00%	1,44%	4,06%	29,17%
	13	0,00%	0,00%	0,00%	0,00%	3,63%	27,29%
	14	0,00%	0,00%	0,00%	0,97%	2,88%	33,95%
	15	0,00%	0,00%	0,00%	3,38%	2,03%	37,19%
	16	0,00%	0,00%	0,00%	0,00%	4,50%	34,88%
	17	2,20%	0,00%	0,00%	0,04%	3,64%	30,99%
	18	0,00%	0,00%	0,00%	0,00%	6,34%	30,10%
	19	0,00%	0,00%	0,00%	0,57%	1,95%	32,15%
	20	7,29%	1,52%	0,00%	0,59%	1,52%	41,59%
	21	5,97%	0,00%	0,00%	1,52%	2,53%	36,76%
	22	0,00%	0,00%	0,00%	1,12%	1,30%	37,09%
	23	1,49%	0,00%	0,00%	0,51%	2,97%	36,51%
	24	3,50%	0,00%	0,00%	0,50%	5,12%	29,07%
	25	0,00%	0,00%	0,00%	0,64%	2,21%	29,18%
	26	0,00%	0,00%	0,00%	0,00%	2,01%	29,26%
	27	3,56%	0,00%	0,00%	1,59%	1,22%	33,19%
	28	0,00%	0,95%	0,00%	0,65%	1,21%	26,86%
	29	0,00%	0,00%	0,00%	0,00%	2,33%	32,09%
	30	0,57%	0,00%	0,00%	1,92%	4,25%	27,31%
Média	1,67%	0,60%	0,00%	0,84%	3,00%	32,29%	
Mediana	0,00%	0,00%	0,00%	0,50%	2,56%	32,12%	
Desvio padrão	2,46%	1,47%	0,00%	1,24%	1,49%	5,34%	
Variancia	0,06%	0,02%	0,00%	0,02%	0,02%	0,29%	

Tabela 28 - Resultados Amostra 2

	Ints.	Desvio ChatGPT-4o	Desvio WW	Desvio SM	Desvio PPB	Desvio QEP
Amostra 2	1	0,00%	0,00%	3,85%	3,37%	45,12%
	2	0,00%	0,00%	2,03%	0,63%	47,53%
	3	0,00%	0,00%	5,04%	0,00%	29,55%
	4	0,00%	0,00%	4,37%	2,91%	22,25%
	5	0,00%	0,00%	3,38%	0,00%	17,25%
	6	0,00%	0,00%	0,00%	2,68%	27,65%
	7	0,00%	0,00%	0,00%	0,00%	31,87%
	8	0,00%	0,00%	0,00%	4,63%	47,37%
	9	0,00%	0,00%	0,51%	7,01%	46,22%
	10	0,00%	0,00%	2,92%	4,04%	37,23%
	11	0,00%	0,00%	2,77%	6,44%	31,95%
	12	0,00%	0,00%	2,94%	2,50%	29,37%
	13	0,00%	0,00%	0,66%	3,51%	51,14%
	14	13,03%	0,00%	0,55%	0,41%	25,71%
	15	0,00%	0,00%	1,00%	2,44%	38,21%
	16	0,00%	0,00%	4,47%	6,24%	38,57%
	17	3,31%	0,00%	5,97%	1,82%	24,60%
	18	0,00%	0,00%	2,05%	5,13%	35,92%
	19	0,00%	0,00%	1,44%	3,48%	24,56%
	20	6,91%	0,00%	0,00%	4,74%	25,19%
	21	0,00%	0,00%	0,00%	2,28%	42,22%
	22	0,00%	0,00%	1,27%	0,00%	25,36%
	23	0,00%	0,00%	0,79%	1,34%	28,09%
	24	0,00%	0,00%	1,54%	3,89%	29,65%
	25	0,00%	0,00%	0,13%	5,51%	35,59%
	26	0,00%	0,00%	0,34%	3,09%	48,09%
	27	0,00%	0,00%	1,73%	3,21%	39,04%
	28	1,12%	0,00%	2,05%	2,61%	57,79%
	29	0,00%	0,00%	4,80%	0,18%	28,35%
	30	0,00%	0,00%	0,00%	6,06%	40,11%
Média	0,81%	0,00%	1,89%	3,01%	35,05%	
Mediana	0,00%	0,00%	1,49%	3,00%	33,77%	
Desvio padrão	2,69%	0,00%	1,78%	2,11%	9,89%	
Variancia	0,07%	0,00%	0,03%	0,04%	0,98%	

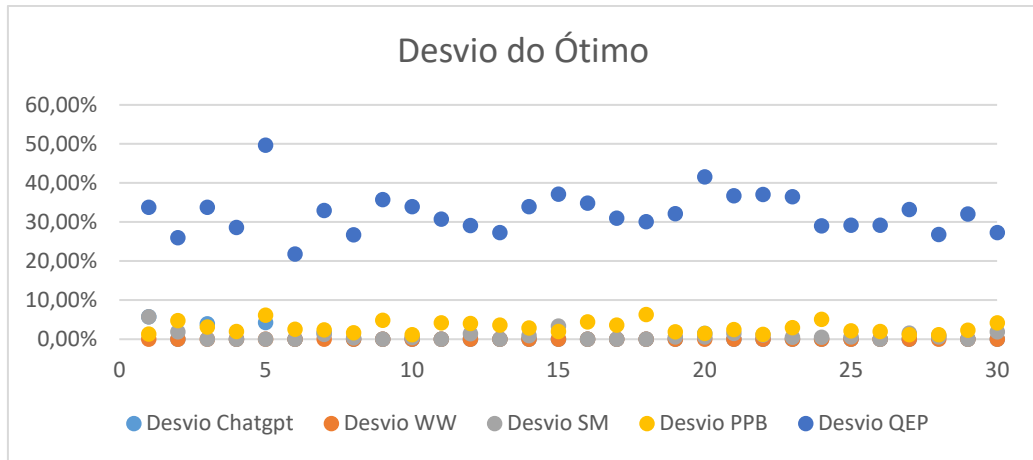


Figura 26 - Gráfico Desvios Amostra 1 (ChatGPT-4o)

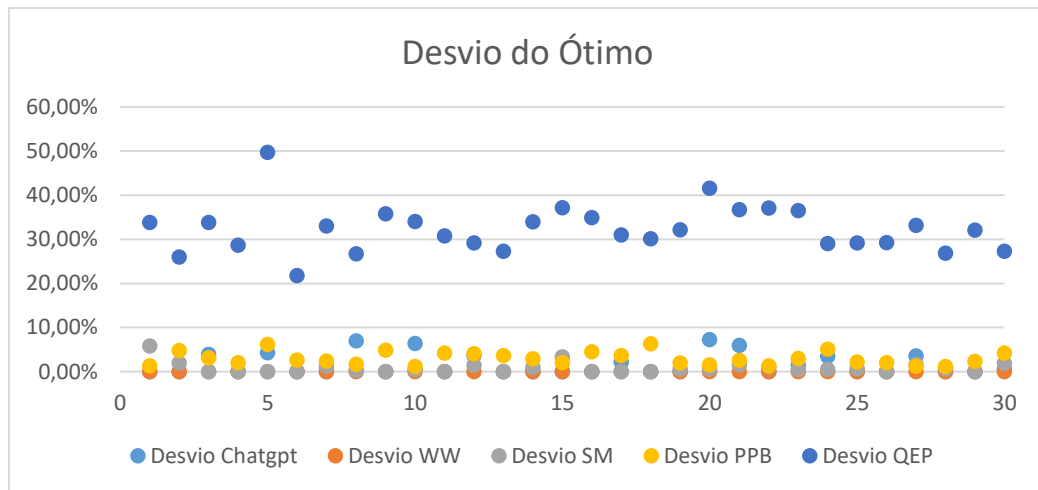


Figura 27 - Gráfico Desvio Amostra 1 (ChatGPT-4)

