



Inteligência Artificial para Otimização de Rotas e Recursos em Apoio ao Domicílio

ANDRÉ FILIPE SERRA TEIXEIRA

Junho de 2026

Inteligência Artificial para Otimização de Rotas e Recursos em Apoio ao Domicílio

André Filipe Serra Teixeira

**Dissertação para obtenção do Grau de Mestre em
Engenharia Informática, Área de Especialização em
Cibersegurança e Administração de Sistemas**

Orientador: Fátima Rodrigues

Co-orientador: Hélder Pinto

Porto, junho 2026

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.

Portanto, 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.

ISEP, Porto, 29 de junho de 2026

Dedicatória

Dedico este trabalho à minha família, mas sobretudo ao meu pai, à minha mãe, à minha irmã, Mariana e à minha namorada, Jéssica.

Resumo

Os serviços de apoio ao domicílio assumem um papel crescente nos sistemas de saúde, respondendo ao envelhecimento populacional, ao aumento de doenças crónicas e à preferência social por cuidados prestados no lar. Estes serviços dependem fortemente de profissionais que realizam visitas programadas a diferentes locais, o que requer uma coordenação exigente de horários, rotas e competências. No entanto, a gestão destas operações apresenta desafios significativos para os quais os métodos de planeamento existentes podem ser insuficientes, sobretudo face à complexidade crescente do setor.

A ineficiência destes métodos origina problemas operacionais significativos, nomeadamente tempos excessivos de deslocação, desequilíbrios na distribuição da carga de trabalho, dificuldades em responder a alterações de última hora e aumento dos custos globais, com impacto direto na qualidade e pontualidade dos cuidados prestados. O problema de investigação emerge da dificuldade em articular, de forma coerente e eficiente, variáveis logísticas, clínicas e humanas, incorporando múltiplos fatores como janelas temporais, prioridades clínicas, disponibilidade dos profissionais e restrições geográficas.

Neste contexto, a presente dissertação tem como objetivo analisar de que forma a Inteligência Artificial pode ser utilizada para otimizar o planeamento de rotas e a alocação de recursos humanos em serviços de apoio ao domicílio. Para tal, é proposta uma abordagem integrada que combina modelos preditivos de procura, baseados em aprendizagem automática, com algoritmos de otimização inspirados em variantes do *Vehicle Routing Problem*, materializada no desenvolvimento de um protótipo funcional.

A solução proposta é avaliada em cenários simulados, recorrendo a dados públicos, permitindo analisar ganhos em termos de eficiência operacional, equilíbrio da carga de trabalho e qualidade do serviço prestado, contribuindo para a aplicação prática de sistemas inteligentes no contexto dos cuidados domiciliários.

Palavras-chave: Inteligência Artificial, Otimização de Rotas, Alocação de Recursos Humanos, Vehicle Routing Problem, Apoio ao Domicílio, Aprendizagem Automática.

Abstract

Home care services play an increasingly important role in healthcare systems, responding to population ageing, the rise in chronic diseases and the social preference for care provided in the home environment. These services rely heavily on professionals who carry out scheduled visits to different locations, requiring demanding coordination of schedules, routes and skills. However, the management of these operations presents significant challenges, for which existing planning methods may be insufficient, particularly given the growing complexity of the sector.

The inefficiency of these methods leads to significant operational problems, namely excessive travel times, imbalances in workload distribution, difficulties in responding to last-minute changes and increased overall costs, with a direct impact on the quality and timeliness of the care provided. The research problem emerges from the difficulty of articulating, in a coherent and efficient manner, logistical, clinical and human variables, incorporating multiple factors such as time windows, clinical priorities, staff availability and geographical constraints.

In this context, this dissertation aims to analyse how Artificial Intelligence can be used to optimise route planning and human resource allocation in home care services. To this end, an integrated approach is proposed that combines demand predictive models, based on machine learning techniques, with optimisation algorithms inspired by variants of the Vehicle Routing Problem, materialised through the development of a functional prototype.

The proposed solution is evaluated in simulated scenarios, using public datasets, allowing the analysis of gains in terms of operational efficiency, workload balance and quality of the service provided, thus contributing to the practical application of intelligent systems in the context of home care.

Keywords: Artificial Intelligence, Route Optimisation, Human Resource Allocation, Vehicle Routing Problem, Home Care, Machine Learning.

Agradecimentos

Com a finalização desta dissertação, marco o culminar de anos de estudo e trabalho, nos quais o esforço investido é agora recompensado com o término da vida académica e o início de uma nova fase. Não seria possível chegar a este momento sem o apoio e a ajuda de todas as pessoas que, direta ou indiretamente, contribuíram para este percurso tão importante na minha vida pessoal e profissional.

Quero expressar a minha profunda gratidão à minha família, cujo apoio incondicional tornou possível a conclusão desta dissertação. Aos meus pais, que sempre acreditaram no meu potencial e me incentivaram a perseguir os meus objetivos com determinação, transmitindo-me valores de esforço, perseverança e coragem. À minha irmã Mariana, agradeço pela presença constante, pelo incentivo e pelas pequenas ajudas que tornaram este percurso mais leve e suportável. O carinho, a dedicação e a confiança que sempre me ofereceram foram fundamentais, funcionando como fonte de motivação e energia em todos os momentos deste desafio académico.

Agradeço também aos professores do ISEP, cujas aulas, conselhos e orientações foram fundamentais para a minha formação académica. Cada discussão em sala, cada desafio proposto e cada esclarecimento contribuíram para o desenvolvimento do meu pensamento crítico e da minha capacidade de análise, preparando-me para enfrentar este projeto com rigor e confiança. A dedicação e paixão pelo ensino que demonstraram ao longo destes anos inspiraram-me a valorizar o conhecimento e a disciplina como ferramentas essenciais para o sucesso profissional.

Agradeço à minha orientadora, Fátima Rodrigues, e ao meu coorientador, Hélder Pinto, pelo apoio, orientação e disponibilidade ao longo de todo o desenvolvimento desta dissertação.

Por fim, mas de forma muito especial, quero deixar um agradecimento muito especial à minha Jéssica. Contigo aprendi a enfrentar desafios com profissionalismo e a perseguir os meus objetivos com coragem e determinação. Nos momentos mais difíceis, quando parecia impossível continuar, foste tu que me sustentaste e me lembraste do meu próprio valor. Obrigado por me inspirares todos os dias, por me motivares a dar sempre o meu melhor e por estares ao meu lado, tornando-me uma versão mais forte e melhor de mim mesmo.

Índice

1	Introdução	1
1.1	Contexto e Motivação	1
1.2	Enquadramento do Problema.....	2
1.3	Objetivos	2
1.4	Questões de Investigação	3
1.5	Contributos	4
1.6	Metodologia.....	5
1.7	Considerações Éticas	6
2	Estado da Arte	9
2.1	Serviços de Apoio ao Domicílio e Desafios Operacionais	9
2.2	Otimização de Rotas em Apoio ao Domicílio	9
2.2.1	Extensões do VRP aplicáveis aos Serviços de apoio ao Domicílio.....	10
2.2.2	Métodos de Solução.....	10
2.2.3	Limitações e Desafios nos Serviços de Apoio ao Domicílio	11
2.3	Previsão de Procura em Serviços de Saúde.....	11
2.3.1	Tipos de Modelos	11
2.3.2	Métricas de Avaliação	12
2.3.3	Limitações e Lacunas	12
2.3.4	Relevância para o projeto	12
2.4	Abordagens Híbridas	13
2.5	Planeamento Adaptativo e Re-Otimização	13
2.6	Segurança e Privacidade de Dados em Sistemas de Saúde	14
2.6.1	Proteção de Dados Sensíveis.....	14
2.6.2	Segurança das Infraestruturas.....	15
2.6.3	Integridade e Fiabilidade dos Modelos de IA	16
2.6.4	Conformidade Legal e Ética	16
2.7	Ferramentas e Tecnologias Seleccionadas	17
2.7.1	Solvers de otimização	17
2.7.2	Modelos de previsão	18
2.7.3	Interpretabilidade dos modelos	18
2.7.4	Ambiente de desenvolvimento	19
3	Definição e caracterização dos Datasets.....	21
3.1	CrITÉrios de seleção.....	22
3.2	Datasets de Otimização de Rotas	22
3.2.1	HHCRSP Dataset	23
3.2.2	Bi-Objective HHC.....	23
3.2.3	HHC Frota Eléctrica	23

3.3	Datasets de Previsão de Procura	24
3.3.1	Healthcare No-Shows	24
3.3.2	Medical Appointment Scheduling	24
3.3.3	CMS Home Health Services	25
3.4	Limitações dos Datasets face ao problema real	25
3.5	Estratégias de Utilização	26
4	Análise de Requisitos	27
4.1	Identificação dos Stakeholders	27
4.2	Requisitos Funcionais	28
4.3	Requisitos Não Funcionais.....	30
4.4	Casos de Uso	32
4.5	Pressupostos e Restrições	33
4.6	Matriz de Rastreabilidade	34
5	Desenho da Solução e Arquitetura	37
5.1	Visão Geral da Arquitetura.....	37
5.2	Módulo 1 - Dados e Pré-processamento	39
5.3	Módulo 2 - Previsão de Procura e de Cancelamentos	40
5.3.1	Previsão de procura	40
5.3.2	Previsão de cancelamentos	40
5.4	Módulo 3 - Otimização de Rotas.....	40
5.4.1	Função Objetivo	41
5.5	Módulo 4 - Re-otimização Adaptativa	42
5.6	Módulo 5 - Avaliação e Comparação	43
5.7	Fluxo de Dados entre Módulos.....	44
5.8	Decisões de Design e Alternativas Consideradas	45
5.9	Diagrama de Casos de Uso	46
6	Implementação	49
6.1	Ambiente de Desenvolvimento e Configuração	49
6.1.1	Estrutura de pastas de projeto	49
6.1.2	Dependências e instalação	50
6.1.3	Configuração global	51
6.2	Módulo 1 - Dados e Pré-processamento	52
6.2.1	Pré-processamento dos datasets de previsão	52
6.2.2	Preparação das instâncias de otimização	53
6.2.3	Análise Exploratória dos Dados	55
6.3	Módulo 2 - Previsão de Procura	58
6.3.1	Preparação dos dados e pré-processamento	58
6.3.2	Pré-processamento dos datasets de previsão	58

6.3.3	Interpretabilidade com SHAP	60
6.3.4	Modelo de previsão de procura – Média Móvel	61
6.3.5	Síntese do Módulo 2.....	63
6.4	Módulo 3 - Otimização de rotas	63
6.4.1	Configuração das instâncias de teste.....	63
6.4.2	Solver OR-Tools (VRPTW com Guided Local Search)	64
6.4.3	Solver PyVRP	65
6.4.4	Algoritmos Genéticos (DEAP)	66
6.4.5	Comparação de solvers	67
6.4.6	Síntese do Módulo 3.....	69
6.5	Módulo 4 - Re-otimização Adaptativa	69
6.5.1	Cenário A - Ausência de visitas de risco	70
6.5.2	Cenário B - Re-otimização por visitas em risco	71
6.5.3	Análise dos resultados e limitações	73
6.5.4	Síntese do Módulo 4.....	73
6.6	Módulo 5 - Avaliação e Comparação.....	74
6.6.1	Métricas de avaliação.....	74
6.6.2	Avaliação do mecanismo de re-otimização.....	74
6.6.3	Comparação dos solvers com métrica unificada	75
6.6.4	Síntese do Módulo 5.....	77
6.7	Pipeline Integrado	78
6.7.1	Estrutura do pipeline	78
6.7.2	Resultados da execução integrada.....	79
6.7.3	Síntese da integração.....	79
6.8	Considerações de Implementação	80
6.8.1	Reprodutibilidade	80
6.8.2	Limitações da matriz de distâncias.....	80
6.8.3	Gestão de dados e conformidade com o RGPD.....	81
7	Avaliação e Resultados	83
7.1	Metodologia de avaliação	84
7.2	Escalabilidade dos solvers	85
7.2.1	OR-Tools.....	85
7.2.2	PyVRP	86
7.2.3	DEAP	86
7.2.4	Comparação global.....	87
7.2.5	Síntese	91
7.3	Avaliação do mecanismo de Re-otimização	92
7.3.1	Metodologia	92
7.3.2	Sensibilidade ao limiar de cancelamento.....	92
7.3.3	Robustez do mecanismo.....	94
7.3.4	Síntese	94
7.4	Avaliação Comparativa Final	95
7.4.1	Comparação de solvers em instâncias com maior escala	95
7.4.2	Avaliação do pipeline completo.....	95
7.4.3	Síntese	99

7.5	Discussão dos Resultados.....	99
7.5.1	Desempenho dos solvers de otimização.....	100
7.5.2	Desempenho do mecanismo de rolling horizon	100
7.5.3	Desempenho do modelo preditivo	100
7.6	Limitações e Trabalho Futuro	101
7.6.1	Limitações identificadas	101
7.6.2	Trabalho futuro.....	101
8	Conclusão.....	103
8.1	Síntese do trabalho realizado	103
8.2	Resposta aos objetivos	103
8.3	Contributos.....	104
8.4	Limitações.....	105
8.5	Trabalho futuro.....	105

Lista de Figuras

Figura 1 – Arquitetura geral do sistema de otimização de rotas em home care	38
Figura 2 - Fluxo de dados entre os módulos do protótipo	44
Figura 3 - Diagrama UML de casos de uso do protótipo	47
Figura 4 - Análise exploratória do dataset Healthcare No-shows.....	56
Figura 5 - Localização geográfica dos nós — instância HHCRSP de 10 pacientes	56
Figura 6 - Matriz de confusão e curva ROC — modelo de cancelamentos XGBoost	60
Figura 7 - Análise SHAP — importância global e impacto por variável.....	61
Figura 8 - Previsão de procura diária — Modelo de Média Móvel MA(3) e MA(5)	62
Figura 9 - Rotas OR-Tools — instância HHCRSP_10_3_100_0.4_C_C.....	65
Figura 10 - Cenário A: probabilidades de cancelamento abaixo do limiar.....	71
Figura 11 - Cenário B: 5 visitas em risco.....	73
Figura 12 - Comparação de solvers com métrica unificada - 10 pacientes.....	77
Figura 13 - Distância total percorrida.....	90
Figura 14 - Equilíbrio de carga (std).....	91
Figura 15 - Cumprimento das janelas temporais (%)	91
Figura 16 - Sensibilidade ao limiar de cancelamento.....	94
Figura 17 - Distância total percorrida: Baseline vs. Re-otimizado.....	98
Figura 18 - Equilíbrio da carga (std): Baseline vs. Re-otimizado	98
Figura 19 - Cumprimento das janelas temporais (%): Baseline vs. Re-otimizado	99

Lista de Tabelas

Tabela 1 Datasets selecionados para o protótipo.....	21
Tabela 2 - Requisitos funcionais do sistema	28
Tabela 3 - Requisitos não funcionais do sistema	30
Tabela 4 - Casos de uso do Sistema.....	32
Tabela 5 - Pressupostos e restrições do sistema.....	34
Tabela 6 - Matriz de rastreabilidade: Requisitos funcionais vs. Objetivos.....	35
Tabela 7 - Resumo dos módulos da arquitetura	38
Tabela 8 - Decisões de design e alternativas consideradas	45
Tabela 9 - Dependências do projeto e comandos de instalação.....	51
Tabela 10 - Resumo das instâncias HHCRSP carregadas na validação do M1.....	57
Tabela 11 - Resultados da previsão de cancelamentos — XGBoost vs. LightGBM	59
Tabela 12 - Resultados do modelo de procura — Média Móvel.....	62
Tabela 13 - Solução gerada pelo OR-Tools para a instância de 10 pacientes	64
Tabela 14 - Solução gerada pelo PyVRP (instância 10 pacientes)	66
Tabela 15 - Melhor solução obtida pelo algoritmo genético (DEAP)	67
Tabela 16 - Restrições e funcionalidades suportadas por cada solver	67
Tabela 17 - Comparação de solvers — instâncias de 10 e 25 pacientes.....	68
Tabela 18 - Resultados do Cenário A (sem visitas de risco)	70
Tabela 19 - Visitas identificadas com elevado risco de cancelamento no Cenário B.....	71
Tabela 20 - Planos ajustados dos três profissionais	72
Tabela 21 - Resultados Globais do Cenário B.....	72
Tabela 22 - Avaliação do mecanismo de rolling horizon	75
Tabela 23 - Comparação de solvers com métrica unificada - 10 pacientes	76
Tabela 24 - Resultados da execução integrada do pipeline	79
Tabela 25 - Instâncias HHCRSP utilizadas no processo de avaliação	84
Tabela 26 - Resultados obtidos para as instâncias de 25 pacientes	87
Tabela 27 - Resultados obtidos para as instâncias de 50 pacientes	88
Tabela 28 - Resultados obtidos para as instâncias de 75 pacientes	89
Tabela 29 - Sensibilidade ao limiar de cancelamento	92
Tabela 30 - Comparação entre plano base e plano re-otimizado	96

Acrónimos e Símbolos

Lista de Acrónimos

IA	<i>Inteligência Artificial</i>
VRP	<i>Vehicle Routing Problem</i>
MAE	<i>Mean Absolute Error</i>
RMSE	<i>Root Mean Squared Error</i>
CRISP-DM	<i>Cross-Industry Standard Process for Data Mining</i>
API	<i>Application Programming Interface</i>
DDoS	<i>Distributed Denial of Service</i>
RGPD	<i>Regulamento Geral sobre a Proteção de Dados</i>
WHO	<i>World Health Organization</i>
HHCSP	<i>Home Health Care Routing and Scheduling Problem</i>
HHC	<i>Home Health Care</i>
RF	<i>Requisito Funcional</i>
RNF	<i>Requisito Não Funcional</i>
UC	<i>Use Case (caso de uso)</i>
ML	<i>Machine Learning</i>
OR	<i>Operations Research</i>
UML	<i>Unified Modeling Language</i>
CSV	<i>Comma-Separated Values (Valores Separados por Vírgulas)</i>
JSON	<i>JavaScript Object Notation (Notação de Objetos JavaScript)</i>
AUC-ROC	<i>Area Under the Receiver Operating Characteristic Curve</i>
SHAP	<i>SHapley Additive exPlanations</i>
CV	<i>Cross-Validation</i>

VRPTW	<i>Vehicle Routing Problem with Time Windows</i>
LSTM	<i>Long Short-Term Memory</i>
MA	<i>Moving Average</i>
TW %	<i>Time Window compliance rate (%)</i>

Lista de Símbolos

1 Introdução

Neste capítulo é apresentado o enquadramento geral da dissertação, começando pela contextualização do problema que motivou a investigação e pelas razões pessoais e profissionais que levaram à escolha do tema. Em seguida, é apresentada a solução proposta e são definidos os objetivos do projeto, bem como as questões de investigação que orientam o trabalho. São também descritas a metodologia adotada, os principais stakeholders envolvidos e os contributos esperados com a realização da dissertação. Por fim, são discutidas as considerações éticas relevantes para a execução do estudo.

1.1 Contexto e Motivação

O envelhecimento da população em Portugal tem vindo a aumentar de forma significativa, gerando uma procura crescente por serviços de apoio ao domicílio, que permitem aos utentes receber cuidados de saúde sem necessidade de deslocação a unidades hospitalares ou centros de saúde (INE, 2025). Esta tendência, aliada à preferência social por cuidados prestados no lar, coloca pressão sobre a organização destes serviços, que dependem de profissionais de saúde e cuidadores para visitar múltiplos utentes de forma eficiente e atempada. A implementação de um sistema inteligente apresenta potencial para reduzir tempos de deslocação, equilibrar a carga de trabalho dos profissionais e melhorar a qualidade e pontualidade dos cuidados prestados aos utentes, reforçando a relevância do estudo para a prática clínica e operacional.

Apesar do desenvolvimento de métodos de planeamento e ferramentas de otimização, a maioria das abordagens atuais baseia-se em heurísticas estáticas ou meta-heurísticas, que não se adaptam dinamicamente às variações na procura nem às competências diferenciadas dos profissionais, limitando a sua aplicabilidade em contextos reais e voláteis. Por este motivo, surge a necessidade de investigar soluções que integrem técnicas de Inteligência Artificial, combinando modelos preditivos de procura com algoritmos de otimização de rotas e alocação de recursos humanos, de forma a aumentar a eficiência operacional e a qualidade dos cuidados prestados.

A motivação pessoal e profissional para a realização desta dissertação advém do interesse em aplicar métodos de IA a problemas complexos de planeamento em saúde, contribuindo simultaneamente para a melhoria da sustentabilidade do sistema de apoio ao domicílio.

1.2 Enquadramento do Problema

Os serviços de apoio ao domicílio assumem um papel crescente nos sistemas de saúde, respondendo ao envelhecimento populacional, ao aumento de doenças crónicas e à preferência social por cuidados prestados no lar. Estes serviços dependem fortemente de profissionais que realizam visitas programadas a diferentes locais, o que requer uma coordenação exigente de horários, rotas e competências. No entanto, a gestão destas operações apresenta desafios significativos para os quais os métodos de planeamento existentes podem ser insuficientes, sobretudo face à complexidade crescente do setor.

A incapacidade de ajustar dinamicamente o planeamento às preferências dos utentes e às competências dos profissionais tende a resultar em soluções pouco robustas, com impacto direto na eficiência operacional e na qualidade do serviço prestado.

A ineficiência destes métodos origina problemas operacionais significativos: tempos excessivos de deslocação, desequilíbrios na distribuição da carga de trabalho, dificuldades em responder a alterações de última hora e aumento dos custos globais. Estas falhas refletem-se diretamente na qualidade e na pontualidade dos cuidados prestados ao utente, podendo comprometer a continuidade terapêutica e o bem-estar das equipas.

O problema de investigação emerge precisamente desta dificuldade em articular, de forma coerente e eficiente, variáveis logísticas, clínicas e humanas. A necessidade de incorporar múltiplos fatores — desde janelas temporais, prioridades clínicas e disponibilidade dos profissionais até restrições geográficas torna este problema particularmente complexo.

Vários grupos de stakeholders são diretamente afetados pelo problema: profissionais de saúde que realizam as visitas, gestores que planeiam as escalas e organizam operações, entidades prestadoras que enfrentam pressões de custos e eficiência, e os próprios pacientes, cuja qualidade de cuidados depende de uma logística eficaz. A estes soma-se a importância institucional de assegurar serviços sustentáveis e equitativos, reforçando a pertinência de uma solução tecnológica que responda a estas necessidades multidimensionais.

1.3 Objetivos

O presente trabalho teve como objetivo desenvolver modelos preditivos capazes de antecipar padrões de procura por serviços ao domicílio, utilizando técnicas de aprendizagem automática. Em seguida, foram implementados algoritmos de otimização de rotas e de alocação de recursos humanos, considerando restrições operacionais, nomeadamente janelas temporais e tempos

de deslocação. A integração da previsão da procura com o planeamento operacional foi materializada num protótipo funcional, cujo desempenho foi avaliado em termos de eficiência operacional, equilíbrio da carga de trabalho e qualidade do serviço prestado aos utentes. Por fim, foram produzidas evidências sobre os ganhos operacionais decorrentes da aplicação do protótipo em cenários simulados.

Neste contexto, a contribuição original desta dissertação consiste no desenvolvimento de uma abordagem integrada de Inteligência Artificial para serviços de apoio ao domicílio, que articula modelos preditivos de procura com algoritmos de otimização multi-objetivo para o planeamento de rotas e a alocação de recursos humanos. Ao contrário das abordagens tradicionais identificadas na literatura, que tendem a tratar a previsão da procura e o planeamento operacional de forma independente ou recorrendo a heurísticas estáticas, a solução proposta permite a adaptação dinâmica do planeamento com base em previsões atualizadas da procura, considerando simultaneamente restrições operacionais, competências profissionais e prioridades clínicas. Enquanto as abordagens heurísticas tradicionais produzem soluções fixas para cenários pré-definidos, a solução proposta permite reajustar o planeamento sempre que ocorrem variações na procura ou na disponibilidade de recursos, aumentando a robustez do sistema em contextos operacionais dinâmicos.