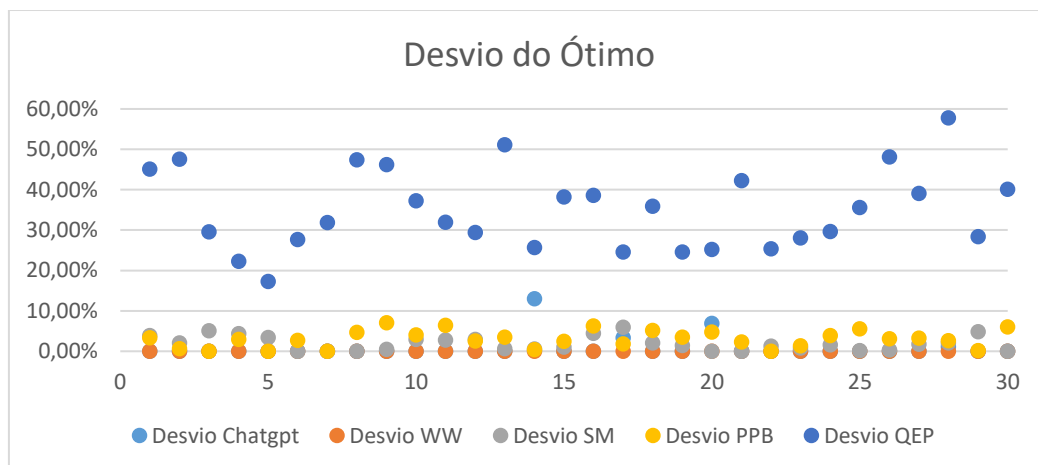


Figura 28 - Gráfico Desvios Amostra 2

## 4.3. Análise dos Resultados

Os resultados obtidos com o modelo GPT-4 mostram-se semelhantes aos alcançados com a técnica SM, embora apresentem uma maior variabilidade. Apesar disso, os desempenhos iniciais do GPT-4 aparentam ser superiores quando comparados com as demais técnicas analisadas. Por outro lado, o modelo GPT-4o parece ser ainda mais eficaz em comparação com o SM, exibindo variações de resultados semelhantes, mas com uma proximidade mais consistente ao resultado ótimo desejado. Este facto pode indicar que, embora ambos os modelos apresentem variações, o GPT-4o consiga manter um desempenho superior de forma mais regular. Além disso, o modelo ChatGPT-4o, assim como o GPT-4, destacam-se por aparentar superar as outras técnicas testadas.

### 4.3.1. Estatística Descritiva da Aplicação Do ChatGPT ao Problema

#### Amostra 1

Com base nos dados anteriores, o ChatGPT-4o apresentou uma melhoria significativa em relação ao ChatGPT-4, com uma média de desvio relativo em relação a solução do WW de 0,60%, comparado a 1,67% no ChatGPT 4. Essa diferença sugere que o ChatGPT-4o oferece maior precisão e confiabilidade nos seus resultados. A mediana do ChatGPT-4o foi 0,00%, o que indica que, na maioria dos casos, os resultados foram ótimos, algo que o ChatGPT-4 não conseguiu alcançar com a mesma consistência. O desvio padrão para o ChatGPT-4o foi menor (1,47%), indicando menos variabilidade nos resultados em comparação ao ChatGPT-4, que teve um desvio padrão de 2,46%. O ChatGPT-4o aproxima-se muito do WW, com desvios pequenos e, em muitos casos, também alcançando um desvio relativo de 0,00%.

O método Silver Meal (SM) apresentou uma média de desvio de 0,84%, mostrando-se relativamente preciso. Foi inferior ao desempenho do ChatGPT-4o, mas superior ao CHatGPT-4. O Parts Period Balancing (PPB) teve uma média de desvio de 3,00%, o que indica menor precisão e maior variabilidade, com desvios padrões e variância levemente superiores. O QEP apresentou o pior desempenho, com uma média de desvio de 32,29%, refletindo a sua baixa precisão e alta variabilidade nos resultados. Este método parece inadequado para cenários representados pela Amostra 1.

Com base na Tabela 29, Tabela 30 e Tabela 31, constata-se que em relação à variabilidade das técnicas utilizadas, na amostra 1, observa-se que o ChatGPT-4o e ChatGPT-4 apresentam em 24 e 18 instâncias os resultados ótimos, respetivamente. Em contrapartida, técnicas como SM e PPB mostram maior variabilidade, com desvios mais distribuídos entre as faixas de 1 a 3% e 4 a 8% de desvios relativos do ótimo. Já o QEP apresenta o pior desempenho, com 100% dos desvios relativos do ótimo acima de 16%, indicando baixa qualidade dos resultados. Na Figura 29 é fornecido um gráfico para melhor compreensão e análise dos dados.

Tabela 29 - Lista de Frequências Amostra 1

ChatGPT 4o - Frequências			ChatGPT 4 - Frequências		
0%	24	80,00%	0%	18	60,00%
1 a 3%	3	10,00%	1 a 3%	3	10,00%
4 a 8%	3	10,00%	4 a 8%	9	30,00%
9 a 15%	0	0,00%	9 a 15%	0	0,00%
>16%	0	0,00%	>16%	0	0,00%

Tabela 30 - Lista de Frequências Amostra 1

WW - Frequências			SM - Frequências		
0%	30	100,00%	0%	11	36,67%
1 a 3%	0	0,00%	1 a 3%	17	56,67%
4 a 8%	0	0,00%	4 a 8%	2	6,67%
9 a 15%	0	0,00%	9 a 15%	0	0,00%
>16%	0	0,00%	>16%	0	0,00%

Tabela 31 - Lista de Frequências Amostra 1

PPB - Frequências			QEP - Frequências		
0%	0	0,00%	0%	0	0,00%
1 a 3%	18	60,00%	1 a 3%	0	0,00%
4 a 8%	11	36,67%	4 a 8%	0	0,00%
9 a 15%	0	0,00%	9 a 15%	0	0,00%
>16%	0	0,00%	>16%	30	100,00%

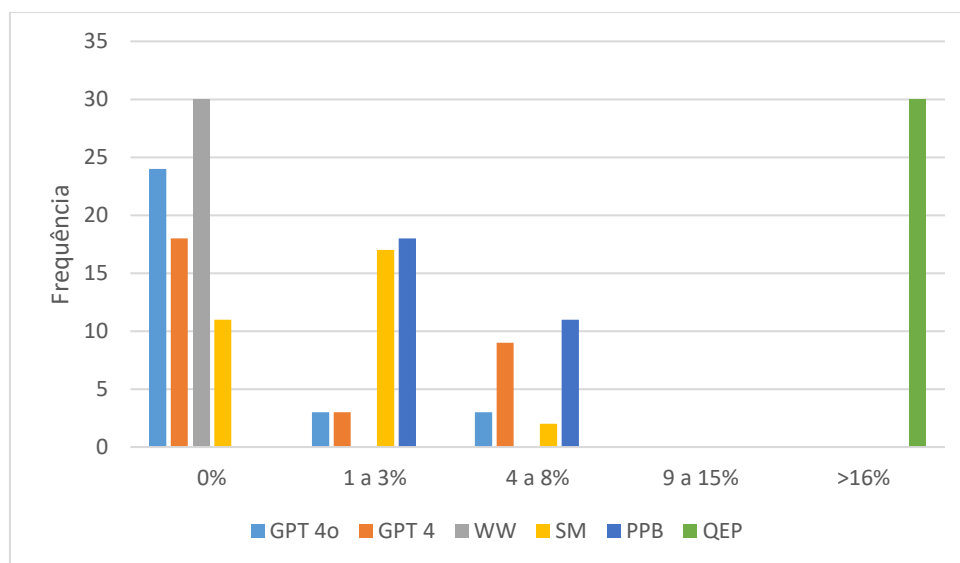


Figura 29 - Gráfico de barras das Frequências do desvio (Amostra 1)

## Amostra 2

O ChatGPT-4o apresentou uma média de desvio relativo em relação ao WW de 0,81%. Apesar de o ChatGPT-4o ter apresentado pequenos desvios, a cerca de 87% das instâncias apresentaram desvio 0,00%, indicando que, na grande maioria dos casos, o seu desempenho foi equivalente ao método WW. O desvio padrão de 2,69% para o ChatGPT-4o sugere que houve variações pontuais, mas, de modo geral, o modelo mostrou alta precisão e consistência próxima ao resultado ótimo representado pelo WW. O método Silver Meal (SM) apresentou uma média de desvio de 1,89%, indicando desempenho razoavelmente preciso, porém inferior ao ChatGPT-4o. O Parts Period Balancing (PPB) teve uma média de desvio de 3,01%, demonstrando maior variação e menor precisão comparado aos métodos anteriores. O Quantidade Económica de Pedido (QEP) exibiu o maior desvio médio, com 35,05%, indicando baixa precisão e significativa discrepância em relação ao resultado ótimo. Os desvios padrões correspondentes mostram que o QEP também teve a maior variabilidade nos resultados, reforçando sua inconsistência em cenários com as características da Amostra 2.

Com base na Tabela 32, Tabela 33, Tabela 34 constata-se que, em relação à variabilidade das técnicas utilizadas na amostra 1, o ChatGPT 4o volta a destacar-se por apresentar elevada precisão, com 26 instâncias a 0% de desvio relativos do ótimo. Em contrapartida, a técnica SM revela uma maior variabilidade, com apenas 6 instâncias a 0% e a maioria dos desvios relativos do ótimo (56,67%) situados entre 1 a 3%. A técnica PPB apresenta resultados moderados, com 50% das instâncias com desvios relativos do ótimo entre 4 a 8% de desvio. Já o QEP volta a obter o pior desempenho, com 100% dos desvios relativos dos ótimos superiores a 16%, indicando baixa precisão. Na Figura 30 é fornecido um gráfico para melhor compreensão e análise dos dados.

Tabela 32 - Lista de Frequências Amostra 2

ChatGPT 4o - Frequências		
0%	26	86,67%
1 a 3%	1	3,33%
4 a 8%	2	6,67%
9 a 15%	1	3,33%
>16%	0	0,00%

Tabela 33- Lista de Frequências Amostra 2

WW - Frequências			SM - Frequências		
0%	30	100,00%	0%	6	20,00%
1 a 3%	0	0,00%	1 a 3%	17	56,67%
4 a 8%	0	0,00%	4 a 8%	7	23,33%
9 a 15%	0	0,00%	9 a 15%	0	0,00%
>16%	0	0,00%	>16%	0	0,00%

Tabela 34 - Lista de Frequências Amostra 2

PPB - Frequências			QEP - Frequências		
0%	4	13,33%	0%	0	0,00%
1 a 3%	11	36,67%	1 a 3%	0	0,00%
4 a 8%	15	50,00%	4 a 8%	0	0,00%
9 a 15%	0	0,00%	9 a 15%	0	0,00%
>16%	0	0,00%	>16%	30	100,00%

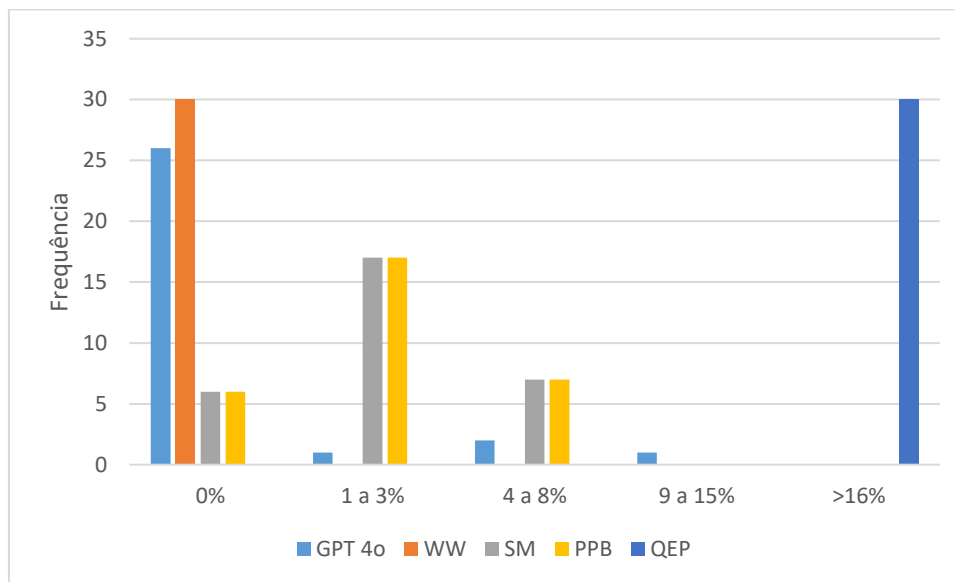


Figura 30 - Gráfico de barras das Frequências do desvio (Amostra 2)

#### 4.3.2. Inferência Estatística da Aplicação Do ChatGPT ao Problema

Para inferir acerca do desempenho do ChatGPT em problemas de dimensionamento de lotes, recorreu-se ao ANOVA. O ANOVA (Analysis of Variance) é um método estatístico utilizado para comparar as médias de três ou mais grupos, a fim de determinar se existem diferenças estatisticamente significativas entre eles. O ANOVA examina a variação entre as médias dos grupos e a variação dentro de cada grupo para avaliar se as diferenças observadas são devidas a variabilidade da amostra ou a fatores específicos. Esse método é amplamente utilizado em pesquisas experimentais onde múltiplos grupos ou condições são analisados simultaneamente. Quando a ANOVA indica a existência de diferenças significativas entre os grupos, é necessário realizar análises adicionais para identificar quais grupos específicos diferem entre si. Para isso, utilizam-se testes *Post Hoc*, que permitem comparações par a par entre os grupos, ajustando os níveis de significância para evitar erros tipo I (falsos positivos). Entre os testes *Post Hoc* mais

comuns estão o teste de Tukey, o teste de Bonferroni e o teste de Scheffé, que será o utilizado neste trabalho (Rochon et al., 2012).

O ANOVA, que é um teste de hipóteses paramétrico, pode ser aplicado aos resultados do ChatGPT e outros métodos, porque, pelo Teorema do Limite Central, pode assumir-se a normalidade das amostras (Rochon et al., 2012). Sendo  $\mu_{ChatGPT40}$ ,  $\mu_{WW}$ ,  $\mu_{SM}$ ,  $\mu_{PPB}$ ,  $\mu_{QEP}$  as médias do desvio relativo da solução do ChatGPT 4o, Wagner Whitin, Silver Meal, Parts Period Balancing e Quantidade Económica de Produção.