Os resultados obtidos demonstraram que esta integração contribui para a melhoria da eficiência operacional, permitindo reduzir o tempo de deslocação e adaptar dinamicamente o planeamento em resposta ao risco de cancelamento das visitas.

1.4 Questões de Investigação

A pergunta central que orienta este trabalho é:

De que forma a Inteligência Artificial pode otimizar o planeamento de rotas e a alocação de recursos humanos em serviços de apoio ao domicílio, considerando métricas de eficiência operacional como o tempo total de deslocação, o equilíbrio da carga de trabalho entre profissionais e cuidadores, e o cumprimento de janelas temporais?

A partir desta questão central, emergem três subperguntas de investigação que estruturam o desenvolvimento do trabalho:

1. Como prever de forma fiável a procura por serviços ao domicílio, de modo a reduzir desequilíbrios na alocação de recursos entre profissionais e cuidadores?
2. Quais os modelos de otimização mais adequados às restrições específicas destes serviços, nomeadamente janelas temporais, prioridades clínicas e heterogeneidade da força de trabalho?

3. De que forma é possível integrar previsões de procura com o planeamento operacional, de modo a suportar a re-otimização periódica das rotas em resposta a variações da procura ou da disponibilidade de recursos?

Estas subperguntas correspondem diretamente às três componentes centrais do protótipo proposto: o módulo preditivo de procura, o módulo de otimização de rotas e alocação de recursos, e o mecanismo de integração e re-otimização adaptativa.

1.5 Contributos

A presente dissertação contribui para a análise crítica do uso de Inteligência Artificial em serviços de apoio ao domicílio, a avaliação do potencial da combinação de algoritmos de aprendizagem automática para previsão da procura com heurísticas/meta-heurísticas para otimização de rotas e alocação de recursos humanos, e o desenvolvimento de um protótipo experimental que permita validar estas abordagens em cenários simulados. Os principais contributos podem ser sumarizados em três pontos:

1. **Desenvolvimento de abordagens híbridas para planeamento operacional:**
Foi proposta uma integração entre modelos preditivos de procura (baseados em aprendizagem automática) e algoritmos heurísticos/meta-heurísticos de otimização de rotas e recursos humanos. Esta abordagem híbrida permite adaptar dinamicamente o planeamento a alterações na procura ou na disponibilidade de profissionais, aumentando a eficiência e robustez do sistema.
2. **Aplicação de heurísticas adaptativas e rolling horizon:**
Em vez de soluções estáticas, o planeamento é reavaliado periodicamente, incorporando novas previsões e ajustando escalas sem reconstruir todo o plano. Esta estratégia garante uma resposta mais flexível às flutuações operacionais, refletindo melhor as necessidades reais do serviço e dos utentes.
3. **Avaliação comparativa rigorosa de soluções:**
Foi realizada uma análise comparativa entre abordagens tradicionais e a solução proposta, recorrendo a métricas operacionais comuns, como o tempo total de deslocação, o equilíbrio da carga de trabalho, os custos operacionais e a qualidade do serviço. Esta avaliação permitiu analisar o impacto da integração entre previsão da procura e otimização adaptativa.

De forma geral, a dissertação contribui para:

- A compreensão do potencial da combinação de aprendizagem automática com heurísticas para problemas de VRP em contextos de saúde domiciliária;
- O desenvolvimento de um protótipo funcional que valida a abordagem em cenários simulados;

- A produção de evidências sobre os ganhos operacionais e melhoria da qualidade do serviço prestado aos utentes.

1.6 Metodologia

A metodologia adotada neste trabalho segue uma abordagem integrada de engenharia e investigação aplicada, assente no desenvolvimento e avaliação de um sistema inteligente para apoio ao planeamento de serviços domiciliários.

O processo inicia com a revisão estruturada da literatura, conduzida em bases de dados científicas de referência, nomeadamente IEEE Xplore, PubMed e ScienceDirect. A pesquisa utilizou queries construídas a partir de palavras-chave como *home care logistics*, *vehicle routing problem*, *healthcare resource allocation*, *machine learning demand prediction*, *AI scheduling*, e *home healthcare optimization*. Os critérios de inclusão consideraram estudos publicados nos últimos vinte anos, com aplicação direta ou potencial a serviços ao domicílio ou problemas logísticos análogos, e com validação empírica. Os critérios de exclusão abrangeram trabalhos genéricos, sem validação, ou exclusivamente hospitalares. Esta revisão permitiu identificar lacunas de investigação, fundamentando o desenvolvimento do protótipo.

Posteriormente, foram utilizados datasets públicos que simulam cenários realistas de prestação de cuidados domiciliários, incluindo a localização geográfica de utentes, perfis genéricos de necessidades, competências dos profissionais, janelas temporais e tempos estimados de deslocação. Os dados foram tratados e analisados para identificar padrões relevantes, avaliar a variabilidade da procura e preparar modelos preditivos, sem recorrer a informação pessoal identificável.

A componente de previsão foi desenvolvida recorrendo a técnicas de aprendizagem automática, seguindo a metodologia CRISP-DM (Data Science Central, 2025), que contempla:

1. Seleção e limpeza dos dados;
2. Desenvolvimento de modelos preditivos;
3. Avaliação do desempenho através de métricas como *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE) e *Area Under the ROC Curve* (AUC-ROC), consoante a natureza da tarefa;
4. Implementação (*deployment*) em cenários simulados para validar a capacidade de generalização.

Em paralelo, foi implementado um modelo de otimização baseado em variantes do *Vehicle Routing Problem with Time Windows* (VRPTW), incorporando restrições típicas dos serviços ao domicílio, como janelas temporais, prioridades de atendimento e múltiplas competências profissionais.

A integração entre previsão e otimização foi materializada num protótipo funcional desenvolvido em Python, permitindo que cenários de procura simulados alimentem dinamicamente o módulo de planeamento. Esta abordagem permite testar, comparar e ajustar diferentes modelos e algoritmos em ambientes controlados, garantindo viabilidade técnica, rigor científico e reprodutibilidade dos resultados.

Deste modo, o contributo inovador da metodologia não reside apenas na aplicação de Inteligência Artificial, mas na sua utilização como mecanismo central de suporte à decisão, integrando previsão e otimização num único ciclo de planeamento operacional.

1.7 Considerações Éticas

O presente projeto foi avaliado à luz dos princípios éticos aplicáveis à investigação em tecnologias de informação e saúde. Não envolve recolha de dados primários, participação direta de indivíduos, nem utilização de dados pessoais ou clínicos reais. Os datasets utilizados são públicos, de referência, não permitindo qualquer identificação direta ou indireta de pessoas ou instituições.

Assim, não se identificam riscos éticos relevantes associados ao tratamento de dados ou à proteção da privacidade. Ainda assim, o projeto segue boas práticas de investigação responsável, transparência metodológica e reprodutibilidade dos resultados.

No que respeita à proteção de dados, o projeto respeita os princípios estabelecidos pelo Regulamento Geral sobre a Proteção de Dados (RGPD), nomeadamente os princípios da minimização de dados, da finalidade e da integridade e confidencialidade. Embora o projeto não processe dados pessoais reais, estes princípios orientam as opções de design do sistema proposto, em particular no que respeita ao tratamento de informação sensível em contextos operacionais futuros.

Adicionalmente, dado que o sistema desenvolvido se enquadra no domínio da saúde, reconhece-se a relevância das diretrizes internacionais sobre ética e governação da Inteligência Artificial, designadamente as orientações da Organização Mundial de Saúde (WHO, 2021), que estabelecem princípios como a transparência algorítmica, a responsabilização e a não maleficência como requisitos fundamentais para sistemas de IA aplicados à saúde.

Do ponto de vista profissional, o desenvolvimento do trabalho assenta em competências técnicas nas áreas de Inteligência Artificial, otimização e engenharia de software, assegurando que as metodologias aplicadas são adequadas, rigorosas e eticamente responsáveis.

No que respeita à utilização de ferramentas de inteligência artificial durante a elaboração desta dissertação, declara-se que foram utilizados assistentes de escrita baseados em modelos de linguagem para apoio à redação e revisão de texto, bem como para suporte ao desenvolvimento de código Python. A utilização destas ferramentas teve carácter auxiliar,

sendo a responsabilidade pela concepção, análise e interpretação dos resultados integralmente do autor.

2 Estado da Arte

2.1 Serviços de Apoio ao Domicílio e Desafios Operacionais

A otimização de operações em serviços de apoio ao domicílio constitui um domínio de investigação marcadamente interdisciplinar, no qual convergem contributos da logística, da investigação operacional, da aprendizagem automática e dos sistemas inteligentes aplicados à saúde (Fikar & Hirsch, 2017). A crescente complexidade destes serviços, associada à diversidade de restrições operacionais e à variabilidade da procura, tem motivado o desenvolvimento de modelos de planeamento mais eficientes e adaptáveis (Pereira, 2022).

No que respeita à otimização de rotas, a literatura identifica o Vehicle Routing Problem como o modelo de referência para o planeamento de deslocações sob múltiplas restrições (Oliveira, 2015). Diversas extensões têm sido propostas para acomodar características específicas dos serviços de saúde ao domicílio, incluindo janelas temporais, prioridades clínicas, múltiplas competências dos profissionais e incerteza nos tempos de deslocação (Jesus, 2023). Dada a complexidade destes problemas, a investigação tem recorrido predominantemente a abordagens heurísticas e meta-heurísticas (Decerle et al., 2018).

Apesar dos avanços alcançados, vários autores sublinham limitações relevantes destas abordagens quando aplicadas a contextos dinâmicos. Em particular, os modelos de VRP dinâmico e com janelas de tempo tendem a produzir soluções eficazes apenas para cenários previamente definidos que não garantem uma adaptação contínua às alterações da procura ou da disponibilidade de recursos (Oliveira, 2022). Esta dependência de heurísticas parametrizadas e de pressupostos relativamente estáticos revela-se especialmente problemática em serviços de apoio ao domicílio, onde cancelamentos, urgências e variações inesperadas são frequentes.

2.2 Otimização de Rotas em Apoio ao Domicílio

O Vehicle Routing Problem (VRP) constitui um dos problemas clássicos de otimização combinatória na logística, cujo objetivo é determinar rotas ótimas para uma frota de veículos

que deve atender a um conjunto de clientes, minimizando custos, distâncias ou tempo total de deslocação (Toth & Vigo, 2014). No contexto de serviços de apoio ao domicílio, o VRP é adaptado para atender às características específicas do domínio, incluindo necessidades clínicas, competências dos profissionais e restrições operacionais.

2.2.1 Extensões do VRP aplicáveis aos Serviços de apoio ao Domicílio

Para refletir a complexidade do planeamento em serviços ao domicílio, o Vehicle Routing Problem tradicional é estendido com várias dimensões:

- Janelas temporais: Cada visita deve ocorrer dentro de períodos específicos definidos pelas necessidades do utente e disponibilidade do profissional. A violação destas janelas compromete a qualidade do serviço e a satisfação do utente (Jesus, 2023).
- Prioridades clínicas: Nem todas as visitas têm o mesmo grau de urgência. O planeamento deve incorporar prioridades, garantindo que utentes com necessidades críticas sejam atendidos primeiro, sem prejudicar a eficiência global da operação.
- Múltiplas competências profissionais: Profissionais possuem competências diferenciadas. Algumas visitas exigem habilidades específicas (enfermeiro, fisioterapeuta), o que transforma o problema num VRP com recursos heterogêneos, aumentando a complexidade da alocação.
- Objetivos múltiplos implícitos: O planeamento deve equilibrar simultaneamente diversos objetivos, incluindo:
 - Minimização do tempo total de deslocação;
 - Redução de custos operacionais;
 - Equilíbrio da carga de trabalho entre profissionais;
 - Maximização da pontualidade e qualidade do atendimento.
- Incerteza nos tempos de deslocação: Condições de tráfego, atrasos em visitas ou cancelamentos introduzem variabilidade, exigindo modelos capazes de lidar com VRP estocástico ou dinâmico (Oliveira, 2022).

2.2.2 Métodos de Solução

Devido à elevada complexidade combinatória e às múltiplas restrições, a maioria das instâncias reais do VRP em home care não pode ser resolvida de forma exata em tempo computacionalmente aceitável. Por isso, predominam abordagens heurísticas e meta-heurísticas, que fornecem soluções próximas do ótimo com eficiência prática:

- Heurísticas clássicas que permitem a construção rápida de soluções viáveis, sendo frequentemente utilizadas como métodos iniciais ou como componentes de algoritmos mais complexos.
- Meta-heurísticas Explorando o espaço de soluções de forma mais abrangente, estas técnicas permitem escapar de ótimos locais e lidar com múltiplas restrições e objetivos complexos.
- Abordagens Híbridas propõem a combinação de modelos preditivos de procura, baseados em técnicas de aprendizagem automática, com heurísticas e meta-heurísticas de otimização. Estas abordagens permitem a re-otimização periódica do planeamento (rolling horizon), incorporando previsões atualizadas e ajustando as rotas e a alocação de recursos em resposta a alterações de última hora, como cancelamentos ou pedidos urgentes (Sales, 2022; Jesus, 2023).

2.2.3 Limitações e Desafios nos Serviços de Apoio ao Domicílio

Apesar dos avanços, os modelos existentes enfrentam desafios significativos:

1. Rigidez dos modelos: muitas soluções assumem cenários estáticos, sem considerar cancelamentos ou alterações de última hora.
2. Integração limitada com previsão de procura: a maioria das abordagens trata a otimização e a previsão de procura de forma independente.
3. Complexidade operacional: múltiplos objetivos, restrições de competências e prioridades dificultam a aplicação prática de métodos tradicionais de VRP.
4. Necessidade de adaptação contínua: contextos reais exigem planeamento adaptativo e re-otimização periódica, o que ainda é pouco explorado na literatura de home care.

2.3 Previsão de Procura em Serviços de Saúde

A previsão da procura em serviços de saúde constitui uma ferramenta essencial para a alocação eficiente de recursos, planeamento de visitas e melhoria da qualidade do serviço prestado (Mendes, 2016). Técnicas de aprendizagem automática (Machine Learning – ML) permitem identificar padrões complexos em séries temporais, lidar com dados heterogéneos e antecipar necessidades futuras de forma mais precisa do que métodos estatísticos tradicionais.

2.3.1 Tipos de Modelos

Entre os modelos mais utilizados destacam-se:

- Modelos de regressão: linear, úteis para relações lineares entre variáveis históricas e procura futura.

- Redes neurais recorrentes (RNN / LSTM): capazes de capturar dependências temporais e padrões não-lineares em séries temporais.
- Algoritmos ensemble: bagging (Random Forest) e boosting (XGBoost, LightGBM), eficazes para combinar múltiplos modelos e reduzir erro de previsão.
- Deep Learning: redes neurais profundas aplicadas a grandes volumes de dados, integrando múltiplas variáveis (idade, histórico clínico, localização, horário).

2.3.2 Métricas de Avaliação

Para avaliar o desempenho dos modelos de previsão, são utilizadas métricas como:

- **MAE (Mean Absolute Error):** erro absoluto médio entre valores previstos e reais;
- **RMSE (Root Mean Squared Error):** penaliza erros grandes, sendo útil quando se deseja maior precisão em picos de procura;

2.3.3 Limitações e Lacunas

Apesar do progresso, os modelos de previsão enfrentam algumas limitações:

1. Independência do planeamento operacional: na maioria dos estudos, as previsões são feitas isoladamente, sem integração com algoritmos de otimização de rotas ou alocação de recursos, limitando seu impacto real.
2. Dados escassos ou heterogêneos: a qualidade da previsão depende da quantidade e consistência de dados históricos, muitas vezes insuficiente em serviços domiciliários.
3. Adaptação a contextos dinâmicos: variações inesperadas de procura, cancelamentos ou urgências exigem atualização contínua das previsões, ainda pouco explorada.

2.3.4 Relevância para o projeto

A integração de modelos de previsão com o VRP e estratégias de otimização adaptativa permite:

- Ajustar proativamente o planeamento de rotas com base em previsões de procura;
- Reduzir tempos de deslocação e sobrecarga de profissionais;
- Aumentar a pontualidade e qualidade do atendimento.

Esta abordagem híbrida, combinando previsão de procura com otimização de rotas e recursos, constitui o núcleo inovador do presente trabalho, preenchendo uma lacuna identificada na literatura de serviços de saúde domiciliários (Fikar & Hirsch, 2017; Algendi et al., 2024).

2.4 Abordagens Híbridas

Estudos mais recentes evidenciam o potencial de abordagens híbridas, nas quais modelos de previsão da procura são integrados com algoritmos de otimização de rotas e de alocação de recursos humanos, permitindo uma adaptação mais eficaz às flutuações operacionais e contribuindo para melhorias significativas na eficiência dos serviços prestados (Sales, 2022; Pedro Seoane Jesus, 2023).

Estas abordagens procuram ultrapassar a separação tradicional entre previsão e planeamento, articulando ambas as componentes num único sistema de apoio à decisão. Tipicamente, a previsão é utilizada para antecipar padrões de procura futura — número de visitas, localização dos utentes ou intensidade dos cuidados — enquanto os algoritmos de otimização exploram essa informação para gerar planos de rotas e escalas mais robustos.

Na literatura, as abordagens híbridas assumem diferentes configurações, destacando-se:

- Integração sequencial, em que os resultados da previsão alimentam diretamente os modelos de otimização;
- Sistemas iterativos, nos quais previsões e planeamento são reavaliados periodicamente;
- Combinação de modelos de Machine Learning com heurísticas e meta-heurísticas para lidar com múltiplas restrições e objetivos.

Apesar dos resultados promissores reportados, a literatura aponta lacunas relevantes. Em particular, observa-se:

1. Escassez de modelos especificamente desenhados para o contexto de serviços de apoio ao domicílio;
2. Reduzida disponibilidade de protótipos validados em cenários próximos da realidade operacional;
3. Exploração limitada de mecanismos que permitam ajustar dinamicamente o planeamento com base em previsões atualizadas e eventos imprevistos (Fikar & Hirsch, 2017; Algendi et al., 2024).

É neste enquadramento que o presente projeto se insere, procurando contribuir para a colmatação desta lacuna através do desenvolvimento de uma abordagem integrada que articule a previsão da procura com a otimização do planeamento de rotas e de recursos humanos em serviços de apoio ao domicílio, com uma orientação clara para a aplicabilidade prática e para a validação experimental dos resultados obtidos.

2.5 Planeamento Adaptativo e Re-Otimização

Os serviços de apoio ao domicílio caracterizam-se por um elevado grau de dinamismo e incerteza, resultante de cancelamentos de visitas, surgimento de pedidos urgentes, alterações na disponibilidade dos profissionais e variações inesperadas nos tempos de deslocação. Neste contexto, abordagens de planeamento estático revelam-se frequentemente inadequadas (Oliveira, 2022).

O planeamento adaptativo surge como uma resposta a estas limitações, permitindo a atualização periódica dos planos de rotas e escalas à medida que nova informação se torna disponível. Entre as estratégias mais utilizadas destacam-se:

- Rolling horizon planning: o planeamento é reavaliado a intervalos regulares, incorporando previsões atualizadas e ajustando apenas parte do horizonte temporal;
- Re-otimização parcial: alterações pontuais são introduzidas nos planos existentes, evitando a reconstrução completa das rotas e minimizando o impacto operacional;
- Planeamento reativo: resposta imediata a eventos inesperados, como cancelamentos ou urgências, com impacto localizado nas rotas afetadas.

Estas estratégias permitem aumentar a robustez das soluções, melhorar a utilização dos recursos humanos e reduzir custos operacionais, mantendo simultaneamente níveis elevados de qualidade e pontualidade no atendimento aos utentes. A literatura evidencia que a combinação de planeamento adaptativo com modelos de previsão da procura constitui uma abordagem particularmente promissora para ambientes voláteis, como os serviços de apoio ao domicílio (Fikar & Hirsch, 2017; Sales, 2022).

Neste sentido, a presente dissertação adota uma perspetiva de planeamento adaptativo, explorando mecanismos de re-otimização periódica suportados por previsões de procura, como alternativa a soluções estáticas tradicionais.

2.6 Segurança e Privacidade de Dados em Sistemas de Saúde

O crescimento do uso de Inteligência Artificial em serviços de saúde domiciliários traz desafios críticos relacionados à segurança, privacidade e integridade dos dados, uma vez que os sistemas processam informações sensíveis de utentes, profissionais e operações logísticas. A proteção destes dados é essencial não apenas por motivos legais e éticos, mas também para garantir a fiabilidade do planeamento e a continuidade dos cuidados prestados.

2.6.1 Proteção de Dados Sensíveis

Os sistemas processam dados pessoais e clínicos, incluindo informações de saúde, localizações geográficas, horários e perfis de competências dos profissionais. O armazenamento e processamento desses dados sem medidas adequadas de segurança expõe os utentes a riscos de violação de privacidade, discriminação e uso indevido de informações.

Para mitigar estes riscos, podem ser aplicadas medidas bem estabelecidas de segurança e conformidade como:

- Anonimização de dados, garantindo que informações identificáveis não possam ser diretamente associadas a indivíduos.
- Encriptação em repouso e em trânsito, para impedir acessos não autorizados em bases de dados ou durante a transmissão de dados entre módulos de previsão e otimização.
- Auditoria e rastreabilidade, como exigido por regulamentações como o Regulamento Geral sobre a Proteção de Dados (RGPD), assegurando que todos os processos automatizados possam ser monitorizados e verificados.

Além disso, práticas de minimização de dados, armazenando apenas o essencial para previsão e otimização, reduzem a superfície de ataque e o risco de exposição indevida.

2.6.2 Segurança das Infraestruturas

A integração de algoritmos de previsão e otimização depende de infraestruturas conectadas, que podem incluir servidores locais, serviços de cloud e Application Programming Interfaces (APIs). As APIs permitem a comunicação entre módulos de software, a troca de dados e a integração de processos automatizados, sendo essenciais para o funcionamento de sistemas integrados. Contudo, estas infraestruturas estão sujeitas a diversos tipos de ataques informáticos, como ransomware, injeção de dados maliciosos ou ataques de negação de serviço distribuído (DDoS), que podem comprometer tanto os resultados do planeamento como a segurança dos utentes e a integridade dos dados processados.

Para mitigar estes riscos, podem ser aplicadas medidas de proteção robustas, tais como:

- Autenticação multifator e controlo de acessos, garantindo que apenas utilizadores autorizados possam aceder a sistemas críticos ou a dados sensíveis.
- Segmentação de redes e sistemas, isolando os dados mais sensíveis de ambientes públicos ou menos confiáveis, reduzindo a superfície de ataque.
- Monitorização contínua e deteção de anomalias, permitindo identificar rapidamente comportamentos suspeitos ou potenciais intrusões e responder de forma proativa.
- Gestão segura de APIs, incluindo autenticação baseada em tokens, limitação de chamadas por utilizador e validação de dados recebidos, para prevenir abusos ou a introdução de dados incorretos no sistema.

A implementação destas práticas permite reforçar a resiliência do sistema, assegurando que os algoritmos de previsão e otimização operam de forma confiável e que os dados sensíveis dos utentes permanecem protegidos.

2.6.3 Integridade e Fiabilidade dos Modelos de IA

Além de proteger os dados, é fundamental assegurar que os modelos de previsão e otimização operam com integridade e robustez, mesmo perante falhas ou ataques. A manipulação dos dados de treino ou a inserção de dados maliciosos pode conduzir a decisões incorretas, prejudicando a eficiência e a qualidade dos serviços.