As hipóteses analisadas pela ANOVA são:

- $H_0: \mu_{ChatGPT40} = \mu_{WW} = \mu_{SM} = \mu_{PPB} = \mu_{QEP}$
- $H_1: \text{Há uma diferença na média do desvio relativo do ótimo da amostra}$

Ou seja, o ANOVA vai analisar a existência de evidências que permitam afirmar que as populações não têm um desempenho médio idêntico, isto é, se pelo menos uma população não tem um desempenho médio idêntico às outras. Os resultados do ANOVA são apresentados na Tabela 35 e Tabela 36 para a Amostra 1 e Tabela 37 e Tabela 38 para a Amostra 2:

#### Amostra I

O resultado da ANOVA é apresentado na Tabela 35 e como é possível analisar foram encontradas evidências estatísticas para rejeitar a hipótese nula, com 95% de confiança. Por outras palavras, não é possível dizer que ChatGPT-4o e as outras técnicas tenham um desempenho médio idêntico, pelo menos uma população não apresentou um desempenho médio idêntico às demais.

Tabela 35 - ANOVA da Média do Desvio Relativo da Amostra 1

ANOVA					
Todos					
	Soma dos Quadrados	df	Quadrado Médio	F	Sig.
Entre Grupos	24298,145	5	4859,629	720,252	<,001
Nos grupos	1174,000	174	6,747		
Total	25472,145	179			

Como foi possível observar, pelo menos uma das populações apresentou um desempenho distinto das demais. Para identificar quais populações tiveram desempenhos estatisticamente diferentes, é necessário realizar um teste *Post Hoc*. Neste caso, foi utilizado o Teste de Scheffé (Rochon et al., 2012). O Teste de Scheffé, conforme mostrado na Tabela 36, permite identificar quais pares de populações possuem médias de desvio relativo em relação à solução ótima que possam ser consideradas semelhantes.

Tabela 36 - Teste de Scheffe da Medias do Desvio Relativo Amostra 1

<b>Comparações múltiplas</b>						
Variável dependente: Amostra 1						
Scheffe						
(I) Grupo	(J) Grupo	Diferença média (I-J)	Erro Padrão	Sig.	Intervalo de Confiança 95%	
					Limite inferior	Limite superior
GPT-4o	GPT-4	-1,06030%	0,67068%	,776	-3,3178%	1,1972%
	WW	0,60488%	0,67068%	,976	-1,6527%	2,8624%
	SM	-0,23454%	0,67068%	1,000	-2,4921%	2,0230%
	PPB	-2,39865%*	0,67068%	,029	-4,6562%	-0,1411%
	QEP	-31,68952%*	0,67068%	<,001	-33,9471%	-29,4320%
GPT-4	GPT-4o	1,06030%	0,67068%	,776	-1,1972%	3,3178%
	WW	1,66518%	0,67068%	,296	-0,5924%	3,9227%
	SM	0,82576%	0,67068%	,910	-1,4318%	3,0833%
	PPB	-1,33836%	0,67068%	,554	-3,5959%	0,9192%
	QEP	-30,62923%*	0,67068%	<,001	-32,8868%	-28,3717%
WW	GPT-4o	-0,60488%	0,67068%	,976	-2,8624%	1,6527%
	GPT-4	-1,66518%	0,67068%	,296	-3,9227%	0,5924%
	SM	-0,83942%	0,67068%	,905	-3,0970%	1,4181%
	PPB	-3,00353%*	0,67068%	,002	-5,2611%	-0,7460%
	QEP	-32,29441%*	0,67068%	<,001	-34,5519%	-30,0369%
SM	GPT-4o	0,23454%	0,67068%	1,000	-2,0230%	2,4921%
	GPT-4	-0,82576%	0,67068%	,910	-3,0833%	1,4318%
	WW	0,83942%	0,67068%	,905	-1,4181%	3,0970%
	PPB	-2,16411%	0,67068%	,070	-4,4217%	0,0934%
	QEP	-31,45498%*	0,67068%	<,001	-33,7125%	-29,1974%
PPB	GPT-4o	2,39865%*	0,67068%	,029	0,1411%	4,6562%
	GPT-4	1,33836%	0,67068%	,554	-0,9192%	3,5959%
	WW	3,00353%*	0,67068%	,002	0,7460%	5,2611%
	SM	2,16411%	0,67068%	,070	-0,0934%	4,4217%
	QEP	-29,29087%*	0,67068%	<,001	-31,5484%	-27,0333%
QEP	GPT-4o	31,68952%*	0,67068%	<,001	29,4320%	33,9471%
	GPT-4	30,62923%*	0,67068%	<,001	28,3717%	32,8868%
	WW	32,29441%*	0,67068%	<,001	30,0369%	34,5519%
	SM	31,45498%*	0,67068%	<,001	29,1974%	33,7125%
	PPB	29,29087%*	0,67068%	<,001	27,0333%	31,5484%

\*. A diferença média é significativa no nível 0.05.

O Teste de Scheffe revela diferenças estatisticamente significativas no desempenho entre o ChatGPT-4o e a técnica QEP, entre o ChatGPT-4o e o PPB, além de diferenças entre o QEP e as demais técnicas. O teste também confirma a existência de diferenças estatisticamente significativas entre o desempenho do ChatGPT-4 e o QEP, assim como entre o ChatGPT-4 e o PPB. Sendo assim, com 95% de confiança, é possível assumir que para a Amostra 1, os resultados indicam que ambos os modelos, ChatGPT-4o e ChatGPT-4, não há diferença significativa no desempenho em termos do seu desvio do valor ótimo, apresentando *p-value* de 0,976 e de 0,296, respectivamente.

### Amostra 2

Para a segunda amostra, o resultado da ANOVA é apresentado na Tabela 37 e como é possível analisar foram novamente encontradas evidências estatísticas para rejeitar a hipótese nula, com 95% de confiança. Não é possível dizer que ChatGPT-4o e as outras técnicas tenham um desempenho médio idêntico.

Tabela 37 - ANOVA da Média do Desvio Relativo da Amostra 2

<b>ANOVA</b>					
Todos					
	Soma dos Quadrados	df	Quadrado Médio	F	Sig.
Entre Grupos	27289,834	4	6822,459	302,618	<,001
Nos grupos	3268,992	145	22,545		
Total	30558,827	149			

O Teste de Scheffé, conforme mostrado na Tabela 38, permite identificar quais pares de populações possuem médias de desvio relativo em relação à solução ótima que possam ser consideradas semelhantes.

Tabela 38 - Teste de Scheffe da Medias do Desvio Relativo Amostra 2

<b>Comparações múltiplas</b>						
Variável dependente: Amostra 2						
Scheffe						
(I) Grupo	(J) Grupo	Diferença média (I-J)	Erro Padrão	Sig.	Intervalo de Confiança 95%	
					Limite inferior	Limite superior
GPT-4o	WW	0,81206%	1,22596%	,979	-3,0133%	4,6374%
	SM	-1,07530%	1,22596%	,942	-4,9007%	2,7501%
	PPB	-2,19326%	1,22596%	,527	-6,0186%	1,6321%
	QEP	-34,23971%*	1,22596%	<,001	-38,0651%	-30,4143%
WW	GPT-4o	-0,81206%	1,22596%	,979	-4,6374%	3,0133%
	SM	-1,88736%	1,22596%	,669	-5,7127%	1,9380%
	PPB	-3,00532%	1,22596%	,205	-6,8307%	0,8201%

Comparações múltiplas						
Variável dependente: Amostra 2						
Scheffe						
(I) Grupo	(J) Grupo	Diferença média (I-J)	Erro Padrão	Sig.	Intervalo de Confiança 95%	
					Limite inferior	Limite superior
	QEP	-35,05176%*	1,22596%	<,001	-38,8771%	-31,2264%
SM	GPT-4o	1,07530%	1,22596%	,942	-2,7501%	4,9007%
	WW	1,88736%	1,22596%	,669	-1,9380%	5,7127%
	PPB	-1,11796%	1,22596%	,934	-4,9433%	2,7074%
	QEP	-33,16440%*	1,22596%	<,001	-36,9898%	-29,3390%
PPB	GPT-4o	2,19326%	1,22596%	,527	-1,6321%	6,0186%
	WW	3,00532%	1,22596%	,205	-0,8201%	6,8307%
	SM	1,11796%	1,22596%	,934	-2,7074%	4,9433%
	QEP	-32,04644%*	1,22596%	<,001	-35,8718%	-28,2211%
QEP	GPT-4o	34,23971%*	1,22596%	<,001	30,4143%	38,0651%
	WW	35,05176%*	1,22596%	<,001	31,2264%	38,8771%
	SM	33,16440%*	1,22596%	<,001	29,3390%	36,9898%
	PPB	32,04644%*	1,22596%	<,001	28,2211%	35,8718%

\*. A diferença média é significativa no nível 0.05.

O Teste de Scheffe revela diferenças estatisticamente significativas no desempenho entre o ChatGPT-4o e a técnica QEP, bem como entre o QEP e as demais técnicas avaliadas. Sendo assim, com 95% de confiança, é possível assumir que para a Amostra 2, não há diferença significativa no desempenho do ChatGPT-4o em termos de desvio do valor ótimo, apresentando *p-value* de 0,979 e quando comparada ao método WW.

Verifica-se assim que para a população analisada, e com 95% de confiança, o ChatGPT 4o obteve um desempenho médio idêntico ao método de Wagner-Whitin (WW), apresentando menores desvios médios e maior *p-value* quando comparado às restantes heurísticas.

#### 4.4. Conclusão final dos Resultados

O problema do dimensionamento de lotes tem sido amplamente estudado ao longo das últimas décadas, com o desenvolvimento de diversas heurísticas e métodos exatos para sua resolução. Por outro lado, nos anos recentes, a inteligência artificial, particularmente com o surgimento da plataforma ChatGPT, tem mostrado grande potencial em diversas áreas de aplicação. Para avaliar a viabilidade do uso do ChatGPT como uma ferramenta na resolução do problema de dimensionamento de lotes, foram conduzidos testes com duas amostras distintas, cada uma com características próprias, para examinar o desempenho do modelo em diferentes cenários. Na Amostra I, que avaliou o desempenho do ChatGPT em condições de baixa variabilidade nas quantidades das ordens de fabrico e custos de posse reduzidos, o modelo ChatGPT-4o evidenciou menor variabilidade e maior precisão nos resultados em comparação ao modelo

anterior, ChatGPT-4, configurando-se como a alternativa mais confiável. Adicionalmente, o ChatGPT-4o apresentou resultados que se aproximam significativamente das soluções ótimas, superando heurísticas tradicionais como o Silver Meal (SM) e o Parts Period Balancing (PPB). O SM obteve desempenho intermediário, enquanto o PPB e o Quantidade Económica de Pedido (QEP) mostraram maior variabilidade e menor precisão, sendo o QEP a técnica com o pior desempenho.

Na Amostra II, o ChatGPT-4o manteve-se como a técnica com o melhor desempenho, apresentando elevada exatidão e consistência, com resultados muito próximos dos valores ótimos estabelecidos pelo método Wagner-Whitin (WW). O desempenho do SM, embora razoável, foi ligeiramente inferior ao observado na Amostra I, enquanto o PPB e o QEP demonstraram desvios mais acentuados e maior variabilidade, com o QEP permanecendo como a técnica menos precisa.

O desempenho do ChatGPT-4 foi semelhante ao do SM, embora com maior variabilidade, apresentando um desvio padrão de 2,46% comparados aos 1,24% do SM. No entanto, os desempenhos iniciais do GPT-4 mostraram-se superiores quando comparados com as restantes técnicas, exceto em relação ao ChatGPT-4o. Este modelo não apenas superou o seu antecessor, o GPT-4, mas também foi mais eficaz do que o SM, exibindo uma proximidade mais consistente aos resultados ótimos. Além disso, para a população analisada, com 95% de confiança, verificou-se que o ChatGPT-4o teve um desempenho médio idêntico ao método de Wagner-Whitin (WW), apresentando menores desvios médios e um *p-value* maior quando comparado às restantes heurísticas.

## 5. Conclusão

### 5.1. Conclusões finais

O dimensionamento de lotes de produção é um processo de decisão crucial que ocorre em qualquer empresa de produção, por vezes, de forma empírica. Este problema está intimamente ligado ao planeamento da produção e é essencial para a criação de um plano eficiente, que minimize os custos de produção, maximize a utilização dos recursos, reduza os níveis de *stock* e, ao mesmo tempo, atenda à procura dos clientes de maneira eficaz.

Embora o problema de dimensionamento de lotes possa parecer simples, envolve uma série de decisões complexas, como determinar a quantidade de produtos a serem fabricados em cada lote. Essa decisão pode impactar significativamente os custos operacionais, o tempo de produção e o nível de stock. Uma simples alteração no tamanho do lote pode levar a variações substanciais nos tempos de ciclo, custos de *setup* e necessidades de armazenamento. Esse processo torna-se ainda mais relevante quando se considera que o dimensionamento de lotes de produção se estende para além do contexto industrial tradicional, como em operações logísticas e de serviços.

Muitas vezes, os métodos exatos, que procuram a solução ótima por meio da enumeração completa das soluções, são inviáveis para problemas de grande escala devido ao aumento exponencial do número de soluções, com o crescimento do tamanho do problema. Essa complexidade faz com que a resolução exata de problemas de dimensionamento de lotes em ambientes reais seja difícil e necessite alto poder de processamento.

Como alternativa, problemas de dimensionamento de lotes podem ser abordados por meio de técnicas aproximadas. Essas técnicas, como heurísticas, oferecem soluções eficientes sem a necessidade de explorar todas as soluções.

Neste contexto, a introdução de métodos de IA, como Algoritmos Genéticos, *Machine Learning* e *Deep Learning*, tem mostrado grande potencial na resolução de problemas complexos de dimensionamento de lotes. O ChatGPT, lançado em 2022, oferece capacidades inovadoras de processamento de linguagem natural, embora sua aplicabilidade direta ao dimensionamento de lotes ainda não tenha sido comprovada.

Embora o ChatGPT se tenha mostrado promissor e com bons resultados nas amostras utilizadas, é importante reconhecer que esta ainda é uma tecnologia recente e que seu desempenho em amostras com diferentes características permanece incerto. A variabilidade nos resultados sugere que, embora eficiente em contextos controlados, o comportamento do ChatGPT pode não ser tão consistente quando exposto a cenários com maior diversidade ou complexos. Além disso, o método empregado para fornecer dados à plataforma foi relativamente arcaico, já que todas as informações foram apresentadas em formato de texto, limitando o potencial total da tecnologia. No futuro, o uso direto do API do ChatGPT pode otimizar este processo, permitindo uma integração mais robusta e eficiente. Com acesso direto às capacidades do ChatGPT em uma aplicação, seria possível explorar plenamente as suas funcionalidades, proporcionando uma interação mais dinâmica e adaptativa na resolução de problemas de dimensionamento de lotes de produção.