Algumas técnicas de mitigação aplicáveis incluem:

- Validação contínua e testes de robustez dos modelos em cenários simulados antes da sua implementação operacional.
- Redundância de dados, garantindo que falhas numa fonte não comprometem o sistema global.
- Auditorias periódicas e gestão de versões dos modelos, permitindo identificar alterações inesperadas e reverter para versões confiáveis.

2.6.4 Conformidade Legal e Ética

Os sistemas na área da saúde domiciliária devem respeitar de forma rigorosa as regulamentações legais e os princípios éticos aplicáveis. Isto implica:

- Conformidade com o RGPD, incluindo o direito de acesso, retificação e eliminação de dados pessoais;
- Transparência nos processos de decisão, especialmente quando os algoritmos influenciam o planeamento de visitas ou a alocação de recursos;
- Responsabilização (accountability), garantindo que as decisões automatizadas possam ser justificadas e auditadas.

A evidência disponível indica que a negligência destes aspetos pode conduzir não apenas a incumprimentos legais, mas também à perda de confiança por parte dos profissionais e dos utentes.

2.7 Ferramentas e Tecnologias Selecionadas

A implementação do protótipo exigiu a seleção de um conjunto de ferramentas compatíveis com os requisitos identificados ao longo deste capítulo, nomeadamente a necessidade de resolver problemas de otimização de rotas com restrições temporais, desenvolver modelos preditivos baseados em aprendizagem automática e garantir a interpretabilidade das previsões produzidas. A seleção privilegiou soluções open-source amplamente utilizadas na comunidade científica, com documentação consolidada e evidência de aplicação em problemas de logística e saúde.

2.7.1 Solvers de otimização

A resolução do problema de otimização de rotas foi suportada por três abordagens distintas: OR-Tools, PyVRP e DEAP, selecionadas por representarem diferentes paradigmas de resolução do Vehicle Routing Problem.

O OR-Tools, desenvolvido pela Google (Perron & Furnon, 2023), constitui uma das bibliotecas de otimização combinatória mais utilizadas na literatura para problemas de Vehicle Routing Problem with Time Windows (VRPTW). A biblioteca disponibiliza suporte nativo para janelas temporais, múltiplos veículos, restrições de capacidade e diversas meta-heurísticas de pesquisa local, permitindo obter soluções de elevada qualidade em tempos computacionais reduzidos. Estas características justificam a sua utilização como solver principal do presente trabalho.

Como termo de comparação foi utilizado o PyVRP (Wouda, Lan & Kool, 2024), uma biblioteca open-source recente especificamente desenvolvida para problemas de roteamento de veículos. O PyVRP implementa algoritmos genéticos híbridos inspirados no trabalho de Vidal (2022), incorporando operadores especializados de cruzamento, mutação e pesquisa local. A literatura evidencia que esta abordagem produz soluções competitivas em diferentes variantes do VRP, constituindo uma referência atual para comparação de desempenho.

Foi ainda considerada uma terceira abordagem baseada na biblioteca DEAP (Distributed Evolutionary Algorithms in Python) (Fortin et al., 2012), amplamente utilizada para desenvolvimento de algoritmos evolucionários. Ao contrário das bibliotecas anteriores, o DEAP não disponibiliza implementações específicas para VRPTW, exigindo o desenvolvimento explícito da representação cromossómica, operadores genéticos e função de avaliação. A sua inclusão permite comparar uma implementação genética personalizada com os solvers especializados utilizados no protótipo.

A utilização destes três solvers possibilita uma avaliação comparativa entre diferentes estratégias de otimização — programação por restrições, algoritmos genéticos híbridos e algoritmos evolucionários — permitindo analisar os respetivos compromissos entre qualidade da solução, equilíbrio da carga de trabalho e tempo de execução.

2.7.2 Modelos de previsão

A componente preditiva do protótipo foi desenvolvida recorrendo a algoritmos de aprendizagem supervisionada adequados à análise de dados tabulares. Entre os modelos identificados na literatura, destacam-se particularmente os métodos de gradient boosting, devido à sua elevada capacidade preditiva, robustez perante dados heterogêneos e reduzida necessidade de pré-processamento.

Os dois modelos selecionados foram o XGBoost (Extreme Gradient Boosting) (Chen e Guestrin, 2016) e o LightGBM (Light Gradient Boosting Machine) (Ke et al., 2017).

O XGBoost constitui atualmente uma das técnicas mais utilizadas em competições de Machine Learning e aplicações industriais, destacando-se pela sua capacidade de lidar com relações não lineares entre variáveis, tratamento eficiente de valores em falta e mecanismos internos de regularização que reduzem o risco de sobreajuste. Estas características tornam-no particularmente adequado para problemas de classificação e regressão com datasets de dimensão média, como os utilizados nesta dissertação.

O LightGBM apresenta uma arquitetura semelhante, baseada em gradient boosting sobre árvores de decisão, mas incorpora estratégias de crescimento leaf-wise que reduzem significativamente o tempo de treino em datasets de grande dimensão. Embora apresente frequentemente desempenho semelhante ao XGBoost, o seu comportamento depende das características específicas dos dados utilizados.

A utilização destes dois algoritmos permitiu comparar diferentes abordagens de gradient boosting, selecionando posteriormente o modelo com melhor desempenho experimental para integração no pipeline do protótipo.

2.7.3 Interpretabilidade dos modelos

A crescente utilização de modelos de Inteligência Artificial em aplicações de saúde reforça a necessidade de mecanismos que permitam compreender as decisões produzidas pelos algoritmos. A interpretabilidade constitui atualmente um dos princípios fundamentais para a utilização responsável de sistemas inteligentes neste domínio (WHO, 2021).

Para esse efeito foi utilizada a biblioteca SHAP (SHapley Additive exPlanations) (Lundberg e Lee, 2017), baseada nos valores de Shapley da teoria dos jogos cooperativos. O SHAP permite quantificar a contribuição individual de cada variável para cada previsão produzida pelo modelo, possibilitando interpretar tanto o comportamento global como decisões individuais.

A integração desta ferramenta permite identificar quais as características dos pacientes que mais influenciam a probabilidade prevista de cancelamento, aumentando a transparência do sistema e facilitando a análise crítica do desempenho dos modelos desenvolvidos.

2.7.4 Ambiente de desenvolvimento

O protótipo foi desenvolvido em Python, devido à vasta disponibilidade de bibliotecas para aprendizagem automática, otimização combinatória e análise de dados. O processamento dos datasets foi realizado com recurso às bibliotecas pandas e NumPy, enquanto os modelos preditivos foram implementados com scikit-learn, XGBoost e LightGBM.

A implementação experimental foi organizada em Jupyter Notebook, permitindo integrar código, documentação e visualização de resultados num único ambiente de trabalho, favorecendo a reprodutibilidade das experiências realizadas.

A visualização dos resultados foi efetuada recorrendo às bibliotecas Matplotlib e Seaborn, utilizadas na construção dos gráficos comparativos apresentados nos capítulos experimentais.

3 Definição e caracterização dos Datasets

Neste capítulo são identificados e caracterizados os conjuntos de dados públicos selecionados para suportar o desenvolvimento e avaliação do protótipo proposto. A escolha recaiu exclusivamente sobre dados de acesso livre, em conformidade com os princípios éticos enunciados no Capítulo 1 e com os requisitos do Regulamento Geral sobre a Proteção de Dados (RGPD), designadamente os princípios da minimização e da finalidade.

Os datasets encontram-se organizados em dois grupos distintos, correspondentes aos dois módulos principais do protótipo: os datasets de otimização de rotas, destinados a alimentar o módulo de planeamento e alocação de recursos, e os datasets de previsão de procura, utilizados no desenvolvimento e validação dos modelos preditivos.

A Tabela 1 apresenta uma visão global dos seis datasets selecionados, com indicação da fonte, módulo de utilização, ano de publicação e licença de utilização.

Tabela 1 Datasets selecionados para o protótipo

Dataset	Fonte	Módulo	Ano	Licença
HHCRSP Dataset	GitHub (afkummer)	Otimização	2021	MIT
Bi-Objective HHC	figshare	Otimização	2022	CC BY 4.0
HHC Frota Elétrica (Ankara)	Mendeley Data	Otimização	2023	CC BY 4.0
Healthcare No-Shows	Kaggle	Previsão	2016	CC0

Dataset	Fonte	Módulo	Ano	Licença
Medical Appointment Scheduling	Kaggle	Previsão	2024	CC0
CMS Home Health Services	data.cms.gov	Previsão	Atual	Público

3.1 Critérios de seleção

A seleção dos datasets foi orientada por um conjunto de critérios técnicos e acadêmicos, definidos com base nos objetivos do projeto e nas características do problema em estudo.

Em primeiro lugar, privilegiou-se a relevância direta para o domínio dos serviços de apoio ao domicílio. Foram excluídos datasets exclusivamente hospitalares ou de contexto logístico genérico, sem restrições operacionais comparáveis às dos serviços domiciliários. Em segundo lugar, exigiu-se a disponibilidade pública e acesso livre, com licenças que permitam a utilização para fins de investigação académica. Em terceiro lugar, valorizou-se a diversidade de características, nomeadamente a presença de janelas temporais, competências de profissionais, dados geográficos reais e variabilidade de procura.

Reconhece-se, contudo, uma limitação estrutural comum a toda a literatura da área: não existe, até à data, um dataset público amplamente citado que capture especificamente padrões de procura em serviços de apoio ao domicílio com granularidade suficiente para treino de modelos preditivos. Esta lacuna é assinalada por vários autores na literatura (Fikar & Hirsch, 2017; Algendi et al., 2024) e constitui uma das motivações para o recurso a dados sintéticos gerados a partir das instâncias de otimização.

3.2 Datasets de Otimização de Rotas

Os três datasets selecionados para o módulo de otimização cobrem diferentes dimensões do problema: especificidade para home care, incerteza nos tempos de serviço e heterogeneidade de recursos.

3.2.1 HHCRSP Dataset

O HHCRSP Dataset 2021 disponibilizado na plataforma GitHub por Afkummer et al. (2021) constitui o conjunto de dados mais diretamente alinhado com o problema em estudo. As instâncias foram geradas com recurso a dados geográficos reais da cidade de Porto Alegre (Brasil), obtidos através da base de dados OpenAddresses, simulando cenários realistas de prestação de cuidados ao domicílio. O dataset introduz múltiplos tipos de serviço, múltiplas visitas por paciente e restrições de dependência entre rotas, características que refletem a complexidade operacional dos contextos reais de apoio ao domicílio (Afkummer et al., 2021).

A metodologia de geração das instâncias seguiu uma abordagem seletiva: foram calculados limites inferior e superior para todas as instâncias geradas, tendo sido selecionadas, por dimensão, as dez instâncias com maior gap em relação à solução ótima e dez instâncias adicionais de forma aleatória, garantindo assim diversidade e desafio algorítmico. Este dataset foi utilizado como referência principal para avaliação do módulo de otimização do protótipo.

As instâncias encontram-se disponíveis sob licença MIT, permitindo utilização académica e reprodutibilidade dos resultados.

3.2.2 Bi-Objective HHC

O *Bi-Objective HHC Dataset*, publicado na plataforma figshare (2022) e associado a um artigo publicado no International Journal of Environmental Research and Public Health, disponibiliza instâncias bi-objetivo para o planeamento de cuidados ao domicílio sob incerteza. As instâncias foram geradas a partir do benchmark de Solomon, com adaptações específicas para o contexto de home care, nomeadamente a introdução de incerteza nos tempos de serviço aos pacientes — uma característica frequentemente presente em contextos reais, onde variações clínicas e logísticas são imprevisíveis.

O dataset contempla três dimensões de instâncias: 25, 50 e 100 pacientes, organizadas em seis tipos (C1, C2, R1, R2, RC1, RC2), diferenciados pela distribuição geográfica dos clientes e pela rigidez das janelas temporais. A sua utilização no presente projeto permite avaliar o comportamento dos algoritmos de otimização em cenários com variabilidade nos tempos de serviço, simulando condições mais próximas da realidade operacional dos serviços domiciliários.

3.2.3 HHC Frota Elétrica

O dataset, desenvolvido por Yucel, Cebeci e Koc (2023) e disponibilizado na plataforma Mendeley Data aborda o estudo sobre o agendamento e o planeamento de rotas para cuidados ao domicílio com frota elétrica heterogénea, na cidade de Ancara (Turquia). As instâncias foram geradas com recurso a dados reais de localização geográfica provenientes da base de dados do Sistema de Registo de População por Morada (TUIK), combinando seleção aleatória e seleção ponderada por densidade populacional.

O dataset inclui dois níveis de competência para os profissionais de saúde, com três graus de especialização cada, e contempla custos fixos por profissional calculados com base em dados salariais reais. Para efeitos do presente projeto, são utilizadas apenas as componentes de localização geográfica dos pacientes, janelas temporais, competências dos profissionais e estrutura de visitas, excluindo a dimensão específica dos veículos elétricos, que não é relevante para o problema em estudo. Esta delimitação é academicamente justificável, sendo prática comum na literatura adaptar instâncias existentes ao contexto específico da investigação.

3.3 Datasets de Previsão de Procura

A componente de previsão de procura enfrenta uma limitação estrutural reconhecida na literatura: a escassez de datasets públicos específicos para serviços de apoio ao domicílio. Em resposta a esta lacuna, adotou-se uma estratégia tripartida que combina dados reais de padrões de marcação e cancelamento de consultas, dados sintéticos com estrutura de agendamento, e dados governamentais reais de prestação de cuidados domiciliários.

3.3.1 Healthcare No-Shows

O dataset Healthcare Appointment No-Shows, disponibilizado na plataforma Kaggle (2016), contém aproximadamente 110 000 registos de consultas médicas reais, recolhidos no estado do Espírito Santo (Brasil) entre abril e junho de 2016. Para cada consulta, estão disponíveis variáveis sociodemográficas do paciente (idade, género, bairro de residência), condições clínicas (hipertensão, diabetes, alcoolismo, deficiência), informação sobre a marcação (data de agendamento, data da consulta, receção de SMS de lembrete) e o resultado (comparência ou falta).

No contexto do presente projeto, este dataset é utilizado para modelar a probabilidade de cancelamento ou ausência a visitas agendadas — uma variável crítica para o módulo de re-otimização adaptativa do protótipo. A capacidade de antecipar faltas permite ajustar proativamente as rotas e a alocação de recursos, evitando deslocações desnecessárias e melhorando a eficiência operacional.

3.3.2 Medical Appointment Scheduling

Este dataset sintético, publicado em 2024, simula um sistema de agendamento de consultas médicas e inclui variáveis estruturalmente próximas das necessidades de planeamento em serviços domiciliários: identificação do paciente, tipo de serviço requerido, disponibilidade temporal, localização e frequência de utilização. A sua natureza sintética, embora represente uma limitação em termos de validade externa, confere uma vantagem prática relevante: a ausência de ruído e de valores omissos, o que facilita o desenvolvimento e teste inicial dos modelos preditivos antes da sua aplicação a dados reais.

Este dataset é utilizado como complemento ao dataset de No-Shows, permitindo explorar padrões de agendamento e frequência de visitas em condições controladas, servindo como base para a fase de desenvolvimento dos modelos de previsão de procura.

3.3.3 CMS Home Health Services

O CMS Provider Data Catalog, disponibilizado pelos Centers for Medicare & Medicaid Services do governo federal dos Estados Unidos, constitui o dataset mais diretamente relevante para a componente de previsão de procura em cuidados domiciliários. Inclui dados reais de agências de prestação de cuidados ao domicílio a nível nacional, com informação sobre o número de visitas por tipo de serviço, qualidade do atendimento, indicadores de desempenho clínico e operacional, e comparações com médias nacionais e estaduais.

A relevância deste dataset para o projeto assenta em três aspetos fundamentais: a autenticidade dos dados, provenientes de um contexto operacional real de home care; a abrangência temporal, com atualizações regulares que permitem analisar tendências e sazonalidade na procura e a credibilidade académica, decorrente da sua origem governamental e da sua utilização em investigação publicada na área da saúde domiciliária. Este dataset é utilizado para modelar padrões de procura por tipo de serviço e para calibrar os modelos preditivos desenvolvidos no módulo de previsão do protótipo.

3.4 Limitações dos Datasets face ao problema real

Embora existam fontes europeias de dados estatísticos e estudos aplicados ao contexto de cuidados domiciliários, não foi identificado nenhum dataset público europeu que disponibilize, em simultâneo, informação operacional detalhada (localização de pacientes, janelas temporais, competências e padrões de procura) com granularidade suficiente para suportar o treino e validação de modelos integrados de otimização e previsão.

As instâncias do HHCRSP 2021 são baseadas em dados de Porto Alegre (Brasil), as do Bi-Objective HHC derivam de coordenadas sintéticas do benchmark de Solomon, e as de Ankara refletem a realidade geográfica e demográfica turca. A transposição direta destes dados para o contexto nacional implicaria diferenças na densidade populacional e nos padrões de mobilidade.

Em segundo lugar, os dados de previsão de procura disponíveis publicamente não capturam especificamente visitas domiciliárias com o nível de granularidade necessário, designadamente, a duração das visitas por tipo de cuidado, as competências requeridas por paciente ou os padrões de cancelamento específicos de serviços de home care. Os datasets de No-Shows e do CMS fornecem aproximações úteis, mas não substituem dados operacionais reais de uma organização prestadora de cuidados ao domicílio.

Em terceiro lugar, nenhum dos datasets contém informação clínica identificável, o que, embora conforme com os requisitos éticos e legais do projeto, limita a possibilidade de modelar restrições clínicas individuais com elevado realismo.

Estas limitações são reconhecidas como inerentes à natureza académica do projeto e ao requisito de utilizar exclusivamente dados públicos. A avaliação do protótipo foi realizada em cenários simulados, sendo os resultados interpretados com as devidas ressalvas quanto à generalização para contextos operacionais reais.

3.5 Estratégias de Utilização

Os seis datasets articulam-se com os dois módulos principais do protótipo de forma complementar e sequencial, conforme ilustrado na estratégia de utilização definida abaixo.

No módulo de otimização, o HHCRSP Dataset 2021 foi utilizado como conjunto de instâncias de referência para o desenvolvimento e avaliação dos algoritmos de otimização de rotas e alocação de recursos. O Bi-Objective HHC Dataset permitiu testar a robustez dos algoritmos em cenários com incerteza nos tempos de serviço, enquanto o dataset de Ankara contribuiu com a dimensão de heterogeneidade de competências profissionais e dados geográficos realistas.

No módulo de previsão de procura, o dataset de No-Shows foi utilizado para treinar modelos de previsão de cancelamentos, alimentando o mecanismo de re-otimização adaptativa. O Medical Appointment Scheduling Dataset, de natureza sintética, serviu como ambiente controlado para desenvolvimento e teste inicial dos modelos. O CMS Home Health Dataset forneceu dados reais de volume e distribuição de visitas domiciliárias, permitindo calibrar os modelos preditivos em padrões operacionais autênticos.

A integração dos dois módulos foi materializada num ciclo de planeamento em que as previsões de procura e cancelamento alimentam dinamicamente os algoritmos de otimização, possibilitando a re-otimização periódica das rotas e escalas em resposta a variações da procura ou da disponibilidade de recursos — estratégia designada na literatura por *rolling horizon planning* (Oliveira, 2022).

4 Análise de Requisitos

Neste capítulo são identificados e documentados formalmente os requisitos do sistema proposto.

A análise de requisitos constitui uma etapa fundamental no processo de engenharia de software, permitindo delimitar o âmbito do protótipo, orientar as decisões de design e estabelecer critérios mensuráveis para a avaliação da solução desenvolvida. Os requisitos identificados foram definidos tendo em conta as características e limitações dos datasets descritos no Capítulo 3, assegurando coerência entre os dados disponíveis, os objetivos do projeto e as funcionalidades propostas.

Os requisitos encontram-se organizados em quatro categorias: requisitos funcionais (RF), que especificam o que o sistema deve fazer; requisitos não funcionais (RNF), que definem as propriedades de qualidade e conformidade do sistema; casos de uso (UC), que descrevem cenários típicos de interação com o protótipo; e pressupostos e restrições, que delimitam explicitamente o âmbito e as condições de operação do sistema.

4.1 Identificação dos Stakeholders

A identificação dos stakeholders precede o levantamento de requisitos, na medida em que os diferentes grupos de partes interessadas apresentam perspetivas, necessidades e restrições distintas que influenciam diretamente os requisitos do sistema. Com base na caracterização do problema e nos objetivos do projeto, identificam-se quatro grupos principais de stakeholders:

- **Profissionais de saúde e cuidadores** — constituem o grupo operacional mais afetado pelo sistema. O seu interesse centra-se na qualidade e equilíbrio das rotas atribuídas, na clareza da informação disponibilizada e na capacidade do sistema de reagir a imprevistos sem gerar sobrecarga.

- **Gestores operacionais** — são os utilizadores primários do protótipo. O seu foco incide na eficiência do planeamento, na redução do tempo despendido em tarefas manuais e na disponibilidade de métricas de desempenho que suportem a tomada de decisão.
- **Pacientes e utentes** — são os beneficiários finais do serviço. O seu interesse reflete-se na pontualidade e na qualidade dos cuidados prestados, dependentes da eficácia do planeamento.
- **Entidades prestadoras de cuidados domiciliários** — visam a redução de custos operacionais, a melhoria da qualidade do serviço prestado e a sustentabilidade do modelo de prestação de cuidados.

No contexto do presente projeto, de natureza exclusivamente académica e baseado em dados públicos e cenários simulados, os gestores operacionais e o investigador/avaliador constituem os principais atores diretamente envolvidos na interação com o protótipo.

4.2 Requisitos Funcionais

Os requisitos funcionais especificam as funcionalidades que o sistema deve disponibilizar, organizadas de acordo com os três módulos principais do protótipo: o módulo de previsão de procura, o módulo de otimização de rotas e alocação de recursos e o módulo de integração e re-otimização adaptativa. A prioridade de cada requisito foi definida com base na sua relevância para os objetivos centrais do projeto, sendo classificada como Alta, Média ou Baixa.

Tabela 2 - Requisitos funcionais do sistema

ID	Designação	Descrição	Módulo	Prioridade
RF01	Importação e pré-processamento de dados.	O sistema deve importar e pré-processar dados de pacientes, profissionais e janelas temporais a partir de datasets externos.	Dados	Alta
RF02	Previsão de padrões de procura.	O sistema deve prever padrões de procura por serviços domiciliários com base em dados históricos, utilizando modelos de aprendizagem automática e respetiva validação do desempenho preditivo.	Previsão	Alta
RF03	Previsão de cancelamentos.	O sistema deve prever a probabilidade de cancelamento ou ausência de visitas agendadas.	Previsão	Alta

ID	Designação	Descrição	Módulo	Prioridade
RF04	Geração de rotas otimizadas.	O sistema deve gerar planos diários de rotas otimizadas para um conjunto de profissionais, minimizando o tempo total de deslocação e/ou outros critérios operacionais definidos.	Otimização	Alta
RF05	Alocação de profissionais baseada em competências.	O sistema deve alocar profissionais a visitas respeitando as suas competências e disponibilidade horária.	Otimização	Alta
RF06	Gestão de janelas temporais.	O sistema deve respeitar as janelas temporais de atendimento definidas para cada paciente.	Otimização	Alta
RF07	Equilíbrio de carga de trabalho.	O sistema deve minimizar o desvio-padrão da carga de trabalho entre os profissionais disponíveis.	Otimização	Alta
RF08	Incorporação de prioridades clínicas.	O sistema deve incorporar prioridades clínicas na ordenação e alocação das visitas, atribuindo penalizações mais elevadas ao incumprimento de visitas prioritárias.	Otimização	Alta
RF09	Re-otimização (rolling horizon).	O sistema deve suportar re-otimização periódica das rotas com base em previsões de procura atualizadas (rolling horizon).	Integração	Alta
RF10	Ajuste dinâmico do planeamento.	O sistema deve ajustar o planeamento em resposta a cancelamentos ou alterações de disponibilidade de profissionais.	Integração	Alta

ID	Designação	Descrição	Módulo	Prioridade
RF11	Comparação com cenário de referência (baseline).	O sistema deve avaliar e comparar o desempenho das soluções geradas face a um cenário de planeamento estático sem previsão de procura nem re-otimização adaptativa.	Avaliação	Média
RF12	Produção de relatórios de desempenho.	O sistema deve produzir relatórios de desempenho com métricas de eficiência operacional, incluindo métricas como tempo de deslocação, equilíbrio de carga e taxa de cancelamentos.	Avaliação	Média

Os requisitos definidos dizem respeito a um protótipo de investigação, não tendo como objetivo a implementação de um sistema operacional completo.

4.3 Requisitos Não Funcionais

Os requisitos não funcionais definem as propriedades de qualidade, conformidade e operação do sistema, complementando os requisitos funcionais com critérios que não se traduzem diretamente em funcionalidades, mas que condicionam o design e a implementação do protótipo. São consideradas oito categorias: desempenho, privacidade e conformidade com o RGPD, reprodutibilidade, manutenibilidade, escalabilidade e transparência algorítmica.

Tabela 3 - Requisitos não funcionais do sistema

ID	Categoria	Descrição
RNF01	Desempenho	O módulo de otimização deve gerar uma solução viável para instâncias de até 100 pacientes em tempo computacionalmente aceitável (inferior a 5 minutos).
RNF02	Desempenho	O módulo de previsão deve produzir resultados com erro médio absoluto (MAE) e erro quadrático médio (RMSE) avaliados e documentados para cada modelo desenvolvido comparando diferentes modelos de previsão.

ID	Categoria	Descrição
RNF03	Privacidade e RGPD	O sistema não deve processar dados pessoais ou clínicos reais identificáveis. Todos os dados utilizados devem ser públicos ou sintéticos, em conformidade com o RGPD.
RNF04	Privacidade e RGPD	O sistema deve aplicar os princípios de minimização de dados e de finalidade, não recolhendo nem retendo informação além do estritamente necessário para o planeamento. (RGPD, Art. 5.º)
RNF05	Reprodutibilidade	O código e os datasets utilizados devem estar documentados e ser reprodutíveis, permitindo a validação independente dos resultados.
RNF06	Manutenibilidade	O protótipo deve ser desenvolvido em Python, com estrutura modular que permita substituir ou atualizar independentemente do módulo de previsão e do módulo de otimização.
RNF07	Escalabilidade	A arquitetura do protótipo deve suportar a extensão para instâncias maiores e a incorporação de novos tipos de restrições sem necessidade de reformulação estrutural.
RNF08	Transparência	Os modelos de IA desenvolvidos devem ser documentados de forma a permitir a interpretação dos seus resultados por utilizadores não especialistas, em conformidade com as diretrizes da WHO (2021) sobre transparência algorítmica, recorrendo a técnicas como feature importance e métodos de explicabilidade simples dos modelos.

Os requisitos de privacidade (RNF03 e RNF04) assumem particular relevância no contexto de um sistema aplicado à saúde, mesmo quando os dados utilizados são públicos ou sintéticos. A adoção dos princípios do RGPD no design do sistema constitui simultaneamente uma obrigação legal e uma boa prática de engenharia, preparando o protótipo para uma eventual transição para contextos operacionais reais com dados sensíveis.

O requisito de transparência algorítmica (RNF08) fundamenta-se nas diretrizes da Organização Mundial da Saúde (2021) para sistemas de IA em contextos de saúde, que estabelecem a interpretabilidade e a responsabilização como princípios fundamentais para a adoção segura de sistemas inteligentes em contextos clínicos e operacionais.

4.4 Casos de Uso

Os casos de uso descrevem os cenários típicos de interação entre os atores identificados e o sistema, permitindo contextualizar os requisitos funcionais em situações concretas de utilização. No contexto do presente projeto, de natureza académica e baseado em dados simulados, os casos de uso representam os cenários previstos para a validação experimental da solução.

Tabela 4 - Casos de uso do Sistema

ID	Ator	Caso de Uso	Evento Desencadeador	Resultado Esperado
UC01	Gestor Operacional	Gerar plano diário de rotas	Início do dia operacional ou pedido manual.	Plano de rotas otimizadas gerado e disponível para consulta.
UC02	Gestor Operacional	Re-otimizar plano por eventos operacionais (cancelamentos ou alterações de disponibilidade)	Cancelamento, ausência de visita ou alteração de disponibilidade de profissionais.	Plano ajustado com redistribuição das visitas afetadas.
UC03	Gestor Operacional	Consultar relatório de desempenho	Fim do ciclo de planeamento.	Relatório com métricas de eficiência, carga e qualidade do serviço.
UC04	Sistema (automático)	Prever procura para o período seguinte	Execução periódica agendada (rolling horizon).	Previsão de procura atualizada disponível para o módulo de otimização.

ID	Ator	Caso de Uso	Evento Desencadeador	Resultado Esperado
UC05	Sistema (automático)	Prever probabilidade de cancelamento	Execução antes da geração do plano diário.	Lista de visitas com probabilidade de falta associada.
UC06	Investigador / Avaliador	Comparar solução proposta com baseline	Execução de cenário experimental.	Resultados comparativos documentados com métricas MAE, RMSE, tempo de deslocação e equilíbrio de carga.
UC07	Investigador / Avaliador	Configurar cenário experimental	Início de execução de testes ou simulação.	Parâmetros do cenário definidos (nº de pacientes, parâmetros, horizonte temporal, etc.).

Os casos de uso UC01 e UC02 representam os cenários centrais de utilização pelo gestor operacional, cobrindo o fluxo normal de geração do plano diário e o tratamento de exceções decorrentes de cancelamentos. Os casos UC04 e UC05, de execução automática pelo sistema, materializam o mecanismo de *rolling horizon*, que distingue a abordagem proposta das soluções estáticas identificadas na literatura. O caso UC06 é específico do contexto de avaliação académica, permitindo a comparação sistemática da solução proposta com um *baseline* de referência. Adicionalmente, o caso de uso UC07 permite a parametrização dos testes, incluindo o número de pacientes, o horizonte temporal e os parâmetros operacionais do sistema, assegurando a reprodutibilidade e controlo dos cenários de avaliação.

O diagrama de casos de uso, representando as relações entre atores e funcionalidades do sistema, é apresentado na secção 5.9.

4.5 Pressupostos e Restrições

A delimitação explícita dos pressupostos e restrições do sistema é essencial para evitar ambiguidades no âmbito do projeto e orientar as decisões de design. Os pressupostos definem

as condições assumidas como verdadeiras para o funcionamento do sistema, enquanto as restrições estabelecem os limites dentro dos quais o protótipo opera.

Tabela 5 - Pressupostos e restrições do sistema

ID	Pressuposto / Restrição
P01	O protótipo é avaliado exclusivamente em cenários simulados, sem integração com sistemas operacionais reais.
P02	Os dados utilizados são públicos e não contêm informação pessoal identificável.
P03	O módulo de otimização baseia-se em variantes do VRP, não incluindo outros paradigmas de otimização como programação por restrições.
P04	O protótipo é desenvolvido em Python, sem requisito de interface gráfica de utilizador.
P05	A re-otimização é periódica (rolling horizon), não em tempo real contínuo.
P06	O sistema não gere autenticação de utilizadores nem controlo de acessos.
P07	A avaliação de desempenho é realizada por comparação com um baseline de planeamento estático, sem envolvimento de utilizadores reais.
P08	Os datasets de previsão e otimização são independentes, sendo a integração realizada através de cenários simulados.

Os pressupostos P01 e P02 decorrem diretamente das considerações éticas e legais do projeto, sendo inegociáveis no contexto académico em que a dissertação se insere. Os pressupostos P03 a P08 refletem decisões de âmbito adotadas para garantir a viabilidade técnica e temporal do projeto, reconhecendo que um protótipo de investigação não tem como objetivo cobrir a totalidade dos requisitos de um sistema em produção.

4.6 Matriz de Rastreabilidade

A matriz de rastreabilidade estabelece a correspondência entre os requisitos funcionais definidos e os objetivos específicos do projeto, permitindo verificar que cada objetivo está coberto por pelo menos um requisito e que cada requisito se justifica por pelo menos um objetivo.

Tabela 6 - Matriz de rastreabilidade: Requisitos funcionais vs. Objetivos

RF	Objetivo Específico Associado	Caso de Uso Relacionado
RF01	Recolher e estruturar dados de localização, disponibilidade e competências dos profissionais, e necessidades dos pacientes.	UC04, UC05
RF02	Aplicar técnicas de Aprendizagem Automática para previsão da procura.	UC04
RF03	Aplicar técnicas de Aprendizagem Automática para previsão de cancelamentos ou ausência de visitas.	UC05
RF04	Aplicar algoritmos de otimização (VRP) para planeamento de rotas.	UC01, UC02
RF05	Aplicar algoritmos de otimização para alocação de recursos humanos.	UC01, UC02
RF06	Aplicar algoritmos de otimização incorporando restrições operacionais.	UC01
RF07	Avaliar o desempenho em termos de equilíbrio de carga de trabalho.	UC01, UC03
RF08	Aplicar algoritmos de otimização incorporando prioridades clínicas.	UC01, UC02
RF09	Desenvolver um protótipo que integre previsão e planeamento operacional.	UC04, UC01
RF10	Desenvolver um protótipo com re-otimização adaptativa.	UC02
RF11	Avaliar o desempenho do modelo em cenários simulados.	UC06
RF12	Disponibilizar evidências sobre ganhos operacionais.	UC03, UC06

A análise da matriz confirma que todos os objetivos específicos definidos na formalização do projeto se encontram cobertos por pelo menos um requisito funcional e que todos os requisitos funcionais se articulam com pelo menos um caso de uso.

5 Desenho da Solução e Arquitetura

Neste capítulo é apresentado o desenho do protótipo desenvolvido, integrando as tecnologias selecionadas no Capítulo 2, os datasets caracterizados no Capítulo 3 e os requisitos definidos no Capítulo 4. O desenho da solução constitui a ponte entre a fase de análise e a fase de implementação, definindo a arquitetura do sistema, o fluxo de dados entre módulos, as decisões de concepção adotadas e as alternativas consideradas.

A solução é estruturada em cinco módulos funcionais distintos, organizados numa sequência de processamento com um mecanismo de retroalimentação adaptativa. Esta organização modular responde ao requisito de manutenibilidade (RNF06) e facilita a substituição ou atualização independente de cada componente.

5.1 Visão Geral da Arquitetura

A arquitetura do protótipo segue um modelo modular sequencial, no qual os dados fluem entre cinco módulos principais. O Módulo 1 é responsável pela ingestão e pré-processamento dos dados; o Módulo 2 produz previsões de procura e de cancelamentos; o Módulo 3 gera planos de rotas otimizadas; o Módulo 4 realiza a re-otimização adaptativa do plano em resposta a eventos e previsões atualizadas; e o Módulo 5 avalia e compara os resultados com um cenário de referência estático.

O mecanismo de *rolling horizon*, implementado no Módulo 4, introduz um ciclo de retroalimentação entre os módulos de previsão e de otimização, permitindo que o sistema se adapte dinamicamente a variações da procura ou a cancelamentos de visitas, distinguindo a abordagem proposta das soluções estáticas identificadas no estado da arte (Fikar & Hirsch, 2017; Algendi et al., 2024).

A Figura 1 apresenta o diagrama da arquitetura geral do sistema, com indicação das tecnologias associadas a cada módulo.

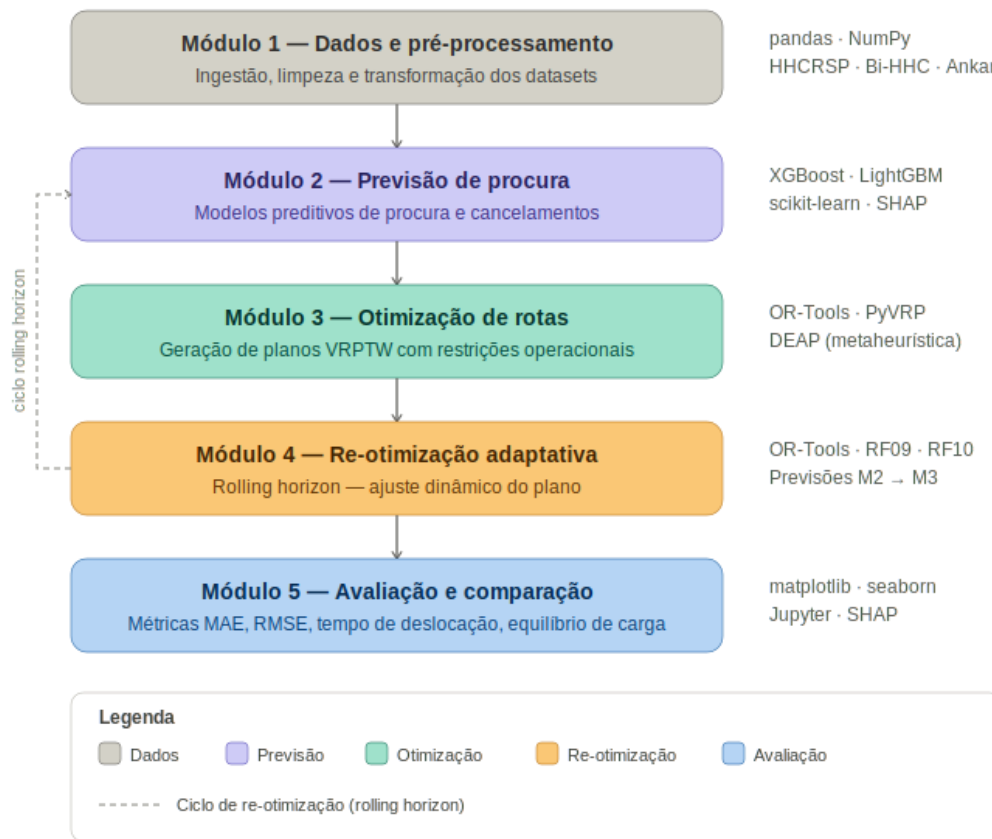


Figura 1 – Arquitetura geral do sistema de otimização de rotas em home care

A Tabela 7 resume as entradas, saídas, tecnologias e requisitos funcionais cobertos por cada módulo, proporcionando uma visão integrada da arquitetura.

Tabela 7 - Resumo dos módulos da arquitetura

Módulo	Entrada	Saída	Tecnologias	RF cobertos
M1 — Pré-processamento	Datasets públicos (CSV, JSON).	Features normalizadas e codificadas.	pandas, NumPy.	RF01
M2 — Previsão	Features tratadas (M1).	Procura prevista + prob. Cancelamento.	XGBoost, LightGBM, scikit-learn, SHAP.	RF02, RF03
M3 — Otimização	Procura prevista (M2) + instâncias.	Plano de rotas VRPTW.	OR-Tools, PyVRP, DEAP.	RF04–RF08

Módulo	Entrada	Saída	Tecnologias	RF cobertos
M4 — Re-otimização	Plano (M3) + prob. cancelamento (M2).	Plano ajustado dinamicamente.	OR-Tools.	RF09, RF10
M5 — Avaliação	Plano final (M4) + baseline estático.	Relatório de desempenho e métricas.	matplotlib, seaborn, Jupyter, SHAP.	RF11, RF12

5.2 Módulo 1 – Dados e Pré-processamento

O Módulo 1 é responsável pela ingestão e preparação dos seis datasets selecionados, transformando-os em representações adequadas para os módulos subsequentes. Este módulo segue as etapas iniciais da metodologia CRISP-DM (compreensão e preparação dos dados) e constitui a base sobre a qual assentam os módulos de previsão e de otimização.

As principais operações realizadas neste módulo são:

- **Limpeza e tratamento de valores omissos** — identificação e imputação de valores em falta nos datasets de previsão (No-Shows e CMS), recorrendo a estratégias adequadas ao tipo de variável (mediana para variáveis numéricas e moda para variáveis categóricas).
- **Codificação de variáveis categóricas** — transformação de variáveis nominais (tipo de serviço, condição clínica) em representações numéricas compatíveis com os modelos de aprendizagem automática, utilizando *one-hot encoding* ou *label encoding*, conforme o contexto.
- **Normalização e escalonamento** — normalização de variáveis numéricas contínuas (idade, tempo de deslocação e distância), de forma a evitar que diferenças de escala influenciem indevidamente o desempenho dos modelos preditivos.
- **Transformação dos datasets de otimização** — conversão das instâncias HHCRSP, Bi-Objective HHC e Ankara para o formato de entrada exigido pelo OR-Tools, incluindo a construção de matrizes de distâncias, definição de janelas temporais e modelação das restrições de competências.

A separação deste módulo dos restantes assegura que eventuais alterações nos datasets ou nos procedimentos de pré-processamento podem ser efetuadas sem impacto nos módulos subsequentes, em conformidade com o requisito de manutenibilidade (RNF06).

5.3 Módulo 2 – Previsão de Procura e de Cancelamentos

O Módulo 2 implementa os modelos preditivos que suportam o ciclo de planeamento adaptativo do protótipo. Produz dois tipos de resultados distintos, correspondentes a dois problemas de previsão complementares: a previsão de procura por tipo de serviço (para alimentar o Módulo 3) e a previsão da probabilidade de cancelamento de visitas agendadas (para alimentar o Módulo 4).

5.3.1 Previsão de procura

A previsão de procura é formulada como um problema de regressão supervisionada, no qual o modelo estima o volume esperado de visitas domiciliárias por tipo de serviço e por período temporal. Os datasets CMS Home Health e Medical Appointment Scheduling fornecem os dados históricos de volume de visitas utilizados no treino.

O pipeline CRISP-DM aplicado inclui a seleção de variáveis relevantes, o treino de modelos baseados em XGBoost (Chen e Guestrin, 2016) e LightGBM (Ke et al., 2017), a avaliação através de validação cruzada utilizando as métricas MAE e RMSE, e a seleção do modelo com melhor capacidade de generalização.

5.3.2 Previsão de cancelamentos

A previsão de cancelamentos é formulada como um problema de classificação binária, no qual o modelo estima a probabilidade de ausência para cada visita agendada. O dataset Healthcare No-Shows fornece os dados de treino.

O resultado deste submodelo, uma probabilidade associada a cada visita, é utilizado pelo Módulo 4 para priorizar a re-otimização das visitas com maior risco de cancelamento.

A interpretabilidade de ambos os modelos são assegurados pela integração da biblioteca SHAP (Lundberg e Lee, 2017), que permite quantificar a contribuição de cada variável de entrada para as previsões individuais, em conformidade com o requisito não funcional RNF08.

5.4 Módulo 3 – Otimização de Rotas

O Módulo 3 constitui o núcleo computacional do protótipo, sendo responsável pela geração de planos de rotas otimizadas para o conjunto de profissionais disponíveis. O problema é modelado como um *Vehicle Routing Problem with Time Windows* (VRPTW), incorporando as restrições operacionais identificadas nos requisitos RF04 a RF08.

A formulação do problema inclui os seguintes elementos:

- **Conjunto de nós** — ponto de partida e de chegada (base de operações) e conjunto de pacientes a visitar, com coordenadas geográficas extraídas dos datasets de otimização.
- **Janelas temporais** — intervalos de atendimento definidos por paciente, derivados dos datasets HHCRSP e Bi-Objective HHC. Visitas realizadas fora da janela temporal incorrem em penalizações.
- **Restrições de competências** — cada profissional possui um perfil de competências; cada visita requer um nível mínimo de qualificação, conforme o modelo do dataset de Ankara.
- **Função objetivo** — minimização do tempo total de deslocação, assegurando simultaneamente o cumprimento das janelas temporais (RF06). O equilíbrio da carga de trabalho é modelado através de penalizações adicionadas à função objetivo principal, permitindo incorporar este critério na função de otimização de forma agregada (RF07).

São utilizados três algoritmos de otimização em regime comparativo: o OR-Tools (Perron & Furnon, 2023) como *solver* principal com *Guided Local Search*, o PyVRP (Wouda, Lan & Kool, 2024) como alternativa de estado da arte, e o DEAP para a implementação de algoritmos genéticos. Os resultados obtidos por cada abordagem são avaliados no Módulo 5, permitindo a análise de compromissos entre qualidade da solução e tempo de execução.

5.4.1 Função Objetivo

Formalmente, o problema VRPTW pode ser definido sobre um grafo direcionado $G = (V, E)$, onde $V = \{0, 1, \dots, n\}$ representa o conjunto de nós, sendo o nó 0 o depósito e os restantes nós correspondentes às visitas a realizar. O conjunto K representa os profissionais disponíveis.

Define-se a variável de decisão binária x_{ijk} , que assume o valor 1 se o profissional $k \in K$ viaja diretamente do nó $i \in V$ para o nó $j \in V$, e 0 caso contrário:

$$x_{ijk} = \begin{cases} 1, & \text{se o profissional } k \text{ viaja diretamente do nó } i \text{ para } j \\ 0, & \text{caso contrário} \end{cases}$$

Define-se ainda a variável contínua $t_i \geq 0$, que representa o instante de início de atendimento no nó i .

A função objetivo do problema consiste na minimização de um critério agregado que combina eficiência operacional, cumprimento de janelas temporais e equilíbrio da carga de trabalho entre profissionais:

$$\min Z = \alpha_1 \sum_{k \in K} \sum_{i \in V} \sum_{j \in V} d_{ij} x_{ijk} + \alpha_2 \sum_{i \in V} P_i^{tw} + \alpha_3 \sigma(L_k)$$

onde:

- d_{ij} representa o tempo ou distância de deslocação entre os nós i e j ;
- x_{ijk} indica se o profissional k percorre o arco $i \rightarrow j$;

- p_i^{tw} corresponde à penalização associada a eventuais violações das janelas temporais do nó i ;
- L_k representa a carga de trabalho atribuída ao profissional k ;
- $\sigma(L_k)$ corresponde ao desvio padrão das cargas de trabalho, utilizado como medida de equilíbrio entre profissionais;
- a_1, a_2, a_3 são parâmetros de ponderação definidos empiricamente.

Restrições do problema

1. Visita única a cada nó

Cada nó (exceto o depósito) deve ser visitado exatamente uma vez:

$$\sum_{k \in K} \sum_{j \in V} x_{ijk} = 1, \quad \forall i \in V \setminus \{0\}$$

2. Conservação de fluxo

Para cada profissional, garante-se a conservação do fluxo de entrada e saída em cada nó:

$$\sum_{j \in V} x_{ijk} = \sum_{j \in V} x_{jik}, \quad \forall i \in V, \forall k \in K$$

3. Janelas temporais

Cada visita deve ocorrer dentro do intervalo temporal admissível:

$$a_i \leq t_i \leq b_i, \quad \forall i \in V$$

onde a_i e b_i representam, respetivamente, o início e o fim da janela temporal associada ao nó i .

4. Consistência temporal das rotas

O início do atendimento no nó seguinte deve respeitar os tempos de deslocação e de serviço:

$$t_j \geq t_i + s_i + d_{ij} - M(1 - x_{ijk})$$

onde:

- s_i corresponde ao tempo de serviço no nó i ;
- M é uma constante suficientemente grande (*big-M*), utilizada para linearizar a restrição.