O ChatGPT foi aplicado ao problema de dimensionamento de lotes na sua forma mais simples, abordando o problema para um único artigo e sem considerar restrições de capacidade de produção. Esse cenário simplificado permitiu explorar as capacidades iniciais do ChatGPT na otimização de lotes, mas ainda permanece incerto como a tecnologia se comportaria ao enfrentar variações mais complexas do problema, como aquelas que envolvem múltiplos artigos e restrições de capacidade. Implementar o uso do ChatGPT em contextos mais complexos exigiria uma adaptação da abordagem, e a eficácia de tal implementação ainda precisa ser validada. A transição para esses cenários mais desafiadores iria necessitar de uma validação das técnicas utilizadas, bem como uma maior integração com dados específicos e estruturados que reflitam a realidade operacional das empresas, para garantir que o ChatGPT pudesse oferecer soluções eficazes e otimizadas em situações de maior complexidade.

Esta implementação do ChatGPT no problema mais simples de dimensionamento de lotes, se refinada e tornada mais prática de utilizar, tem o potencial de se tornar uma ferramenta valiosa para pequenas empresas e negócios. Muitas dessas empresas procuram controlar melhor os seus *stocks*, mas frequentemente carecem de conhecimento nas técnicas de dimensionamento de lotes ou simplesmente não dispõem de tempo para se dedicar a essas análises. Ao simplificar a interface e automatizar o processo com o uso do ChatGPT, essas empresas poderiam facilmente acessar *insights* valiosos sobre a otimização de seus *stocks*, permitindo decisões mais informadas e eficientes sem a necessidade de análise técnica aprofundada. Assim, o ChatGPT poderia atuar como um assistente acessível e poderoso para melhorar a gestão, atendendo às necessidades específicas de pequenos negócios e contribuindo para sua competitividade e eficiência operacional.

Os resultados obtidos a partir das análises das Amostras 1 e 2 demonstram que o ChatGPT, particularmente na versão ChatGPT-4o, possui um potencial significativo na resolução do problema de dimensionamento de lotes, superando métodos tradicionais como o Silver Meal, Parts Period Balancing e Quantidade Económica de Produção. Na Amostra 1, o ChatGPT-4o apresentou menor variabilidade e maior precisão nos resultados em comparação com o método SM, amplamente reconhecido como uma das melhores heurísticas em termos de qualidade de solução. Este modelo demonstrou uma aproximação substancial às soluções ótimas, destacando-se não apenas em relação ao seu antecessor, mas também frente às

heurísticas convencionais. De forma similar, na Amostra 2, o ChatGPT-4o manteve um desempenho consistente e elevado, confirmando-se como a técnica que mais se aproxima dos resultados ótimos. No estudo de inferência estatística, foi verificado que, em ambas as amostras, o ChatGPT-4o não apresentou diferenças significativas em relação aos resultados ótimos, com valores de *p-value* de 0,976 e 0,979, respetivamente. Os desvios do ChatGPT-4o foram inferiores a 1% em ambas as amostras, com 0,6% na amostra 1 e 0,81% na amostra 2, em comparação com o desvio do SM, que foi de 0,84% na Amostra 1 e 1,89% na Amostra 2.

## 5.2. Trabalhos futuros

Como extensão deste trabalho, será interessante avaliar o desempenho do ChatGPT em outras amostras com características diferentes. No entanto, para tal, seria fundamental encontrar uma forma de automatizar o processo de *input* de dados para o cálculo do dimensionamento de lotes, já que alterar manualmente os valores das quantidades para cada instância torna-se moroso, especialmente quando se avaliam grandes quantidades de dados. A automatização deste processo, poderia facilitar significativamente a análise.

Outra direção futura promissora para a continuidade deste trabalho é a utilização direta do API (Interface de Programação de Aplicações) do ChatGPT, com o objetivo de tornar a sua aplicação na resolução do problema de dimensionamento de lotes mais prática e acessível. O API é uma forma de interação direta com o ChatGPT, possibilitando a integração das suas funcionalidades em sistemas externos, aplicações ou plataformas personalizadas. Ao invés de depender de interações manuais e ajustes repetitivos de *prompts* na interface tradicional, o uso do API permite a automatização dos processos, tornando a utilização mais eficiente e ajustada às necessidades específicas dos utilizadores. Com o API, é possível desenvolver interfaces dedicadas que guiem o utilizador através do processo, minimizando a necessidade de intervenção manual e reduzindo a margem de erro, o que aumenta a eficiência e a precisão na resolução de problemas de dimensionamento de lotes.



## Referências

- Aggarwal, A., & Park, J. K. (1993). Improved Algorithms for Economic Lot Size Problems. *Operations Research*, 41(3), 549–571. <https://doi.org/10.1287/opre.41.3.549>
- Alkaissi, H., & McFarlane, S. I. (2023). Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. *Cureus*, 15(2), e35179. <https://doi.org/10.7759/cureus.35179>
- Alp, O., Erkip, N. K., & Güllü, R. (2003). Optimal Lot-Sizing/Vehicle-Dispatching Policies Under Stochastic Lead Times and Stepwise Fixed Costs. *Operations Research*, 51(1), 160–166. <https://doi.org/10.1287/opre.51.1.160.12799>
- Alves, F. F., Nogueira, T. H., de Souza, M. C., & Ravetti, M. G. (2021). Approaches for the joint resolution of lot-sizing and scheduling with infeasibilities occurrences. *Computers & Industrial Engineering*, 155, 107176. <https://doi.org/10.1016/j.cie.2021.107176>
- Arora, M., Dhawan, S., & Singh, K. (2018). Deep Learning: Overview, Architecture, Framework & Applications. *International Journal of Latest Trends in Engineering and Technology*, 10(1), 379–384. <https://doi.org/http://dx.doi.org/10.21172/1.101.69>
- Baciarello, L., D’Avino, M., Onori, R., & Schiraldi, M. M. (2013). Lot Sizing Heuristics Performance. *International Journal of Engineering Business Management*, 5, 5. <https://doi.org/10.5772/56004>
- Bahl, H. C., Ritzman, L. P., & Gupta, J. N. D. (1987). OR Practice—Determining Lot Sizes and Resource Requirements: A Review. *Operations Research*, 35(3), 329–345. <https://doi.org/10.1287/opre.35.3.329>
- Bansal, J. (2023). *89+ Amazing ChatGPT Stats For 2023 (GPT-4 Update)*. DISCOVERTHETECH. <https://discoverthetech.com/chatgpt-stats/>
- Bhoskar, M. T., Kulkarni, M. O. K., Kulkarni, M. N. K., Patekar, M. S. L., Kakandikar, G. M., & Nandedkar, V. M. (2015). Genetic Algorithm and its Applications to Mechanical Engineering: A Review. *Materials Today: Proceedings*, 2(4–5), 2624–2630. <https://doi.org/10.1016/j.matpr.2015.07.219>
- Blackburn, J. D., & Millen, R. A. (1980). HEURISTIC LOT-SIZING PERFORMANCE IN A ROLLING-SCHEDULE ENVIRONMENT\*. *Decision Sciences*, 11(4), 691–701. <https://doi.org/10.1111/j.1540-5915.1980.tb01170.x>
- Bockting, C. L., van Dis, E. A. M., Bollen, J., van Rooij, R., & Zuidema, W. (2023). ChatGPT: five priorities for research. *NATURE*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7> WE - Science Citation Index Expanded (SCI-EXPANDED)
- Bommarito, M., & Katz, D. M. (2022). GPT Takes the Bar Exam. *SSRN Electronic Journal*, November 2022, 1–15. University of North Dakota School of Law
- Bookbinder, J. H., & H’ng, B.-T. (1986). Production lot sizing for deterministic rolling schedules.

- Journal of Operations Management*, 6(3–4), 349–362. [https://doi.org/10.1016/0272-6963\(86\)90009-4](https://doi.org/10.1016/0272-6963(86)90009-4)
- Brahimi, N., Absi, N., Dauzère-Pérès, S., & Nordli, A. (2017). Single-item dynamic lot-sizing problems: An updated survey. *European Journal of Operational Research*, 263(3), 838–863. <https://doi.org/10.1016/j.ejor.2017.05.008>
- Brahimi, N., Dauzere-Peres, S., Najid, N. M., & Nordli, A. (2006). Single item lot sizing problems. *European Journal of Operational Research*, 168(1), 1–16. <https://doi.org/10.1016/j.ejor.2004.01.054>
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT. *Journal of the ACM*, 37(4). <http://arxiv.org/abs/2303.04226>
- Chan-Olmsted, S. M. (2019). A Review of Artificial Intelligence Adoptions in the Media Industry. *International Journal on Media Management*, 21(3–4), 193–215. <https://doi.org/10.1080/14241277.2019.1695619>
- Charlie Giattino, Edouard Mathieu, V. S. and M. R. (2023). *Artificial Intelligence. OurWorldInData.Org*. <https://ourworldindata.org/artificial-intelligence>
- Chatterjee, J., & Dethlefs, N. (2023). This new conversational AI model can be your friend, philosopher, and guide ... and even your worst enemy. *Patterns*, 4(1), 100676. <https://doi.org/10.1016/j.patter.2022.100676>
- Copil, K., Wörbelauer, M., Meyr, H., & Tempelmeier, H. (2017). Simultaneous lotsizing and scheduling problems: a classification and review of models. *OR Spectrum*, 39(1), 1–64. <https://doi.org/10.1007/s00291-015-0429-4>
- Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of Artificial Intelligence in Machine Learning: Review and Prospect. *International Journal of Computer Applications*, 115(9), 31–41. <https://doi.org/10.5120/20182-2402>
- Drexl, A., & Kimms, A. (1997). Lot sizing and scheduling — Survey and extensions. *European Journal of Operational Research*, 99(2), 221–235. [https://doi.org/10.1016/S0377-2217\(97\)00030-1](https://doi.org/10.1016/S0377-2217(97)00030-1)
- Edward A. Silver, David F. Pyke, & Rein Peterson. (1998). *Inventory Management and Production Planning and Scheduling* (3rd ed.). Wiley.
- Erlenkotter, D. (1990). Ford Whitman Harris and the Economic Order Quantity Model. *Operations Research*, 38(6), 937–946. <https://doi.org/10.1287/opre.38.6.937>
- Evans, J. R. (1985). An efficient implementation of the Wagner-Whitin algorithm for dynamic lot-sizing. *Journal of Operations Management*, 5(2), 229–235. [https://doi.org/10.1016/0272-6963\(85\)90009-9](https://doi.org/10.1016/0272-6963(85)90009-9)
- Florim, W., Dias, P., Santos, A. S., Varela, L. R., Madureira, A. M., & Putnik, G. D. (2019). Analysis of lot-sizing methods' suitability for different manufacturing application scenarios oriented to MRP and JIT/Kanban environments. *Brazilian Journal of Operations & Production Management*, 16(4), 638–649. <https://doi.org/10.14488/BJOPM.2019.v16.n4.a9>
- Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277–304.

- <https://doi.org/10.1080/15228053.2023.2233814>
- Gill, S. S., & Kaur, R. (2023). ChatGPT: Vision and challenges. *Internet of Things and Cyber-Physical Systems*, 3, 262–271.  
<https://doi.org/https://doi.org/10.1016/j.iotcps.2023.05.004>
- Gordijn, B., & Have, H. ten. (2023). ChatGPT: evolution or revolution? *Medicine, Health Care and Philosophy*, 26(1), 1–2. <https://doi.org/10.1007/s11019-023-10136-0>
- Guimarães, L., Klabjan, D., & Almada-Lobo, B. (2014). Modeling lotsizing and scheduling problems with sequence dependent setups. *European Journal of Operational Research*, 239(3), 644–662. <https://doi.org/10.1016/j.ejor.2014.05.018>
- Guner Goren, H., Tunali, S., & Jans, R. (2010). A review of applications of genetic algorithms in lot sizing. *Journal of Intelligent Manufacturing*, 21(4), 575–590.  
<https://doi.org/10.1007/s10845-008-0205-2>
- Harris, F. W. (1990). How Many Parts to Make at Once. *Operations Research*, 38(6), 947–950.  
<https://doi.org/10.1287/opre.38.6.947>
- Hoyer, W. D., Kroschke, M., Schmitt, B., Kraume, K., & Shankar, V. (2020). Transforming the Customer Experience Through New Technologies. *Journal of Interactive Marketing*, 51(1), 57–71. <https://doi.org/10.1016/j.intmar.2020.04.001>
- Jain, V., Rai, H., . P., & Mogaji, E. (2023). The Prospects and Challenges of ChatGPT on Marketing Research and Practices. *SSRN Electronic Journal*.  
<https://doi.org/10.2139/ssrn.4398033>
- Jans, R., & Degraeve, Z. (2007). Modeling Industrial Lot Sizing Problems: A Review. *International Journal of Production Research*, 46, 1619–1643.  
<https://doi.org/10.1080/00207540600902262>
- Jens, & Meyr, H. (2023). *A machine learning approach for predicting the best solution heuristic for a large scaled Capacitated Lotsizing Problem*. <https://doi.org/10.21203/rs.3.rs-3709286/v1>
- José Fernando Gonçalves. (2010). *Gestão de Aprovisionamentos* (Publindústria (ed.); 2ª).
- K Kar, U. (2018). The Future of Health and Healthcare in a World of Artificial Intelligence. *Archives in Biomedical Engineering & Biotechnology*, 1(1).  
<https://doi.org/10.33552/ABEB.2018.01.000503>
- Kar, U., Dash, R., McMurtrey, M., & Rebman, C. (2019). Application of Artificial Intelligence in Automation of Supply Chain Management. *Journal of Strategic Innovation and Sustainability*, 14(3). <https://doi.org/10.33423/jsis.v14i3.2105>
- Karimi, B., Fatemi Ghomi, S. M. T., & Wilson, J. M. (2003). The capacitated lot sizing problem: a review of models and algorithms. *Omega*, 31(5), 365–378.  
[https://doi.org/10.1016/S0305-0483\(03\)00059-8](https://doi.org/10.1016/S0305-0483(03)00059-8)
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.  
<https://doi.org/10.1038/nature14539>
- Lee, I., Gupta, J. N. D., & Amar, A. D. (2001). A multi-neural-network learning for lot sizing and sequencing on a flow-shop. *Proceedings of the 2001 ACM Symposium on Applied Computing*, 36–40. <https://doi.org/10.1145/372202.372210>
- Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., Cao, K., Liu, D.,