5.5 Módulo 4 – Re-otimização Adaptativa

O Módulo 4 implementa o mecanismo de *rolling horizon*, que distingue a abordagem proposta das soluções estáticas identificadas na literatura. Este módulo utiliza o plano de rotas gerado

pelo Módulo 3 e as previsões de cancelamento produzidas pelo Módulo 2, ajustando dinamicamente o planeamento em resposta a dois tipos de eventos:

- **Cancelamentos previstos** — visitas com probabilidade de cancelamento superior a um limiar configurável são reagendadas preventivamente ou redistribuídas por outros profissionais com capacidade disponível.
- **Alterações de disponibilidade** — em caso de indisponibilidade de um profissional, o módulo re-otimiza a alocação das visitas afetadas, recorrendo aos recursos disponíveis e minimizando o impacto na qualidade do serviço.

O ciclo de *rolling horizon* opera com uma janela temporal configurável, tipicamente correspondente a um período de planeamento diário, com re-otimizações periódicas ao longo desse período. Esta abordagem permite equilibrar a capacidade de adaptação do sistema com o custo computacional associado à re-otimização, em conformidade com o pressuposto P05.

5.6 Módulo 5 – Avaliação e Comparação

O Módulo 5 é responsável pela avaliação experimental do protótipo, comparando o desempenho da solução proposta com um cenário de referência de planeamento estático — correspondente à geração de um único plano diário sem re-otimização adaptativa. Esta comparação é essencial para demonstrar os ganhos operacionais da abordagem integrada, conforme definido nos objetivos do projeto.

As métricas de avaliação utilizadas são:

- **MAE e RMSE** — avaliação do desempenho dos modelos preditivos do Módulo 2.
- **Tempo total de deslocação** — distância e tempo percorridos pelo conjunto de profissionais durante o período de planeamento.
- **Equilíbrio de carga** — desvio padrão do número de visitas atribuídas a cada profissional, como indicador de equidade na distribuição do trabalho.
- **Taxa de cumprimento de janelas temporais** — percentagem de visitas realizadas dentro do intervalo temporal definido pelo paciente.
- **Custo operacional estimado** — indicador aproximado do custo de deslocação, calculado com base no tempo total de viagem.
- **Ganho percentual face ao planeamento estático** — melhoria relativa das principais métricas de desempenho em relação ao planeamento estático, permitindo quantificar de forma direta os benefícios da abordagem proposta.

Os resultados são visualizados com as bibliotecas matplotlib e seaborn, e documentados em Jupyter Notebook, garantindo a reprodutibilidade integral das experiências, em conformidade com o requisito RNF05.

5.7 Fluxo de Dados entre Módulos

A Figura 2 apresenta o fluxo de dados detalhado entre os cinco módulos do protótipo, evidenciando as entradas e saídas de cada módulo, bem como os mecanismos de retroalimentação associados ao ciclo de *rolling horizon*.

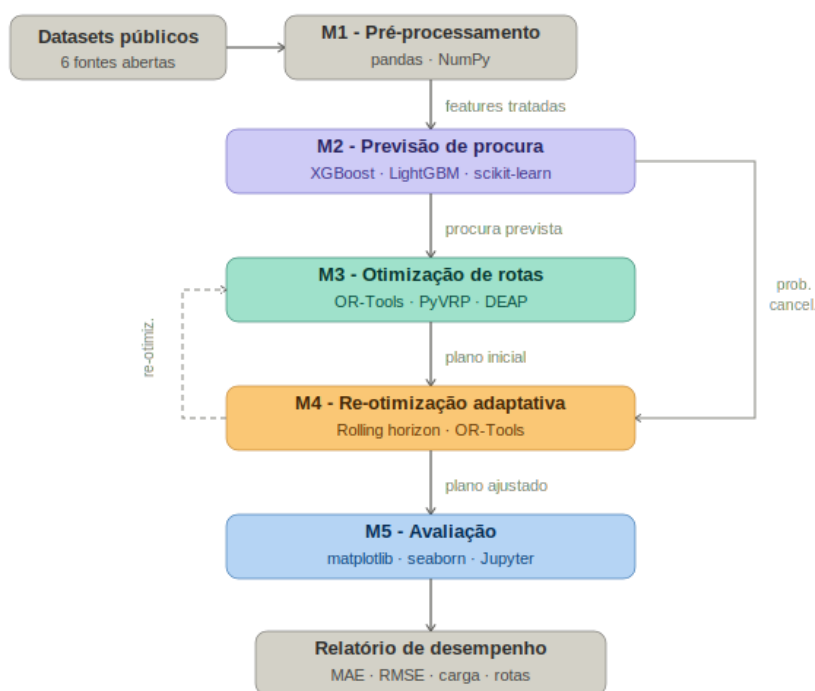


Figura 2 - Fluxo de dados entre os módulos do protótipo

O fluxo principal é sequencial e descendente: os datasets públicos são processados pelo Módulo 1, sendo preparados dois tipos distintos de dados: os datasets de previsão, que alimentam o Módulo 2 (previsão), e os datasets de otimização, que fornecem as instâncias utilizadas pelo Módulo 3 (otimização de rotas). O Módulo 2 produz dois resultados distintos: a procura prevista, que alimenta o Módulo 3, e a probabilidade de cancelamento, que alimenta diretamente o Módulo 4. O Módulo 3 gera o plano inicial de rotas, o qual é posteriormente ajustado pelo Módulo 4 através do ciclo de *rolling horizon*. O plano final é avaliado pelo Módulo 5, que produz o relatório de desempenho.

O mecanismo de retroalimentação entre o Módulo 4 e o Módulo 3 — representado pela seta tracejada no diagrama — materializa o ciclo de re-otimização adaptativa, no qual o plano ajustado pode ser submetido a uma nova iteração de otimização, caso as alterações sejam suficientemente significativas para justificar um recálculo completo das rotas.

5.8 Decisões de Design e Alternativas Consideradas

O processo de concepção do protótipo envolveu um conjunto de decisões técnicas para as quais foram analisadas diferentes alternativas antes da definição da solução final. A Tabela 8 apresenta as principais decisões, as opções selecionadas, as alternativas consideradas e a respetiva justificação.

Tabela 8 - Decisões de design e alternativas consideradas

Decisão	Opção escolhida	Alternativas consideradas	Justificação
Solver principal VRP	OR-Tools (Google).	PyVRP, DEAP isolado, CPLEX.	Open-source; suporte nativo a VRPTW;
Modelo preditivo principal	XGBoost.	LightGBM, Random Forest, LSTM.	Melhor interpretabilidade em datasets médios; integração SHAP; RNF08.
Estratégia de re-otimização	Rolling horizon periódico.	Re-otimização contínua, estática.	Equilíbrio entre adaptabilidade e custo computacional; P05.
Metaheurística de comparação	Algoritmos Genéticos (DEAP).	Simulated Annealing, Tabu Search.	Ecossistema Python maduro.
Ambiente de desenvolvimento	Jupyter Notebook.	Scripts Python puros, R Markdown.	Reprodutibilidade; combinação de código, visualização e texto; RNF05.

A decisão mais crítica do ponto de vista arquitetural foi a adoção do mecanismo de *rolling horizon* para a re-otimização, em detrimento de uma abordagem de re-otimização contínua em tempo real. Esta escolha fundamenta-se em dois argumentos principais.

Em primeiro lugar, a re-otimização contínua exigiria um *solver* capaz de produzir soluções viáveis num intervalo de tempo muito reduzido, o que excede o âmbito de um protótipo académico com recursos computacionais limitados.

Em segundo lugar, na literatura de planeamento e escalonamento em cuidados ao domicílio, o *rolling horizon* constitui uma abordagem amplamente utilizada para lidar com procura dinâmica

e incerteza, permitindo atualizar periodicamente o plano com base na informação mais recente, mantendo simultaneamente a eficiência computacional.

Assim, esta opção representa um compromisso adequado entre adaptabilidade do sistema e custo computacional (Fikar & Hirsch, 2017).

5.9 Diagrama de Casos de Uso

A Figura 3 apresenta o diagrama UML de casos de uso do sistema, representando as relações entre os três atores identificados e os seis casos de uso definidos anteriormente.

O Gestor Operacional interage diretamente com o UC01 (Gerar plano diário de rotas), UC02 (Re-otimizar plano em resposta a eventos operacionais) e UC03 (Consultar relatório de desempenho).

O ator Sistema (execução automática) é responsável pela execução dos casos UC04 (Prever procura para o período seguinte) e UC05 (Prever probabilidade de cancelamento), que constituem os gatilhos automáticos do ciclo de rolling horizon.

O Investigador/Avaliador interage com o UC06 (Comparar solução proposta com o cenário de referência), destinado à avaliação experimental das soluções produzidas, e com o UC07 (Configurar cenário experimental), que permite definir os parâmetros utilizados nos ensaios, tais como o número de pacientes, a instância de teste, o horizonte temporal e outros parâmetros de configuração necessários à execução das experiências.

As relações “include” representam dependências obrigatórias: UC01 inclui UC04 — a geração do plano diário requer sempre uma previsão de procura atualizada; e UC02 inclui UC05 — a re-otimização em resposta a cancelamentos requer a estimativa da probabilidade de ausência associada a cada visita afetada.

6 Implementação

Este capítulo descreve de forma detalhada o plano de implementação do protótipo, apresentando, módulo a módulo, as decisões técnicas, os algoritmos selecionados, a estrutura do código e exemplos de pseudocódigo e código Python ilustrativos dos principais componentes do sistema.

A implementação foi conduzida seguindo a metodologia CRISP-DM Data Science Central para a componente de aprendizagem automática e princípios de engenharia de software para a componente de otimização, garantindo modularidade, reprodutibilidade e rastreabilidade dos resultados.

6.1 Ambiente de Desenvolvimento e Configuração

O ambiente de desenvolvimento do protótipo é baseado em Python e Jupyter Notebook, conforme definido no Capítulo 2.7. A configuração do ambiente é integralmente reprodutível através de um ficheiro requirements.txt, garantindo o cumprimento do requisito RNF05 relativo à reprodutibilidade dos testes.

6.1.1 Estrutura de pastas de projeto

A organização do projeto segue uma estrutura modular, separando os datasets, os módulos funcionais do sistema, os notebooks de análise e avaliação, e os resultados produzidos durante a fase de testes. Esta separação facilita a manutenção, a reutilização de componentes e a reprodutibilidade das avaliações realizadas, em conformidade com os requisitos RNF05 e RNF06.

```
home_care_optimizer/  
├─ data/  
│   └─ raw/                                # Datasets originais descarregados
```

```

|   └─ processed/          # Datasets após pré-processamento (M1)
|   └─ instances/         # Instâncias VRP prontas para OR-Tools
└─ modules/
|   └─ m1_preprocessing.py
|   └─ m2_forecasting.py
|   └─ m3_optimization.py
|   └─ m4_reoptimization.py
|   └─ m5_evaluation.py
└─ notebooks/
|   └─ 01_exploratory_analysis.ipynb
|   └─ 02_forecasting_experiments.ipynb
|   └─ 03_optimization_experiments.ipynb
|   └─ 04_reoptimization_experiments.ipynb
|   └─ 05_evaluation_results.ipynb
└─ results/              # Métricas, gráficos e relatórios
└─ requirements.txt
└─ main.py               # Ponto de entrada do pipeline completo

```

6.1.2 Dependências e instalação

A implementação do protótipo depende de um conjunto de bibliotecas Python previamente identificadas. A Tabela 9 apresenta as principais dependências do projeto, respectivas versões mínimas e finalidade no contexto do sistema desenvolvido.

A gestão das dependências é realizada através de um ficheiro `requirements.txt`, permitindo reproduzir integralmente o ambiente de desenvolvimento em diferentes sistemas. A instalação das bibliotecas é efetuada com o seguinte comando:

```
pip install -r requirements.txt
```

A utilização de versões mínimas explícitas garante a compatibilidade entre módulos, a estabilidade das experiências e a reprodutibilidade dos resultados obtidos ao longo do desenvolvimento e da avaliação do protótipo.

Tabela 9 - Dependências do projeto e comandos de instalação

Biblioteca	Versão instalada	Instalação (pip)	Módulo(s) utilizador
pandas	3.0.3	pip install pandas	M1
NumPy	2.4.4	pip install numpy	M1, M3
scikit-learn	1.8.0	pip install scikit-learn	M2
XGBoost	3.2.0	pip install xgboost	M2
LightGBM	4.6.0	pip install lightgbm	M2
SHAP	0.51.0	pip install shap	M2, M5
ortools	9.15.6755	pip install ortools	M3, M4
pyvrp	0.13.3	pip install pyvrp	M3
deap	1.4.4	pip install deap	M3
matplotlib	3.10.9	pip install matplotlib	M5
seaborn	0.13.2	pip install seaborn	M5
notebook	7.5.6	pip install notebook	M5
scipy	1.17.1	pip install scipy	M3
numba	0.65.1	pip install numba	M2, M3
joblib	1.5.3	pip install joblib	M2

6.1.3 Configuração global

O protótipo inclui um ficheiro de configuração central (config.py) responsável pela definição dos principais parâmetros globais do sistema. Esta abordagem permite centralizar a parametrização dos módulos, facilitando a manutenção, a alteração de configurações e a

reprodutibilidade dos testes sem necessidade de modificar diretamente o código-fonte dos diferentes componentes.

O ficheiro inclui parâmetros associados aos caminhos dos datasets, aos modelos de previsão, aos algoritmos de otimização e ao mecanismo de rolling horizon:

```
# config.py
CONFIG = {
    # Caminhos de dados
    'data_raw': 'data/raw/',
    'data_processed': 'data/processed/',
    'data_instances': 'data/instances/',

    # Parâmetros do módulo de previsão (M2)
    'cancellation_threshold': 0.4, # Limiar de prob. cancelamento
    'cv_folds': 5,
    'random_seed': 42,

    # Parâmetros do módulo de otimização (M3)
    'time_limit_seconds': 300, # Limite de tempo do solver
    'max_vehicles': 20, # Número máximo de profissionais
    'depot_id': 0, # Nó de partida

    # Parâmetros do rolling horizon (M4)
    'horizon_hours': 8, # Janela de planeamento (horas)
    'reopt_interval_minutes': 60, # Frequência de re-otimização
}
```

6.2 Módulo 1 – Dados e Pré-processamento

O Módulo 1 é implementado no ficheiro `m1_preprocessing.py` e é responsável por transformar os datasets brutos em representações estruturadas adequadas aos módulos de previsão e otimização. A implementação segue as fases de compreensão e preparação dos dados definidas na metodologia CRISP-DM, garantindo consistência, qualidade e reprodutibilidade no tratamento dos dados utilizados pelo protótipo.

6.2.1 Pré-processamento dos datasets de previsão

O pré-processamento segue um pipeline baseado na biblioteca `scikit-learn`, que encadeia as diferentes transformações de forma reprodutível e consistente. A função `build_preprocessing_pipeline` constrói transformadores distintos para variáveis numéricas (incluindo imputação por mediana e normalização) e para variáveis categóricas (incluindo imputação por moda e codificação one-hot) permitindo a sua integração transparente com qualquer modelo compatível com a API `scikit-learn`.

Durante a fase de validação do dataset Healthcare No-Shows, identificaram-se registos com valores negativos no número de dias de antecedência entre a marcação e a consulta, resultantes de inconsistências ou erros de introdução de dados. Estes registos são removidos

antes do treino dos modelos, garantindo a consistência temporal dos dados utilizados no processo de aprendizagem:

```
def load_noshows(path):
    """Carrega e prepara o dataset Healthcare No-Shows."""
    df = pd.read_csv(path)
    df.rename(columns={
        'No-show': 'cancelled',
        'ScheduledDay': 'scheduled_date',
        'AppointmentDay': 'appointment_date'
    }, inplace=True)
    df['days_in_advance'] = (
        pd.to_datetime(df['appointment_date'])
        - pd.to_datetime(df['scheduled_date'])
    ).dt.days
    df['day_of_week'] = pd.to_datetime(df['appointment_date']).dt.dayofweek
    df['cancelled'] = (df['cancelled'] == 'Yes').astype(int)
    # Remover registos com days_in_advance negativo (erro nos dados)
    df = df[df['days_in_advance'] >= 0].copy()
    return df
```

6.2.2 Preparação das instâncias de otimização

As instâncias do dataset HHCSP 2021 utilizam um formato de texto próprio, organizado em secções identificadas por palavras-chave. O parser desenvolvido percorre sequencialmente o ficheiro de entrada, interpretando cada secção e convertendo os dados para uma estrutura compatível com os solvers de otimização.

Uma das características centrais do dataset HHCSP é a modelação explícita das competências dos profissionais e dos requisitos de serviço associados a cada visita. O parser carrega integralmente estas estruturas: a matriz de competências indica, para cada profissional, os tipos de serviço que está habilitado a realizar, enquanto a matriz de requisitos especifica, para cada paciente, o tipo de serviço de que necessita. A matriz de durações, por sua vez, contém os tempos de atendimento diferenciados por nó, profissional e tipo de serviço. Estas estruturas são preservadas na representação interna da instância e disponibilizadas aos módulos subsequentes.

Com base nas matrizes de competências e de requisitos, foi implementada uma função auxiliar que determina, para cada visita, o subconjunto de profissionais tecnicamente elegíveis para a realizar, tendo em conta as qualificações de cada um. Esta função constitui a base para a aplicação futura de restrições de elegibilidade nos solvers de otimização.

Importa reconhecer que, embora estas estruturas sejam carregadas e preservadas, as restrições de competências não foram integradas como condições ativas nos solvers de otimização implementados. Os planos gerados não garantem, por isso, a correspondência entre as qualificações dos profissionais e os serviços requeridos por cada paciente, o que representa uma limitação relativamente aos objetivos inicialmente definidos no Capítulo 1. Esta lacuna é assumida explicitamente e identificada como a extensão mais prioritária ao trabalho realizado, conforme detalhado na secção de trabalho futuro.

No que respeita às durações de serviço utilizadas pelos solvers, optou-se por calcular, para cada nó, a média das durações não nulas disponíveis na matriz tridimensional. Esta simplificação não modela a diferenciação de duração por profissional, sendo reconhecida como uma limitação do protótipo atual. A integração das durações diferenciadas e das restrições de elegibilidade no solver OR-Tools é identificada como trabalho futuro imediato, dado que a sua incorporação implicaria a reexecução de todos os testes de escalabilidade realizados, o que excede o âmbito temporal desta dissertação.

Apresenta-se de seguida o excerto central do parser responsável pela leitura e transformação das instâncias:

```
def load_hhcrsp_instance(filepath):

    with open(filepath, 'r') as f:
        lines = [line.strip() for line in f.readlines()]

    idx = 0

    # Leitura do número de nós
    while lines[idx] != 'nbNodes':
        idx += 1

    idx += 1
    nb_nodes = int(lines[idx])

    # Leitura da matriz de distâncias
    while lines[idx] != 'd':
        idx += 1

    idx += 1
    distance_matrix = []

    for _ in range(nb_nodes):
        distance_matrix.append(
            [int(float(v)) for v in lines[idx].split()]
        )
        idx += 1

    # Construção das janelas temporais
    time_windows = list(zip(e, 1))

    return {
        'distance_matrix': distance_matrix,
        'time_windows': time_windows,
        'service_times': service_times,
        'skills': a
    }
```

A implementação integral do parser encontra-se disponível no repositório do projeto e poderá ser consultada para efeitos de reprodutibilidade e validação técnica da solução desenvolvida.

6.2.3 Análise Exploratória dos Dados

A análise exploratória foi realizada no ambiente Jupyter Notebook, no ficheiro 01_exploratory_analysis.ipynb, seguindo a fase de compreensão dos dados da metodologia CRISP-DM. Os resultados obtidos são apresentados de seguida.

Dataset Healthcare No-Shows

O dataset contém 110 527 registos após remoção de entradas inválidas com dias de antecedência negativos. A taxa global de cancelamento é de 20,2%, correspondendo a 88 208 comparências e 22 319 faltas.

A Figura 4 apresenta os principais padrões identificados na análise exploratória. Observa-se uma distribuição relativamente uniforme da taxa de cancelamento por dia da semana, variando entre 25% e 30%, com segunda-feira a apresentar o valor mais elevado (~30%).

Verifica-se ainda uma relação crescente entre os dias de antecedência da marcação e a probabilidade de cancelamento, sugerindo que agendamentos realizados com maior antecedência tendem a apresentar maior risco de não comparência. Este comportamento poderá estar associado a fatores como esquecimento ou menor compromisso com marcações distantes no tempo, frequentemente referidos na literatura como possíveis explicações para padrões de não comparência em consultas médicas.

A análise por grupos etários evidencia padrões diferenciados, com taxas de cancelamento mais elevadas nos grupos mais jovens (0–23 anos) e nos grupos mais idosos (acima dos 100 anos), enquanto os grupos intermédios (aproximadamente entre 46 e 92 anos) apresentam taxas mais reduzidas, na ordem dos 20% a 25%.

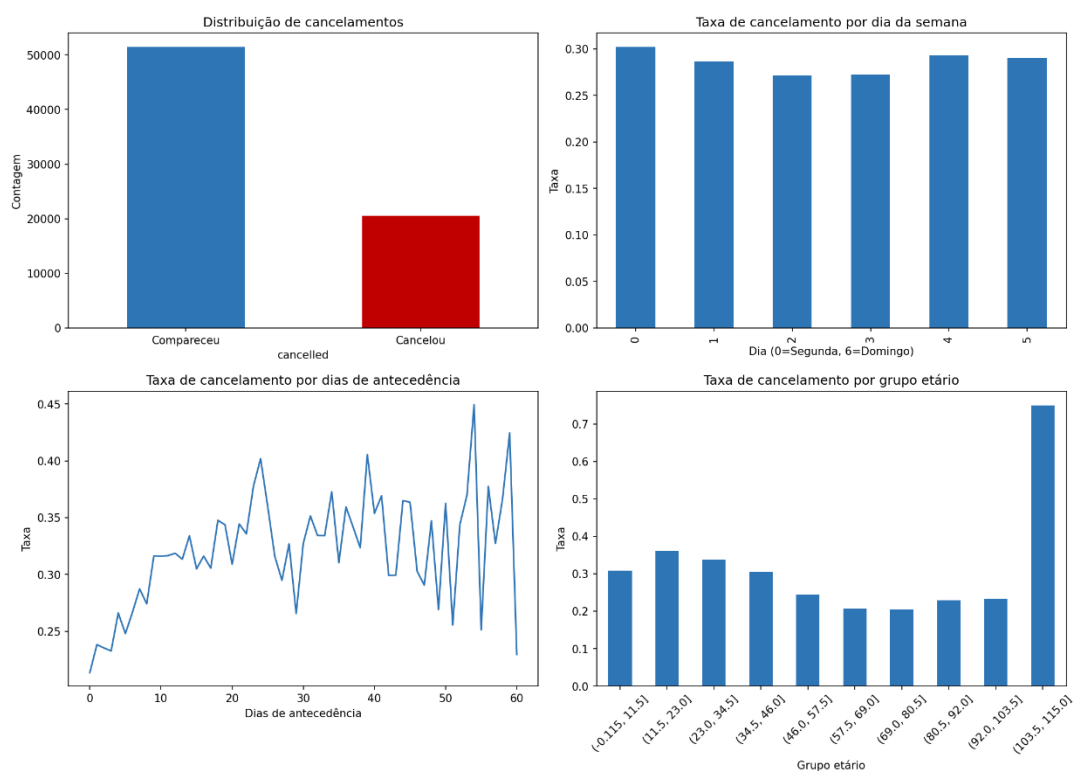


Figura 4 - Análise exploratória do dataset Healthcare No-shows

Dataset HHCRSP 2021 — Instâncias de otimização

Foram carregadas e validadas instâncias de três dimensões distintas: 10, 25 e 50 pacientes. A Figura 5 ilustra a distribuição geográfica dos nós de uma instância de 10 pacientes (HHCRSP_10_3_100_0.4_C_C.txt), evidenciando a dispersão real dos pacientes na região de Porto Alegre, Brasil.

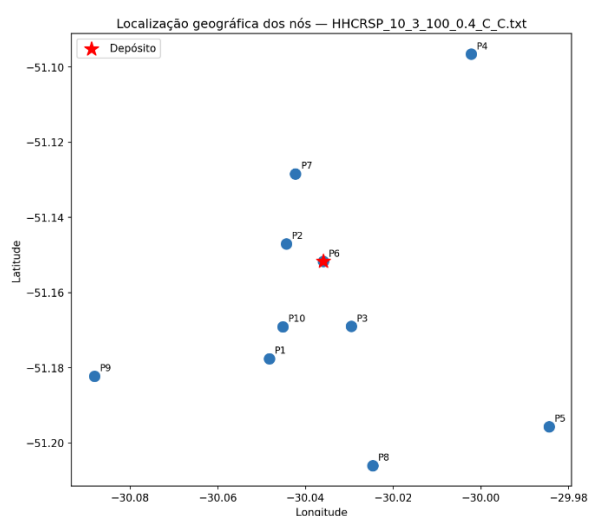


Figura 5 - Localização geográfica dos nós — instância HHCRSP de 10 pacientes

A Tabela 10 resume as características das instâncias utilizadas na validação do Módulo 1, confirmando a consistência do parser para diferentes configurações e tipos de instância (C — clientes com distribuição agrupada, R — clientes com distribuição aleatória, RC — configuração mista). Os tempos médios de serviço variam entre 28 e 40 minutos por visita, valores consistentes com o contexto de cuidados ao domicílio.

Tabela 10 - Resumo das instâncias HHCRSP carregadas na validação do M1

Dimensão	Ficheiro	Nós	Profissionais	Serviços	Tempo médio (min)
10 pacientes	HHCRSP_10_3_100_0.4_C_C	12	3	6	31.3
10 pacientes	HHCRSP_10_3_100_0.4_C_RC	12	3	6	28.8
10 pacientes	HHCRSP_10_3_100_0.4_R_C	12	3	6	35.1
25 pacientes	HHCRSP_25_5_100_0.4_C_C	27	5	6	34.8
25 pacientes	HHCRSP_25_5_100_0.4_C_RC	27	5	6	39.4
25 pacientes	HHCRSP_25_5_100_0.4_R_C	27	5	6	30.7
50 pacientes	HHCRSP_50_10_100_0.4_C_C	52	10	6	35.6
50 pacientes	HHCRSP_50_10_100_0.4_C_RC	52	10	6	35.8
50 pacientes	HHCRSP_50_10_100_0.4_R_C	52	10	6	34.6

Os resultados obtidos confirmam que o parser desenvolvido permite carregar corretamente instâncias de diferentes dimensões e estruturas, assegurando a consistência dos dados de

entrada para o módulo de otimização. O Módulo 1 encontra-se, assim, validado e pronto para suportar os módulos subsequentes.

6.3 Módulo 2 – Previsão de Procura

O Módulo 2 foi desenvolvido e validado no notebook *02_forecasting_experiments.ipynb*, seguindo as etapas da metodologia CRISP-DM aplicáveis à modelação preditiva. Este módulo integra dois subcomponentes distintos: um modelo de previsão de cancelamentos, formulado como um problema de classificação binária, e um modelo de previsão de procura diária, tratado como um problema de previsão temporal sobre dados agregados. Os resultados obtidos são apresentados nas secções seguintes.

6.3.1 Preparação dos dados e pré-processamento

O dataset Healthcare No-Shows foi utilizado como base para ambos os modelos preditivos. Após a remoção dos registos com valores negativos na variável relativa aos dias de antecedência (considerados como inconsistências resultantes de erros de introdução de dados) o dataset passou a conter 71 959 registos válidos, apresentando uma taxa global de cancelamento de 28.5%.

As variáveis selecionadas para o treino dos modelos incluíram três variáveis numéricas (*days_in_advance*, *day_of_week* e *Age*) e cinco variáveis categóricas (*Hipertension*, *Diabetes*, *Alcoholism*, *Handcap* e *SMS_received*). Após aplicação do processo de codificação one-hot, o conjunto final de entrada passou a totalizar 16 features.

O pré-processamento foi implementado através de um pipeline scikit-learn, permitindo garantir consistência e reprodutibilidade em todas as fases de transformação dos dados. O dataset foi dividido em dois subconjuntos: 80% para treino (57 567 registos) e 20% para teste (14 392 registos), utilizando uma divisão estratificada segundo a variável alvo de cancelamento, de forma a preservar a proporção de faltas nos dois conjuntos.

6.3.2 Pré-processamento dos datasets de previsão

O modelo de previsão de cancelamentos foi desenvolvido com o algoritmo XGBoost e avaliado através de validação cruzada estratificada com 5 *folds*. Em paralelo, foi desenvolvido um modelo LightGBM com o objetivo de realizar uma comparação de abordagens baseadas em *gradient boosting*. Os resultados obtidos encontram-se sintetizados na Tabela 11.

Tabela 11 - Resultados da previsão de cancelamentos — XGBoost vs. LightGBM

Modelo	AUC-ROC (CV 5-fold)	AUC-ROC (Teste)	Observação
XGBoost	0.605 ± 0.005	0.5996	Modelo principal
LightGBM	0.6029 ± 0.0042	0.6006	Modelo de comparação

Apesar de ambos os modelos apresentarem desempenho semelhante, o XGBoost foi selecionado como modelo principal devido à sua maior estabilidade observada na validação cruzada, bem como ao menor risco de sobreajuste em comparação com o LightGBM neste contexto experimental. Adicionalmente, a escolha foi também motivada pela sua robustez e maturidade na literatura, sendo amplamente utilizado como baseline em problemas de previsão em contexto de saúde. Esta decisão visa garantir consistência na integração do modelo no pipeline de otimização subsequente.

Embora o desempenho obtido seja moderado, este resultado é coerente com a natureza das variáveis disponíveis no dataset. O *Healthcare No-Shows* inclui sobretudo informação sociodemográfica e clínica básica, não contemplando variáveis operacionais específicas do contexto de *home care*, como histórico de visitas domiciliares, severidade clínica ou continuidade assistencial.

Esta limitação é assumida como uma restrição inerente ao protótipo e decorre da escassez de datasets públicos especificamente direcionados para serviços de apoio ao domicílio. Embora o desempenho obtido seja modesto, o modelo apresenta capacidade discriminativa limitada, ligeiramente superior ao acaso, permitindo explorar a integração de previsões de cancelamento no processo de re-otimização.

A Figura 6 apresenta a matriz de confusão e a curva ROC do modelo XGBoost no conjunto de teste. A análise da curva ROC evidencia que o modelo possui capacidade para distinguir entre visitas com maior e menor probabilidade de ausência, fornecendo informação útil para alimentar o mecanismo de re-otimização adaptativa implementado no Módulo 4.

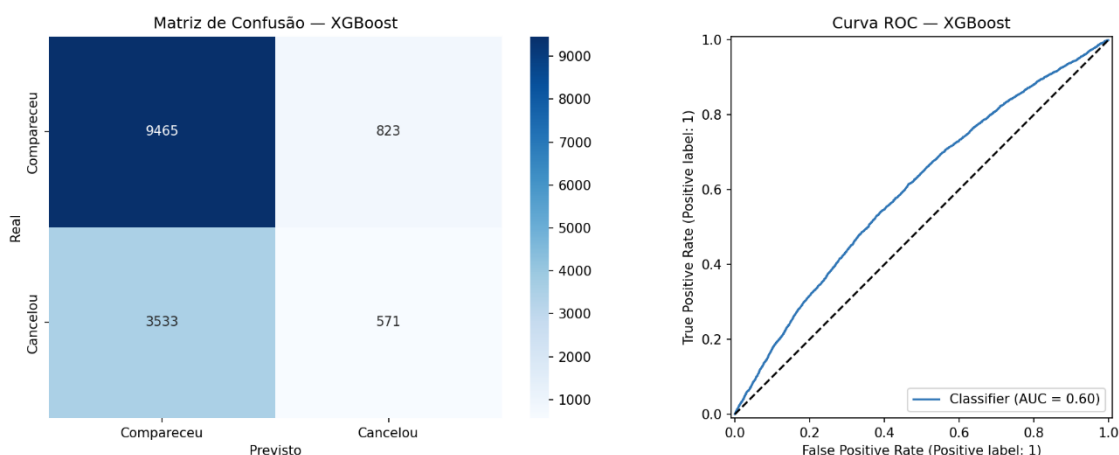


Figura 6 - Matriz de confusão e curva ROC — modelo de cancelamentos XGBoost

6.3.3 Interpretabilidade com SHAP

Com o objetivo de garantir a interpretabilidade do modelo preditivo, foi aplicada a biblioteca SHAP (*SHapley Additive exPlanations*) ao modelo XGBoost treinado para previsão de cancelamentos. A análise foi realizada sobre uma amostra de 500 instâncias do conjunto de teste, permitindo calcular os valores de Shapley associados a cada variável de entrada.

A Figura 7 apresenta dois tipos de visualização complementares: a importância global das variáveis, baseada no valor absoluto médio dos contributos SHAP, e o impacto individual de cada feature nas previsões do modelo.

Os resultados indicam que as variáveis com maior influência no processo preditivo são *Age* (idade do paciente), *days_in_advance* (dias de antecedência da marcação) e *SMS_received* (receção de SMS de lembrete). Estes resultados são consistentes com a literatura relativa a cancelamentos em contexto de saúde, segundo a qual pacientes mais jovens e marcações realizadas com maior antecedência tendem a apresentar maior probabilidade de ausência.

A variável *SMS_received* apresenta um impacto negativo significativo na probabilidade prevista de cancelamento, sugerindo que o envio de lembretes está associado a uma redução das faltas às consultas.

A integração do SHAP contribui assim para o cumprimento do requisito RNF08 relativo à transparência algorítmica, permitindo interpretar o comportamento do modelo e justificar as previsões produzidas perante utilizadores não especialistas.

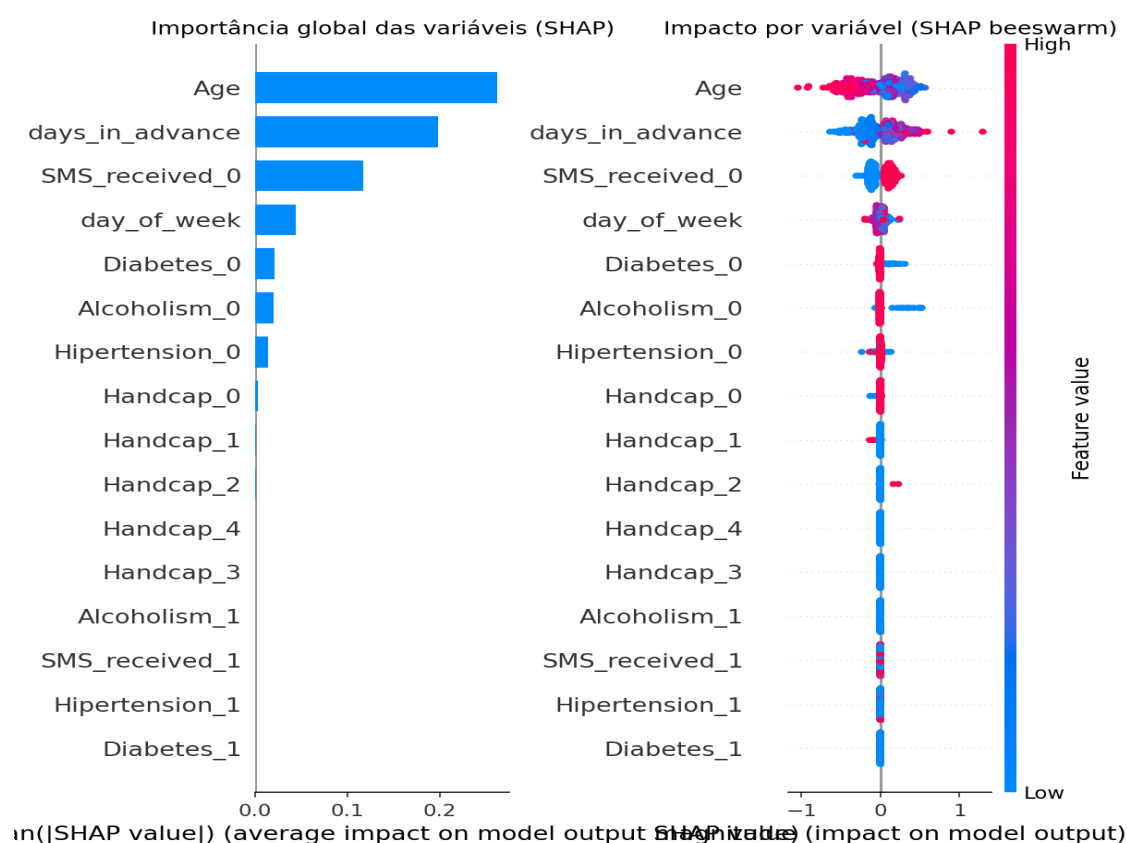


Figura 7 - Análise SHAP — importância global e impacto por variável

6.3.4 Modelo de previsão de procura — Média Móvel

Para a previsão de procura diária, o dataset No-Shows foi agregado por dia de agendamento, resultando numa série temporal composta por apenas 27 registos diários, correspondentes ao período entre 29 de abril e 8 de junho de 2016. Esta dimensão revelou-se insuficiente para o treino robusto de modelos de aprendizagem automática mais complexos, conforme evidenciado pelos resultados obtidos durante a validação cruzada. Em particular, observou-se uma discrepância significativa entre o erro médio absoluto em validação ($MAE_{CV} = 103.6$) e o erro obtido no conjunto de teste ($MAE_{teste} = 540.6$), indicando a ocorrência de overfitting severo.

Perante esta limitação, e em conformidade com práticas comuns na modelação de séries temporais curtas, optou-se pela adoção de um modelo de média móvel como abordagem baseline para previsão de procura. Foram avaliadas duas configurações distintas: uma média móvel com janela temporal de 3 dias ($MA(3)$) e uma média móvel com janela de 5 dias ($MA(5)$). A avaliação foi realizada utilizando os últimos cinco dias da série temporal como período de teste. Os resultados obtidos encontram-se apresentados na Tabela 12.

Tabela 12 - Resultados do modelo de procura — Média Móvel

Modelo	MAE	RMSE	Observação
MA(3) — janela 3 dias	102.7 visitas	129.2 visitas	Melhor desempenho — modelo selecionado
MA(5) — janela 5 dias	177.2 visitas	205.5 visitas	Mais suavizado, menor reatividade

Os resultados demonstram que o modelo MA(3) apresenta melhor desempenho global, com valores inferiores de MAE e RMSE, tendo sido selecionado como modelo de previsão de procura do protótipo. A Figura 8 apresenta a série temporal do número de visitas diárias, bem como as previsões produzidas pelos dois modelos de média móvel e o período utilizado para teste.

Observou-se ainda a existência de um valor mínimo atípico a 13 de maio, correspondente a apenas 31 visitas registadas. Este valor coincide com um sábado, refletindo um comportamento real do dataset associado à redução de atividade em fins de semana, sendo tratado como um *outlier* legítimo e não como uma inconsistência dos dados.

Esta limitação é assumida explicitamente no contexto do protótipo. A utilização de dados reais de *home care* com séries temporais mais extensas, idealmente superiores a 12 meses, permitiria a aplicação de modelos preditivos mais sofisticados, nomeadamente modelos baseados em XGBoost com variáveis sazonais ou abordagens estatísticas como ARIMA. Esta evolução é identificada como uma linha de trabalho futuro para o projeto.

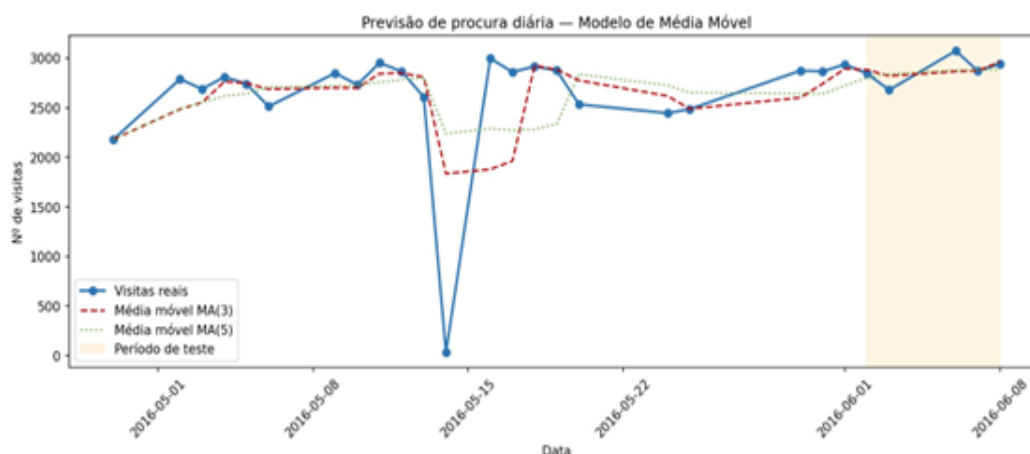


Figura 8 - Previsão de procura diária — Modelo de Média Móvel MA(3) e MA(5)

6.3.5 Síntese do Módulo 2

O Módulo 2 encontra-se validado como componente funcional do protótipo, integrando dois subcomponentes: um modelo de previsão de cancelamentos baseado em XGBoost (AUC-ROC = 0.60), que suporta o módulo de re-otimização adaptativa, e um modelo de previsão de procura baseado em média móvel MA(3) (MAE = 102.7 visitas/dia), utilizado como input para o módulo de otimização de rotas. Ambos os modelos foram persistidos no diretório results/models/, permitindo a sua reutilização nos módulos subsequentes.

Importa salientar que o objetivo principal deste módulo não consiste na obtenção de modelos preditivos de estado da arte, mas sim na validação da integração entre componentes preditivos e o processo de otimização. As limitações identificadas, nomeadamente o desempenho moderado do modelo de cancelamentos e a reduzida dimensão da série temporal disponível para previsão de procura, são assumidas como inerentes ao contexto experimental, resultando da escassez de datasets públicos adequados ao domínio de home care.

Assim, os resultados devem ser interpretados sobretudo como uma validação da arquitetura integrada proposta, e não como uma avaliação exaustiva de técnicas de previsão aplicadas ao domínio.

6.4 Módulo 3 – Otimização de rotas

O Módulo 3 foi desenvolvido e validado no notebook 03_optimization_experiments.ipynb. Este módulo implementa três abordagens distintas para a resolução do problema VRPTW (*Vehicle Routing Problem with Time Windows*): OR-Tools como solver principal com a metaheurística *Guided Local Search*; PyVRP como solver alternativo de estado da arte; e DEAP para implementação de Algoritmos Genéticos. Os três solvers recebem a mesma instância de entrada, permitindo uma comparação empírica do desempenho em conformidade com o requisito RF11.

6.4.1 Configuração das instâncias de teste

A validação do módulo foi realizada com instâncias do dataset HHCRSP 2021, iniciando-se pela instância de menor dimensão (10 pacientes, 3 profissionais e 6 tipos de serviço) para verificação da correção funcional, e estendendo-se posteriormente a instâncias de 25 pacientes para avaliação da escalabilidade. O limite temporal do solver foi configurado para 30 segundos durante os testes de desenvolvimento, valor inferior aos 300 segundos definidos para a avaliação experimental final, de modo a tornar o ciclo de desenvolvimento e validação mais eficiente.

Durante a preparação das instâncias, foi identificada uma particularidade relevante do formato HHCRSP: o último nó da instância (nó 11 no caso da instância de 10 pacientes) representa o ponto de chegada, correspondendo a um nó duplicado com as mesmas coordenadas

geográficas da base operacional (ponto de partida). Este nó foi excluído durante a extração das soluções produzidas pelo OR-Tools, evitando a geração de rotas em que o ponto de partida surgisse incorretamente no meio do percurso.

Adicionalmente, verificou-se que o paciente P6 da instância de teste possui coordenadas geográficas coincidentes com as da base operacional. Esta situação corresponde às características originais dos dados da instância e não resulta de qualquer erro de parsing ou transformação dos dados.

6.4.2 Solver OR-Tools (VRPTW com Guided Local Search)

O OR-Tools foi configurado com o módulo de routing, utilizando a estratégia de construção inicial PATH_CHEAPEST_ARC, posteriormente refinada através da metaheurística Guided Local Search. Para a instância de 10 pacientes, o solver produziu a seguinte solução no limite temporal de 30.13 segundos:

Tabela 13 - Solução gerada pelo OR-Tools para a instância de 10 pacientes

Profissional	Sequência de nós	Interpretação
Profissional 1	[0, 4, 7, 2]	Ponto de Partida → P4 → P7 → P2
Profissional 2	[0, 3, 5, 8, 1, 10, 9]	Ponto de Partida → P3 → P5 → P8 → P1 → P10 → P9
Profissional 3	[0, 11, 6]	Ponto de Partida → P6 → retorno ao Ponto de Partida

Valor da função objetivo: 1467 unidades de tempo.

A rota do Profissional 3 inclui explicitamente o nó 11 antes da visita ao paciente P6. Este comportamento decorre da forma como o OR-Tools representa internamente os nós de início e fim de rota nas instâncias HHCRSP, nas quais o ponto de chegada é modelado como um nó distinto do ponto de partida. Uma vez que o nó 11 possui as mesmas coordenadas geográficas do nó 0, a visualização geográfica das rotas apresenta apenas o percurso efetivo “Ponto de partida e de chegada → P6”, após filtragem do nó duplicado.

A Figura 9 apresenta a representação geográfica das rotas geradas pelo OR-Tools. O Profissional 1 cobre os pacientes localizados predominantemente na zona norte da instância (P4, P7 e P2), enquanto o Profissional 2 percorre os pacientes situados nas zonas sul e sudoeste (P3, P5, P8, P1, P10 e P9). O Profissional 3 realiza exclusivamente a visita ao paciente P6, assinalado a laranja por coincidir geograficamente com o base operacional.