- Wang, G., Xu, Q., Fang, X., Zhang, S., Xia, J., & Xia, J. (2020). Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. *Radiology*, *296*(2), E65–E71. <https://doi.org/10.1148/radiol.2020200905>
- Lin, J. C., Younessi, D. N., Kurapati, S. S., Tang, O. Y., & Scott, I. U. (2023). Comparison of GPT-3.5, GPT-4, and human user performance on a practice ophthalmology written examination. *Eye*. <https://doi.org/10.1038/s41433-023-02564-2>
- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023). Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology*, *1*(2), 100017. <https://doi.org/10.1016/j.metrad.2023.100017>
- Ludermir, T. B. (2021). Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. In *Estudos Avançados* (Vol. 35). scielo .
- Maravelias, C. T., & Sung, C. (2009). Integration of production planning and scheduling: Overview, challenges and opportunities. *Computers & Chemical Engineering*, *33*(12), 1919–1930. <https://doi.org/10.1016/j.compchemeng.2009.06.007>
- Mehdi, Y. (2023). *Confirmed: the new Bing runs on OpenAI's GPT-4*. Microsoft Bing Blogs. [https://blogs.bing.com/search/march\\_2023/Confirmed-the-new-Bing-runs-on-OpenAI's-GPT-4](https://blogs.bing.com/search/march_2023/Confirmed-the-new-Bing-runs-on-OpenAI's-GPT-4)
- Miclo, R., Fontanili, F., Lauras, M., Lamothe, J., & Milian, B. (2015). MRP vs. demand-driven MRP: Towards an objective comparison. *2015 International Conference on Industrial Engineering and Systems Management (IESM)*, 1072–1080. <https://doi.org/10.1109/IESM.2015.7380288>
- Miclo, R., Lauras, M., Fontanili, F., Lamothe, J., & Melnyk, S. A. (2019). Demand Driven MRP: assessment of a new approach to materials management. *International Journal of Production Research*, *57*(1), 166–181. <https://doi.org/10.1080/00207543.2018.1464230>
- MidJourney. (2024). *MidJourney*. <https://www.midjourney.com/home>
- Open AI. (2024). *Prompt Engineering*. <https://doi.org/10.53349/schuleverantworten.2024.i1.a405>
- OpenAI. (2022). *ChatGPT: Optimizing Language Models for Dialogue*. <https://openai.com/blog/chatgpt/>
- OpenAI. (2023). *GPT-4 Technical Report*. 4, 1–100. <http://arxiv.org/abs/2303.08774>
- Pisinger, D., & Ropke, S. (2007). A general heuristic for vehicle routing problems. *Computers & Operations Research*, *34*(8), 2403–2435. <https://doi.org/10.1016/j.cor.2005.09.012>
- Pochet, Y., & Wolsey, L. A. (2006). Production Planning by Mixed Integer Programming. In *Springer Series in Operations Research and Financial Engineering*. Springer New York. <https://doi.org/10.1007/0-387-33477-7>
- Reçus, L. (1964). F. Hanssmann Operations Research in Production and Inventory Control. New York, London, John Wiley and Sons, Inc., 1962, XII p. 254 p., 64/- . *Recherches Économiques de Louvain*, *30*(2), 150–150. <https://doi.org/10.1017/S0770451800015761>
- Rochon, J., Gondan, M., & Kieser, M. (2012). To test or not to test: Preliminary assessment of normality when comparing two independent samples. *BMC Medical Research Methodology*, *12*(1), 81. <https://doi.org/10.1186/1471-2288-12-81>

- Rogers, J. (1958). A Computational Approach to the Economic Lot Scheduling Problem. *Management Science*, 4(3), 264–291. <https://doi.org/10.1287/mnsc.4.3.264>
- Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). *A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications*. <http://arxiv.org/abs/2402.07927>
- Sallam, M. (2023). ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *HEALTHCARE*, 11(6). <https://doi.org/10.3390/healthcare11060887> WE - Science Citation Index Expanded (SCI-EXPANDED) WE - Social Science Citation Index (SSCI)
- Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, 2(6), 420. <https://doi.org/10.1007/s42979-021-00815-1>
- Şenyiğit, E., & Atici, U. (2013). Artificial neural network models for lot-sizing problem: A case study. *Neural Computing and Applications*, 22(6), 1039–1047. <https://doi.org/10.1007/s00521-012-0863-z>
- Sheikh, H., Prins, C., & Schrijvers, E. (2023). Mission AI: The New System Technology. In *Mission AI*.
- Simpson, N. C. (2001). Questioning the relative virtues of dynamic lot sizing rules. *Computers & Operations Research*, 28(9), 899–914. [https://doi.org/10.1016/S0305-0548\(00\)00015-0](https://doi.org/10.1016/S0305-0548(00)00015-0)
- The Lancet Digital Health. (2023). ChatGPT: friend or foe? *The Lancet Digital Health*, 5(3), e102. [https://doi.org/10.1016/S2589-7500\(23\)00023-7](https://doi.org/10.1016/S2589-7500(23)00023-7)
- Thorp, H. H. (2023). ChatGPT is fun, but not an author. *SCIENCE*, 379(6630), 313. <https://doi.org/10.1126/science.adg7879> WE - Science Citation Index Expanded (SCI-EXPANDED)
- TURING, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, LIX(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Van Vyve, M. (2007). Algorithms for Single-Item Lot-Sizing Problems with Constant Batch Size. *Mathematics of Operations Research*, 32(3), 594–613. <https://doi.org/10.1287/moor.1070.0257>
- Wagelmans, A., van Hoesel, S., & Kolen, A. (1992). Economic Lot Sizing: An  $O(n \log n)$  Algorithm That Runs in Linear Time in the Wagner-Whitin Case. *Operations Research*, 40(1-supplement-1), S145–S156. <https://doi.org/10.1287/opre.40.1.S145>
- Wagner, H. M., & Whitin, T. M. (1958). Dynamic Version of the Economic Lot Size Model. *Management Science*, 5(1), 89–96. <https://doi.org/10.1287/mnsc.5.1.89>
- Wang, S., Scells, H., Koopman, B., & Zuccon, G. (2023). Can ChatGPT Write a Good Boolean Query for Systematic Review Literature Search? *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1(1), 1426–1436. <https://doi.org/10.1145/3539618.3591703>
- World Economic Forum. (2023). Future of Jobs Report. In *World Economic Forum* (Vol. 59, Issue May). <https://www.weforum.org/reports/the-future-of-jobs-report-2023/>



## Declaração de Integridade

Declaro ter conduzido este trabalho académico com integridade. Não plagiei ou apliquei qualquer forma de uso indevido de informações ou falsificação de resultados ao longo do processo que levou à sua elaboração.

Declaro que o trabalho apresentado neste documento é original e de minha autoria, não tendo sido utilizado anteriormente para nenhum outro fim.

Declaro ainda que tenho pleno conhecimento do Código de Conduta Ética do P.PORTO.

NOME: André Da Silva Ribeiro

ISEP, Porto, 11 de Setembro de 2024



## Anexos A

Anexo A (Código criado pelo ChatGPT 4.0)

```
def find_optimal_ordering(setup_cost, holding_cost, demands):
    n = len(demands)
    min_cost = [float('inf')] * (n + 1)
    order_strategy = [-1] * n

    min_cost[n] = 0 # Boundary condition

    for t in range(n - 1, -1, -1):
        cumulative_demand = 0
        cumulative_holding_cost = 0

        for i in range(t, n):
            cumulative_demand += demands[i]
            if i > t:
                cumulative_holding_cost += (i - t) * demands[i] *
holding_cost

            order_cost = setup_cost + cumulative_holding_cost
            total_cost = order_cost + min_cost[i + 1]

            if total_cost < min_cost[t]:
                min_cost[t] = total_cost
                order_strategy[t] = (cumulative_demand, i + 1)

    actual_orders = []
    t = 0
    while t < n:
        if order_strategy[t][0] > 0:
            order_qty, next_period = order_strategy[t]
            actual_orders.append((t + 1, order_qty))
            t = next_period
        else:
            t += 1

    return min_cost[0], actual_orders
```

```
def main():
    # Example usage:
    setup_cost = 300
    holding_cost = 2
    demands = [10, 10, 15, 20, 70, 180, 250, 280, 230, 40, 0, 10]

    total_cost, orders = find_optimal_ordering(setup_cost, holding_cost,
demands)
    print("Total Cost:", total_cost)
    print("Orders:", orders)

if __name__ == "__main__":
    main()
```