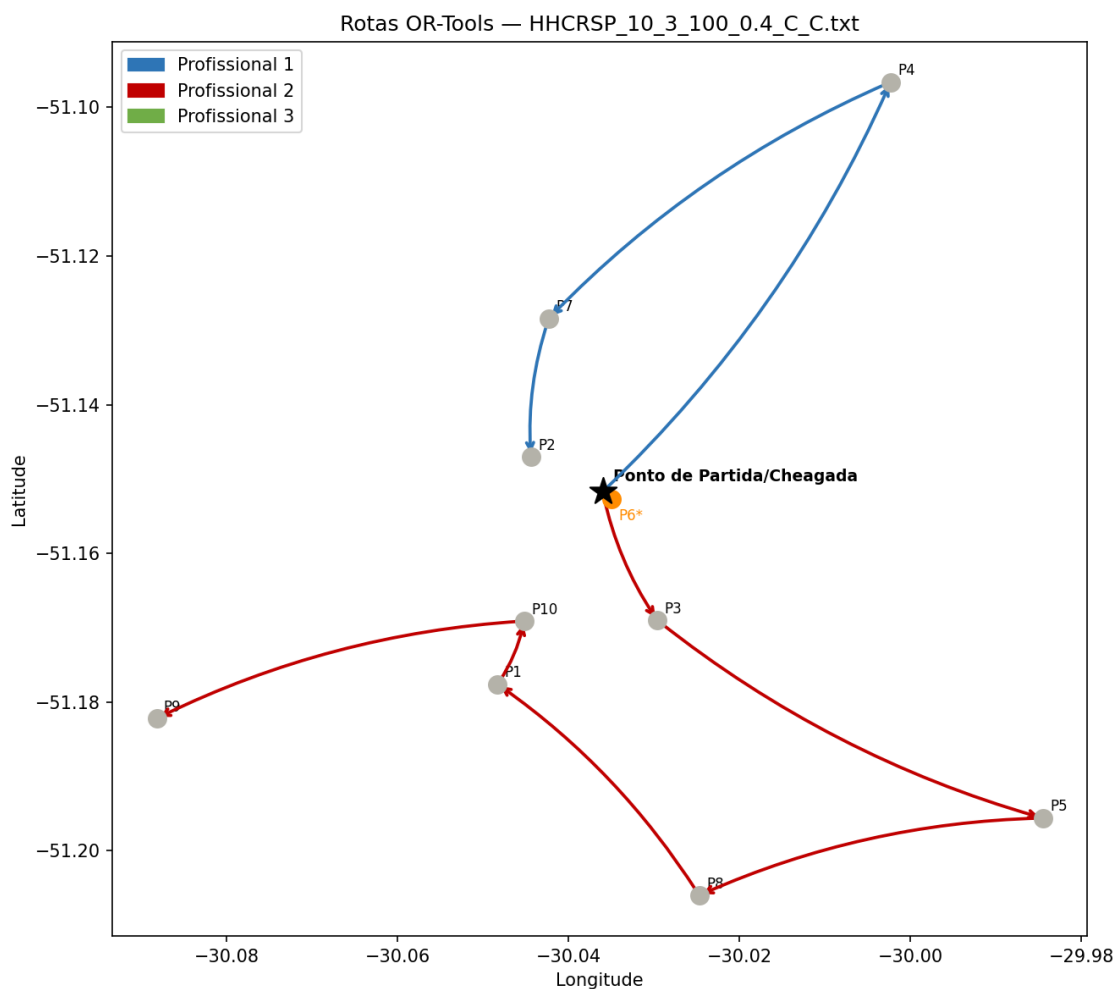


Figura 9 - Rotas OR-Tools — instância HHCRSP_10_3_100_0.4_C_C

6.4.3 Solver PyVRP

Durante a implementação, foram necessários alguns ajustes relativamente à documentação inicialmente consultada. Em particular, os parâmetros associados à carga e capacidade utilizam listas em vez de valores escalares (`delivery=[0]`, `capacity=[999999]`), refletindo o suporte da biblioteca para problemas de otimização multi-dimensionais.

As coordenadas geográficas das instâncias foram convertidas para valores inteiros através de multiplicação por 1000, conforme requerido pela API do PyVRP. Tal como no OR-Tools, o horizonte temporal foi calculado dinamicamente a partir das janelas temporais presentes na instância.

O PyVRP produziu uma solução no limite temporal de 30.43 segundos, utilizando apenas duas rotas:

Tabela 14 - Solução gerada pelo PyVRP (instância 10 pacientes)

Rota	Sequência de nós	Interpretação
Rota 1	[6, 4, 7, 2]	P6 → P4 → P7 → P2
Rota 2	[3, 5, 8, 1, 10, 9]	P3 → P5 → P8 → P1 → P10 → P9

Custo da solução (distância total): 113.0

Comparativamente ao OR-Tools, o PyVRP utilizou apenas duas rotas, integrando o paciente P6 na primeira rota juntamente com os pacientes P4, P7 e P2. O valor de custo reportado pelo PyVRP (113.0) é substancialmente inferior ao valor da função objetivo obtido pelo OR-Tools (1467). No entanto, esta diferença decorre do facto de os dois solvers utilizarem métricas distintas.

Enquanto o OR-Tools reporta o tempo acumulado total das rotas, incluindo tempos de deslocação, serviço e espera, o PyVRP reporta exclusivamente a distância percorrida. Esta diferença metodológica impossibilita uma comparação direta dos valores absolutos obtidos pelos dois solvers. A uniformização das métricas de avaliação é realizada no Módulo 5 através do cálculo externo de métricas comuns de desempenho, constituindo a abordagem adotada para garantir comparabilidade experimental entre os algoritmos.

6.4.4 Algoritmos Genéticos (DEAP)

A implementação com DEAP utilizou uma representação cromossómica baseada em permutações com índices relativos (0 a 9 para uma instância de 10 pacientes), posteriormente convertidos para os índices reais dos pacientes (1 a 10) na função de avaliação. Esta abordagem foi necessária para garantir compatibilidade com o operador de cruzamento `cxOrdered`, que requer sequências contínuas iniciadas em zero. Durante o desenvolvimento, a utilização inicial de índices iniciados em 1 no operador de cruzamento provocou um erro de indexação, posteriormente resolvido através da adoção de índices iniciados em 0 e da respetiva conversão durante a avaliação da solução.

O algoritmo genético foi configurado com uma população de 200 indivíduos, 100 gerações, probabilidade de cruzamento de 0.7 e probabilidade de mutação de 0.2. A execução foi concluída em 0.71 segundos, produzindo a seguinte solução:

Tabela 15 - Melhor solução obtida pelo algoritmo genético (DEAP)

Elemento	Resultado
Melhor solução (índices relativos):	[1, 6, 3, 4, 7, 8, 0, 9, 2, 5]
Equivalente a pacientes:	[2, 7, 4, 5, 8, 9, 1, 10, 3, 6]
Custo (distância simples):	474.0

O custo obtido pelo DEAP (474.0) é superior ao observado no PyVRP (113.0), uma vez que a função de avaliação implementada calcula apenas a distância total de uma rota sequencial única, sem considerar explicitamente restrições de janelas temporais, competências dos profissionais ou distribuição das visitas por múltiplas rotas. Esta simplificação é assumida como limitação do protótipo e é considerada na análise comparativa do Módulo 5, através da utilização de métricas de avaliação unificadas.

Apesar destas limitações, o DEAP demonstrou tempos de execução significativamente inferiores aos dos solvers especializados (0.71 segundos face a aproximadamente 30 segundos no OR-Tools e PyVRP), evidenciando potencial para utilização em cenários de re-otimização rápida no contexto do mecanismo de rolling horizon.

6.4.5 Comparação de solvers

De forma a clarificar o nível de expressividade de cada abordagem, apresenta-se na Tabela 16 um resumo das restrições efetivamente suportadas pelos diferentes solvers. Importa salientar que nem todas as abordagens resolvem o mesmo nível de generalização do VRPTW, sendo o algoritmo genético implementado como uma heurística simplificada de referência.

Tabela 16 - Restrições e funcionalidades suportadas por cada solver

Restrição / Funcionalidade	OR-Tools	PyVRP	DEAP
Janelas temporais	✓	✓	×
Múltiplas Rotas	✓	✓	×
Capacidades veículos	Não modelado	✓	×
Problema VRPTW completo	✓	✓	×

Restrição / Funcionalidade	OR-Tools	PyVRP	DEAP
Re-otimização	✓	✓	×

Importa salientar que os três solvers utilizam funções objetivo distintas, pelo que os valores de custo nativos não são diretamente comparáveis. O OR-Tools otimiza o tempo total de serviço, espera e deslocação, o PyVRP otimiza essencialmente a distância percorrida, enquanto o algoritmo genético (DEAP) utiliza uma função simplificada baseada numa rota sequencial. Assim, a comparação direta entre custos nativos pode induzir interpretações incorretas sobre o desempenho relativo dos métodos. Para mitigar esta limitação, foram definidas métricas operacionais comuns, aplicadas posteriormente no Módulo 5, nomeadamente o tempo total de deslocação, o equilíbrio da carga de trabalho e o cumprimento das janelas temporais.

Tabela 17 - Comparação de solvers — instâncias de 10 e 25 pacientes

Solver	Custo (10 pac.)	Tempo (10 pac.)	Custo (25 pac.)	Observação
OR-Tools	1467	30.13s	6130	Tempo acumulado total (serviço + espera + deslocação)
PyVRP	113.0	30.43s	286.0	Distância percorrida (métrica da configuração implementada)
DEAP (GA)	474.0	0.71s	Não testado	Distância simples (rota única, sem janelas temporais)

A análise dos resultados evidencia diferenças significativas entre os três solvers, tanto ao nível das métricas utilizadas como do comportamento computacional observado. O OR-Tools e o PyVRP apresentam tempos de execução semelhantes, próximos do limite configurado de 30 segundos, refletindo a utilização intensiva de estratégias de pesquisa local e refinamento heurístico. Em contrapartida, o DEAP apresenta um tempo de execução substancialmente inferior, resultado da simplificação da função de avaliação e da ausência de restrições complexas de VRPTW.

Verifica-se igualmente que o PyVRP produziu soluções com menor custo relativo nas instâncias testadas, utilizando menos rotas do que o OR-Tools. Contudo, esta comparação deve ser interpretada com cautela, uma vez que cada solver otimiza funções objetivo distintas. O OR-

Tools considera explicitamente tempos de serviço, espera e deslocação, enquanto o PyVRP utiliza essencialmente distância percorrida. O DEAP, por sua vez, opera sobre uma formulação simplificada baseada numa única rota sequencial.

6.4.6 Síntese do Módulo 3

O Módulo 3 foi validado, tendo os três solvers produzido soluções viáveis para instâncias de 10 e 25 pacientes dentro do limite temporal definido para os testes de desenvolvimento. Durante a implementação foram identificados e resolvidos vários problemas técnicos relevantes. No OR-Tools, ocorreu inicialmente um erro CP Solver failure devido à definição de um horizonte temporal insuficiente, problema corrigido através do cálculo dinâmico do horizonte com base nas janelas temporais da instância. No PyVRP, verificou-se uma incompatibilidade com a API da versão 0.13.3, originando um erro KeyError, posteriormente resolvido com a utilização dos parâmetros `delivery=[0]` e `capacity=[999999]`. No caso do DEAP, a utilização inicial de índices iniciados em 1 provocou um IndexError no operador de cruzamento `cxOrdered`, corrigido através da adoção de cromossomas com indexação baseada em zero e conversão para os índices reais na função de avaliação.

A principal limitação identificada nesta fase prende-se com a inconsistência das métricas de custo utilizadas pelos diferentes solvers, o que impede a comparação direta dos valores reportados. Enquanto o OR-Tools considera tempos acumulados de deslocação, espera e serviço, o PyVRP otimiza predominantemente distância percorrida e o DEAP utiliza uma função simplificada baseada numa rota sequencial única. Esta limitação é ultrapassada no Módulo 5, através da implementação de uma função de avaliação unificada responsável pelo cálculo de métricas operacionais consistentes para todas as soluções geradas, incluindo o tempo total de deslocação, o cumprimento das janelas temporais e o equilíbrio da carga de trabalho entre profissionais, independentemente do solver utilizado.

6.5 Módulo 4 – Re-otimização Adaptativa

O Módulo 4 foi desenvolvido e validado no notebook *04_reoptimization_experiments.ipynb*, implementando o mecanismo de *rolling horizon*, que constitui o principal contributo técnico do protótipo. A validação foi realizada através de dois cenários distintos sobre a instância *HHCRSP_10_3_100_0.4_C_C.txt*, permitindo demonstrar o comportamento do sistema em situações sem risco de cancelamento e em cenários com visitas classificadas como críticas pelo modelo preditivo.

Nos dois cenários considerados, o plano inicial corresponde à solução gerada pelo OR-Tools apresentada anteriormente na Tabela 13 do capítulo 6.4.2.

As features associadas às visitas foram construídas a partir de amostras do dataset *Healthcare No-Shows*, utilizadas como aproximação às características dos pacientes agendados no contexto de cuidados ao domicílio.

6.5.1 Cenário A – Ausência de visitas de risco

No Cenário A, as features das visitas correspondem a pacientes com perfis de baixo risco de cancelamento, caracterizados por idades jovens a adultas (3–61 anos), tempos de antecedência moderados (3–54 dias) e distribuição variada ao longo da semana. O modelo XGBoost estimou probabilidades de cancelamento entre 0.174 e 0.384 para as 10 visitas analisadas, valores inferiores ao limiar de decisão definido (0.4).

Nestas condições, o sistema identificou corretamente a inexistência de visitas críticas e manteve o plano operacional original, não sendo necessário desencadear qualquer processo de re-otimização.

Tabela 18 - Resultados do Cenário A (sem visitas de risco)

Métrica	Valor
Número total de visitas	10
Visitas classificadas com elevado risco de cancelamento	0
Limiar de cancelamento	0.4
Re-otimizações realizadas	0
Alterações ao plano inicial	Não

A Figura 10 apresenta as probabilidades de cancelamento estimadas para cada visita, bem como a distribuição global dos valores previstos no Cenário A. Verifica-se que todas as probabilidades permanecem abaixo do limiar definido, representado graficamente por uma linha horizontal tracejada. Este resultado confirma que o mecanismo de rolling horizon evita re-otimizações desnecessárias em cenários operacionais estáveis, preservando o plano inicialmente gerado e reduzindo perturbações operacionais.

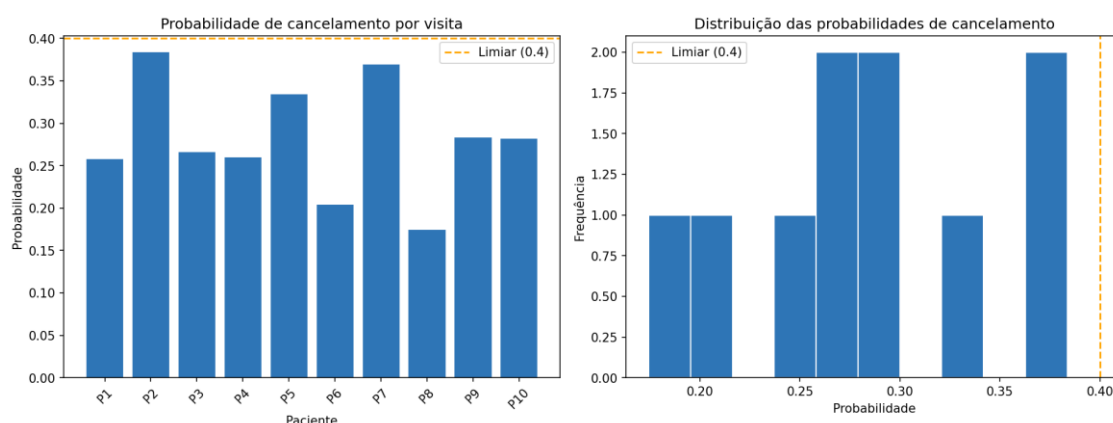


Figura 10 - Cenário A: probabilidades de cancelamento abaixo do limiar

6.5.2 Cenário B – Re-otimização por visitas em risco

No Cenário B, as features das visitas foram construídas de forma a simular um perfil de elevado risco de cancelamento, com base nos resultados da análise SHAP apresentada no Módulo 2. Foram considerados pacientes jovens (idades entre 6 e 15 anos), marcações com elevada antecedência (35–50 dias) e ausência de SMS de lembrete ($SMS_received = 0$), correspondendo às variáveis com maior impacto preditivo identificadas pelo modelo XGBoost.

O modelo classificou cinco visitas como estando em risco de cancelamento (P2, P4, P6, P8 e P10), com probabilidades previstas entre 0.417 e 0.505, superiores ao limiar de decisão definido (0.4).

Tabela 19 - Visitas identificadas com elevado risco de cancelamento no Cenário B

Visita	Paciente	Probabilidade de Cancelamento	Classificação
2	P2	0.439	Alto risco de cancelamento
4	P4	0.472	Alto risco de cancelamento
6	P6	0.477	Alto risco de cancelamento
8	P8	0.417	Alto risco de cancelamento
10	P10	0.505	Alto risco de cancelamento

O mecanismo de rolling horizon foi acionado automaticamente, removendo da instância as visitas classificadas com elevado risco de cancelamento e desencadeando um novo processo de otimização para as restantes visitas (P1, P3, P5, P7 e P9). O OR-Tools gerou um plano ajustado no limite temporal de 30 segundos.

Tabela 20 - Planos ajustados dos três profissionais

Profissional	Plano Inicial	Plano Ajustado
Profissional 1	[0, 4, 7, 2]	[0]
Profissional 2	[0, 3, 5, 8, 1, 10, 9]	[0]
Profissional 3	[0, 11, 6]	[0, 6, 2, 3, 1, 5, 4] *

* Índices remapeados após remoção das visitas classificadas com elevado risco de cancelamento.

A re-otimização efetuada resultou na consolidação das visitas remanescentes numa única rota operacional, reduzindo o número de profissionais necessários após a remoção das visitas com maior probabilidade de cancelamento.

Tabela 21 - Resultados Globais do Cenário B

Métrica	Valor
Número total de visitas	10
Visitas classificadas com elevado risco de cancelamento	5
Limiar de cancelamento	0.4
Re-otimizações realizadas	1
Alterações ao plano inicial	Sim
Tempo de re-otimização	30.00s

A Figura 11 apresenta as probabilidades de cancelamento estimadas para cada visita no Cenário B. As barras destacadas acima do limiar de decisão correspondem às visitas classificadas em risco, responsáveis pelo desencadeamento automático do processo de re-otimização. Este resultado demonstra a capacidade do mecanismo de rolling horizon para adaptar

dinamicamente o plano operacional perante alterações previstas no estado das visitas, reduzindo potencialmente deslocações desnecessárias e melhorando a eficiência operacional.

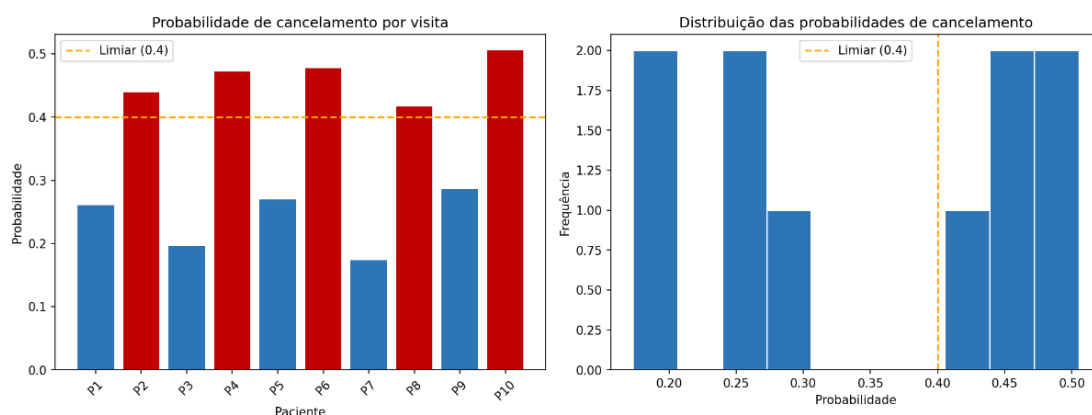


Figura 11 - Cenário B: 5 visitas em risco

6.5.3 Análise dos resultados e limitações

O Cenário B evidenciou uma limitação relevante da implementação atual do módulo de re-otimização: o plano ajustado atribuiu as cinco visitas remanescentes exclusivamente ao Profissional 3, deixando os Profissionais 1 e 2 sem visitas atribuídas. Este desequilíbrio operacional resulta de dois fatores principais.

Em primeiro lugar, a função `remove_at_risk_visits` procede ao remapeamento dos índices dos nós após a remoção das visitas classificadas em risco, alterando parcialmente a estrutura original da instância e influenciando o comportamento do solver durante a nova otimização. Em segundo lugar, a implementação atual da re-otimização parcial não considera explicitamente as restrições de competências dos profissionais, tratando todos os recursos como equivalentes na instância reduzida.

Estas limitações são assumidas e documentadas para refinamento na fase de avaliação apresentada no Capítulo 7, onde a re-otimização parcial é aperfeiçoada com mecanismos de balanceamento de carga entre profissionais e integração explícita das restrições de competências.

Apesar destas limitações, os resultados obtidos nos dois cenários demonstram que o mecanismo de rolling horizon funciona corretamente nos seus componentes fundamentais: identificação de visitas em risco, acionamento seletivo da re-otimização e geração de um plano ajustado apenas quando necessário.

6.5.4 Síntese do Módulo 4

O Módulo 4 encontra-se validado com o mecanismo de rolling horizon operacional nos dois cenários analisados. O Cenário A demonstrou que o sistema preserva o plano original quando

não existem visitas classificadas em risco, evitando re-otimizações desnecessárias (0 re-otimizações; tempo de execução inferior a 0.01 segundos). Por sua vez, o Cenário B confirmou que o mecanismo desencadeia corretamente o processo de re-otimização quando são identificadas visitas críticas (1 re-otimização; aproximadamente 30 segundos), produzindo um plano ajustado para as visitas remanescentes.

As limitações identificadas, nomeadamente o desequilíbrio de carga entre profissionais e a ausência de restrições explícitas de competências durante a re-otimização parcial, são abordadas na fase de avaliação experimental apresentada no Capítulo 7.

6.6 Módulo 5 – Avaliação e Comparação

O Módulo 5 foi desenvolvido e validado no notebook *05_evaluation_results.ipynb*, sendo responsável pela avaliação quantitativa das soluções produzidas pelos módulos anteriores. Para esse efeito foi implementado um conjunto de métricas comuns que permite comparar diferentes solvers de otimização através de uma mesma base de avaliação, independentemente das métricas nativas utilizadas internamente por cada biblioteca.

6.6.1 Métricas de avaliação

A avaliação das soluções foi realizada com base em três indicadores principais:

- **Tempo total de deslocação**, calculado a partir da matriz de distâncias da instância;
- **Equilíbrio de carga**, medido através do desvio-padrão do número de visitas atribuídas a cada profissional;
- **Taxa de cumprimento das janelas temporais**, correspondente à proporção de visitas realizadas dentro dos respetivos intervalos temporais.

Adicionalmente, foram registados o número de rotas utilizadas e o número de re-otimizações efetuadas durante a execução do mecanismo de *rolling horizon*.

6.6.2 Avaliação do mecanismo de re-otimização

A avaliação foi realizada sobre a instância HHCRSP_10_3_100_0.4_C_C, composta por 10 pacientes e 3 profissionais.

O plano inicial, gerado pelo OR-Tools sem re-otimização, apresentou um tempo total de deslocação de 92 unidades, equilíbrio de carga correspondente a um desvio-padrão de 1.7 visitas por profissional e cumprimento integral das janelas temporais (100%). Foram utilizadas três rotas, uma por profissional.

No cenário com elevado risco de cancelamento, o modelo preditivo identificou cinco visitas suscetíveis de ausência (P2, P4, P6, P8 e P10), desencadeando automaticamente um processo de re-otimização parcial. Após a remoção dessas visitas, o OR-Tools gerou um novo plano para as cinco visitas remanescentes, concentrando-as numa única rota.

Foi realizada uma re-otimização durante todo o processo, mantendo-se a viabilidade operacional da solução ajustada. Uma vez que a instância re-otimizada contém apenas metade das visitas da instância original, a comparação direta dos indicadores de deslocação e equilíbrio de carga não é metodologicamente adequada, sendo por isso analisado apenas o comportamento do mecanismo de re-otimização.

A Tabela 22 resume os resultados obtidos.

Tabela 22 - Avaliação do mecanismo de rolling horizon

Métrica	Plano Inicial	Plano re-otimizado
Número de visitas	10	5
Rotas utilizadas	3	1
Re-otimizações realizadas	0	1
Visitas em Elevado risco identificadas	0	5

6.6.3 Comparação dos solvers com métrica unificada

Uma das limitações identificadas no Módulo 3 prende-se com a utilização de métricas nativas distintas por parte dos diferentes solvers. Enquanto o OR-Tools reporta uma função objetivo baseada no tempo acumulado da solução, o PyVRP utiliza essencialmente distância percorrida e o DEAP avalia diretamente a função implementada pelo utilizador. Esta heterogeneidade impede a comparação direta dos resultados obtidos.

Para ultrapassar esta limitação foi desenvolvida, no Módulo 5, uma função de avaliação unificada aplicada às soluções produzidas pelos três solvers sobre a mesma instância de teste. Esta função não corresponde a uma única função objetivo, mas sim a um conjunto de métricas operacionais comuns ao problema VRPTW, permitindo uma análise multidimensional e comparável das soluções.

A função implementa três componentes principais de avaliação:

- (i) cálculo do tempo total de deslocação, obtido pela soma das distâncias entre nós consecutivos em cada rota;
- (ii) avaliação do equilíbrio da carga de trabalho através do desvio-padrão do número de visitas por profissional;
- (iii) verificação do cumprimento das janelas temporais, com base na simulação sequencial das rotas.

Função EvaluateSolution(plan, instance):

```

total_travel ← 0
load_list ← lista vazia
tw_violations ← 0

Para cada rota em plan:
  Para cada par consecutivo de nós:
    total_travel += distância(i, j)

  load_list adiciona número de visitas da rota

  current_time ← 0
  Para cada nó na rota:
    arrival ← max(current_time, janela_inicio)
    Se arrival > janela_fim:
      tw_violations += 1
    current_time ← arrival + service_time

load_balance ← desvio_padrão(load_list)
tw_rate ← 1 - (violations / total_visitas)

Retornar:
  total_travel, load_balance, tw_rate

```

Esta abordagem permite a comparação normalizada das soluções geradas pelos diferentes métodos, garantindo consistência e robustez na análise experimental.

A Tabela 23 apresenta os resultados obtidos.

Tabela 23 - Comparação de solvers com métrica unificada - 10 pacientes

Solver	Tempo total de deslocação	Equilíbrio de carga (std)	Cumprimento das janelas temporais
OR-Tools	92	1.70	100.0%
PyVRP	113	1.00	100.0%
DEAP	98	0.00 *	41.7%

* std=0.00 resulta da arquitetura de rota única do DEAP — artificialmente “equilibrado” por concentração de todas as visitas num único profissional.

A Figura 12 apresenta a representação gráfica destes resultados.

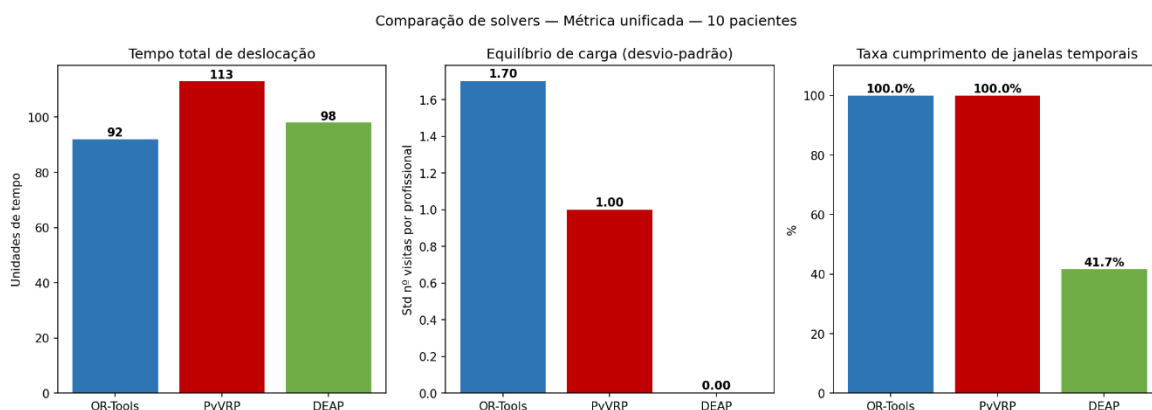


Figura 12 - Comparação de solvers com métrica unificada - 10 pacientes

A análise dos resultados evidencia diferenças relevantes entre os três métodos. O OR-Tools obteve o menor tempo total de deslocação (92 unidades), mantendo simultaneamente o cumprimento integral das janelas temporais. O PyVRP apresentou um desempenho semelhante em termos de viabilidade temporal, embora com um tempo de deslocação aproximadamente 23% superior, compensado por uma distribuição de carga mais equilibrada entre profissionais.

O algoritmo genético (DEAP) apresentou um tempo de deslocação intermédio (98 unidades), mas revelou limitações significativas no cumprimento das janelas temporais (41.7%), resultado da simplificação da sua formulação, que não incorpora explicitamente todas as restrições do VRPTW. O valor de desvio-padrão igual a zero não representa uma solução ótima, mas sim um artefacto da estrutura de rota única utilizada, que concentra todas as visitas num único percurso.

6.6.4 Síntese do Módulo 5

Os resultados obtidos demonstram que a métrica unificada permite comparar de forma consistente soluções geradas por diferentes solvers de otimização. Entre as abordagens avaliadas, o OR-Tools apresentou o melhor desempenho global, combinando o menor tempo de deslocação com o cumprimento integral das janelas temporais, justificando a sua adoção como solver principal do protótipo.

A avaliação confirmou igualmente o correto funcionamento do mecanismo de rolling horizon, capaz de identificar visitas em elevado risco de cancelamento e desencadear automaticamente processos de re-otimização quando necessário.

As limitações identificadas nos módulos anteriores mantêm-se válidas, nomeadamente a concentração de carga após a re-otimização parcial e a ausência de restrições explícitas de competências profissionais na instância reduzida.

6.7 Pipeline Integrado

O protótipo desenvolvido integra os cinco módulos anteriormente apresentados num pipeline operacional único, implementado através do ficheiro *main.py*. Esta integração materializa a arquitetura definida no Capítulo 5 e permite executar, de forma sequencial e reproduzível, todas as fases do processo: carregamento e pré-processamento dos dados, previsão de cancelamentos e procura, otimização de rotas, re-otimização dinâmica e avaliação dos resultados.

6.7.1 Estrutura do pipeline

A execução do pipeline encontra-se organizada em cinco etapas sequenciais.

Na primeira etapa, o Módulo 1 procede ao carregamento e pré-processamento dos dados. São importadas as instâncias HHCRSP e o dataset Healthcare No-Shows, sendo posteriormente aplicadas as transformações de limpeza e preparação definidas durante a fase de compreensão e preparação dos dados. O objeto de pré-processamento gerado nesta fase é reutilizado nos módulos subsequentes, garantindo consistência entre os dados utilizados no treino e na inferência.

Na segunda etapa, correspondente ao Módulo 2, são ajustados os modelos de previsão de cancelamentos e de previsão de procura. O modelo XGBoost é utilizado como classificador principal de cancelamentos, enquanto o modelo de procura fornece estimativas da atividade operacional futura.

Na terceira etapa, referente ao Módulo 3, são gerados planos de rotas através das três abordagens de otimização consideradas: OR-Tools, PyVRP e DEAP. As soluções produzidas são convertidas para um formato comum, permitindo a sua posterior avaliação através de métricas unificadas.

Na quarta etapa, o Módulo 4 aplica o mecanismo de *rolling horizon*, utilizando as probabilidades de cancelamento previstas para identificar visitas suscetíveis de ausência. Sempre que o risco estimado ultrapassa o limiar definido, é desencadeado um processo de re-otimização parcial do plano operacional.

Finalmente, na quinta etapa, o Módulo 5 calcula as métricas de desempenho das soluções geradas e produz os resultados comparativos finais.

Para aumentar a flexibilidade do sistema, o bloco principal da aplicação identifica automaticamente uma instância disponível na diretório de instâncias de testes, eliminando a

necessidade de configuração manual dos caminhos dos ficheiros e facilitando a reutilização do pipeline com diferentes instâncias HHCRSP.

6.7.2 Resultados da execução integrada

O pipeline foi executado com sucesso sobre a instância HHCRSP_10_3_100_0.4_C_C, composta por 10 pacientes e 3 profissionais, utilizando simultaneamente o dataset Healthcare No-Shows com 71 959 registos válidos.

Os resultados obtidos confirmam a consistência entre a execução integrada e os resultados previamente obtidos nos notebooks de validação individual dos módulos.

Tabela 24 - Resultados da execução integrada do pipeline

Módulo	Resultado principal
M1 – Pré-processamento	71 959 registos válidos; 16 features após codificação.
M2 – Previsão de cancelamentos	XGBoost: AUC-ROC = 0.605 ± 0.005
M3 – Otimização de rotas	OR-Tools, PyVRP e DEAP executados com sucesso.
M4 – Rolling Horizon	0 re-otimizações realizadas na execução integrada.
M5 – Avaliação	Métricas calculadas e relatório comparativo gerado.

Os resultados apresentados na Tabela 24 confirmam a execução bem-sucedida de todos os módulos do protótipo no contexto de um pipeline integrado. Os valores obtidos mantiveram-se consistentes com os resultados reportados nas secções anteriores, demonstrando que a integração dos componentes não introduziu alterações no comportamento do sistema nem inconsistências na transferência de dados entre módulos.

6.7.3 Síntese da integração

A execução integrada do protótipo confirmou a correta articulação entre os cinco módulos desenvolvidos ao longo do projeto. O pipeline executou de forma sequencial e sem erros, assegurando a transferência consistente de dados entre as diferentes etapas, desde o pré-processamento dos dados até à avaliação final das soluções geradas.

A integração validou igualmente a arquitetura proposta, demonstrando que os modelos de previsão, os mecanismos de otimização e re-otimização e o sistema de avaliação podem operar

de forma coordenada num único fluxo de execução. Os resultados produzidos mantiveram-se consistentes com aqueles obtidos durante a validação individual dos módulos, não tendo sido identificadas regressões funcionais decorrentes da integração.

Conclui-se, assim, que o protótipo desenvolvido constitui uma solução funcional e reproduzível para apoio ao planeamento e replaneamento de serviços domiciliários. As limitações identificadas ao longo da implementação, nomeadamente as relacionadas com a distribuição da carga de trabalho entre profissionais, a modelação de competências na re-otimização parcial e as restrições impostas pelos datasets disponíveis, são analisadas e discutidas criticamente no capítulo seguinte.

6.8 Considerações de Implementação

6.8.1 Reprodutibilidade

A reprodutibilidade constitui um requisito fundamental do protótipo desenvolvido. Para garantir a consistência dos resultados obtidos, todas as fontes de aleatoriedade presentes no pipeline foram controladas através da definição de um valor fixo do gerador de números aleatórios (*random_seed = 42*), configurada centralmente e utilizada em todos os módulos. Esta parametrização abrange a inicialização dos modelos de aprendizagem automática, a divisão dos dados em conjuntos de treino e teste, os procedimentos de validação cruzada e a geração das populações iniciais dos algoritmos genéticos.

Desta forma, qualquer investigador que utilize os mesmos datasets, dependências de software e ficheiros de configuração poderá reproduzir integralmente os resultados apresentados nesta dissertação, em conformidade com o requisito não funcional RNF05.

6.8.2 Limitações da matriz de distâncias

A modelação das deslocações entre pacientes foi realizada através de uma matriz de distâncias euclidianas calculada a partir das coordenadas geográficas disponibilizadas pelas instâncias HHCRSP. Esta abordagem constitui uma simplificação da realidade operacional, uma vez que não considera a rede rodoviária, as condições de trânsito ou outras restrições geográficas que influenciam os tempos reais de deslocação.

A utilização de distâncias euclidianas é, contudo, uma prática comum em estudos académicos de problemas de otimização de rotas, permitindo a comparação objetiva entre algoritmos e reduzindo a dependência de fontes externas de dados. Em contexto operacional, esta matriz poderia ser substituída por tempos de viagem obtidos através de serviços de cálculo de rotas, como sistemas baseados em OpenStreetMap ou APIs de navegação equivalentes.

6.8.3 Gestão de dados e conformidade com o RGPD

Os datasets utilizados ao longo do desenvolvimento do protótipo são de acesso público e não contêm informação pessoal identificável dos indivíduos representados. Consequentemente, não foram tratados dados pessoais sensíveis nem efetuadas operações suscetíveis de comprometer a privacidade dos utilizadores.

Todos os ficheiros de dados foram armazenados localmente durante a execução dos testes, não sendo transmitidos para serviços externos nem sujeitos a processamento em plataformas de terceiros. Esta abordagem assegura a conformidade com os requisitos RNF03 e RNF04 e está alinhada com os princípios do Regulamento Geral sobre a Proteção de Dados (RGPD), nomeadamente os princípios da minimização de dados e da proteção da privacidade.

7 Avaliação e Resultados

No Capítulo 6, a validação dos módulos foi realizada recorrendo maioritariamente a instâncias HHCRSP de pequena dimensão (10 pacientes), adequadas para verificar a correção da implementação, validar a integração entre componentes e identificar eventuais problemas técnicos durante o desenvolvimento do protótipo.

Concluída essa fase de validação funcional, procede-se neste capítulo à avaliação experimental do sistema. O objetivo consiste em analisar o desempenho dos algoritmos implementados em instâncias de maior dimensão, mais representativas de cenários operacionais reais, avaliando a sua capacidade de escalabilidade e a qualidade das soluções produzidas.

Para este efeito, foram selecionadas instâncias do dataset HHCRSP 2021 com 25, 50 e 75 pacientes. A escolha destas dimensões resulta de um compromisso entre escalabilidade do problema e viabilidade computacional da avaliação experimental, uma vez que configurações de maior dimensão implicariam um aumento substancial do tempo de execução, considerando a avaliação comparativa de múltiplos solvers, a aplicação de métricas unificadas e a execução do mecanismo de re-otimização.

Adicionalmente, importa salientar que o objetivo deste trabalho não consiste na obtenção de resultados de estado da arte para o problema VRPTW, mas sim na validação da arquitetura proposta e na análise do comportamento relativo dos módulos desenvolvidos. Neste contexto, as instâncias selecionadas permitem avaliar de forma progressiva o impacto do aumento da complexidade do problema, garantindo simultaneamente a reprodutibilidade e eficiência dos ensaios experimentais.

As secções seguintes apresentam a metodologia de avaliação, os resultados experimentais obtidos e a respetiva discussão crítica.

7.1 Metodologia de avaliação

A avaliação experimental do protótipo foi realizada através da execução dos três métodos de otimização implementados (OR-Tools, PyVRP e DEAP) sobre instâncias do dataset HHCRSP 2021 com diferentes dimensões. Esta fase complementa a validação efetuada durante a implementação, permitindo analisar o comportamento dos algoritmos em cenários progressivamente mais complexos e avaliar a sua capacidade de escalabilidade.

Para cada dimensão (25, 50 e 75 pacientes) foram considerados três cenários experimentais, correspondentes a diferentes combinações entre a distribuição geográfica dos pacientes e a distribuição das janelas temporais de atendimento. A nomenclatura utilizada segue a convenção utilizada no dataset HHCRSP 2021:

- C_C – pacientes geograficamente agrupados (Clustered) e janelas temporais agrupadas (Clustered);
- R_C – pacientes distribuídos aleatoriamente (Random) e janelas temporais agrupadas (Clustered);
- C_RC – pacientes geograficamente agrupados (Clustered) e janelas temporais de configuração mista (Random-Clustered).

A utilização destes diferentes cenários permite avaliar a robustez dos algoritmos perante distintas condições operacionais e reduzir o impacto das características específicas de uma única instância nos resultados obtidos.

Para cada dimensão foram utilizadas três instâncias, uma por cenário experimental, perfazendo um total de nove instâncias avaliadas.

Tabela 25 - Instâncias HHCRSP utilizadas no processo de avaliação

Dimensão	Cenário	Instância	Pacientes	Profissionais
Pequena	C_C	HHCRSP_25_5_100_0.4_C_C	25	5
Pequena	R_C	HHCRSP_25_5_100_0.4_R_C	25	5
Pequena	C_RC	HHCRSP_25_5_100_0.4_C_RC	25	5
Média	C_C	HHCRSP_50_10_100_0.4_C_C	50	10
Média	R_C	HHCRSP_50_10_100_0.4_R_C	50	10
Média	C_RC	HHCRSP_50_10_100_0.4_C_RC	50	10
Grande	C_C	HHCRSP_75_15_100_0.4_C_C	75	15

Dimensão	Cenário	Instância	Pacientes	Profissionais
Grande	R_C	HHCRSP_75_15_100_0.4_R_C	75	15
Grande	C_RC	HHCRSP_75_15_100_0.4_C_RC	75	15

O tempo máximo de execução foi fixado em 60 segundos por solver e por instância. Este valor é superior ao utilizado durante a fase de desenvolvimento, permitindo uma exploração mais aprofundada do espaço de soluções nas instâncias de maior dimensão.

A comparação entre os diferentes algoritmos foi realizada utilizando a função de avaliação unificada apresentada na Secção 6.6.3, garantindo que todas as soluções são analisadas segundo as mesmas métricas de desempenho. Em particular, foram consideradas métricas relacionadas com o tempo total de deslocação, equilíbrio da carga de trabalho entre profissionais e cumprimento das janelas temporais, permitindo uma comparação direta e consistente entre os diferentes métodos de otimização.

7.2 Escalabilidade dos solvers

7.2.1 OR-Tools

Os resultados obtidos demonstram que o OR-Tools manteve o cumprimento integral das janelas temporais em todas as instâncias avaliadas, independentemente da dimensão do problema ou da configuração espacial dos pacientes. Este comportamento confirma os resultados observados durante a fase de implementação e valida a adequação do solver para problemas de planeamento domiciliário sujeitos a restrições temporais rigorosas.

À medida que a dimensão das instâncias aumentou, verificou-se um crescimento gradual do tempo total de deslocação, passando de valores compreendidos entre 223 e 265 unidades nas instâncias de 25 pacientes para valores entre 558 e 659 unidades nas instâncias de 75 pacientes. Esta evolução é consistente com o aumento do número de visitas a realizar e da complexidade combinatória inerente ao problema.

Em contrapartida, observou-se uma degradação progressiva do equilíbrio de carga entre profissionais. O desvio-padrão do número de visitas atribuídas aumentou de valores entre 3.19 e 4.87 nas instâncias de 25 pacientes para valores próximos de 5 nas instâncias de 75 pacientes. Este resultado era expectável, uma vez que a função objetivo do OR-Tools privilegia a minimização do custo global das rotas, não incorporando explicitamente critérios de equidade na distribuição do trabalho.

Importa ainda salientar que o solver atingiu sistematicamente o limite temporal de 60 segundos definido para a experiência. Este comportamento sugere que o espaço de pesquisa permanece suficientemente vasto para justificar a continuação da exploração de soluções. Apesar disso, todas as soluções produzidas foram viáveis e respeitaram integralmente as restrições temporais do problema.

7.2.2 PyVRP

O PyVRP apresentou um comportamento consistente ao longo de todas as instâncias avaliadas. Embora tenha produzido tempos de deslocação superiores aos obtidos pelo OR-Tools, manteve igualmente uma taxa de cumprimento das janelas temporais de 100% em todos os cenários considerados.

A principal diferença entre os dois solvers observou-se ao nível da distribuição da carga de trabalho. O PyVRP produziu sistematicamente soluções mais equilibradas, apresentando valores de desvio-padrão inferiores aos do OR-Tools em todas as dimensões testadas. Este resultado sugere que a estratégia de pesquisa adotada pelo PyVRP favorece uma distribuição mais homogénea das visitas pelos profissionais, mesmo quando tal implica um aumento moderado do custo total de deslocação.

Os resultados obtidos evidenciam, assim, um compromisso entre eficiência operacional e equidade de distribuição. Enquanto o OR-Tools privilegia a minimização do tempo total de deslocação, o PyVRP tende a gerar soluções mais equilibradas do ponto de vista organizacional. Esta característica pode tornar o PyVRP particularmente interessante em contextos onde a distribuição equitativa da carga de trabalho constitui um requisito operacional relevante.

7.2.3 DEAP

O algoritmo genético implementado com recurso à biblioteca DEAP apresentou tempos de execução substancialmente inferiores aos dos solvers especializados, concluindo a resolução das instâncias em menos de dois segundos, mesmo para os problemas de maior dimensão.

Contudo, esta vantagem computacional foi acompanhada por uma degradação acentuada da qualidade das soluções produzidas. A taxa de cumprimento das janelas temporais diminuiu progressivamente com o aumento da dimensão das instâncias, atingindo valores inferiores a 10% nas instâncias de 75 pacientes. Estes resultados demonstram que a representação e função de avaliação atualmente implementadas não são adequadas para problemas VRPTW com restrições temporais exigentes.

A principal causa deste comportamento reside na arquitetura simplificada adotada para o algoritmo genético. A função de avaliação considera essencialmente a distância percorrida, não penalizando adequadamente as violações das janelas temporais nem contemplando a divisão das visitas por múltiplas rotas. Consequentemente, o algoritmo tende a privilegiar percursos geograficamente curtos, mesmo quando estes são operacionalmente inviáveis.

Um exemplo ilustrativo desta limitação ocorreu na instância R_C de 25 pacientes, na qual o DEAP obteve um tempo de deslocação inferior ao do OR-Tools. Embora este resultado possa sugerir uma solução mais eficiente, trata-se de um efeito artificial decorrente da inexistência de restrições temporais efetivas na função objetivo. Na prática, a solução produzida viola um número significativo de janelas temporais e não pode ser considerada operacionalmente válida.

Em síntese, os resultados demonstram que o DEAP constitui uma abordagem promissora do ponto de vista da rapidez computacional, mas requer uma reformulação substancial da representação das soluções e da função de avaliação antes de poder competir com solvers especializados em problemas de routing com restrições temporais.

7.2.4 Comparação global

As Tabelas 26, 27 e 28 apresentam os resultados obtidos para as instâncias de 25, 50 e 75 pacientes, respectivamente, considerando os três cenários experimentais (C_C, R_C e C_RC). Esta organização permite comparar o desempenho dos algoritmos em diferentes condições operacionais, preservando as particularidades de cada cenário e evitando a agregação de resultados que poderiam ocultar diferenças relevantes.

Os resultados obtidos para as instâncias de 25 pacientes encontram-se apresentados na Tabela 26.

Tabela 26 - Resultados obtidos para as instâncias de 25 pacientes

Cenário	Solver	Deslocação media	Std carga media	Cumprimento TW% média	Tempo(s)
C_C	OR-Tools	223	4.12	100%	60.06
C_C	PyVRP	286	2.49	100%	60.87
C_C	DEAP	247	0.00*	18.5%	0.88
R_C	OR-Tools	263	4.87	100%	60.00
R_C	PyVRP	286	3.30	100%	60.00
R_C	DEAP	207	0.00	25.9%	0.54
C_RC	OR-Tools	265	3.19	100%	60.00
C_RC	PyVRP	317	3.03	100%	59.99

Cenário	Solver	Deslocação media	Std carga media	Cumprimento TW% média	Tempo(s)
C_RC	DEAP	276	0.00	18.5%	0.50

*O valor de equilíbrio igual a 0.00 resulta da utilização de uma única rota, não representando uma distribuição equilibrada das visitas pelos profissionais.

A Tabela 27 apresenta os resultados obtidos para as instâncias de 50 pacientes.

Tabela 27 - Resultados obtidos para as instâncias de 50 pacientes

Cenário	Solver	Deslocação media	Std carga media	Cumprimento TW% média	Tempo(s)
C_C	OR-Tools	366	4.13	100%	60.00
C_C	PyVRP	444	2.95	100%	59.99
C_C	DEAP	388	0.00	11.5%	0.65
R_C	OR-Tools	403	4.39	100%	60.00
R_C	PyVRP	452	2.13	100%	59.99
R_C	DEAP	408	0.00	7.7%	0.69
C_RC	OR-Tools	406	4.53	100%	60.00
C_RC	PyVRP	468	2.69	100%	59.99
C_RC	DEAP	419	0.00	9.6%	0.68

Por fim, a Tabela 28 apresenta os resultados obtidos para as instâncias de 75 pacientes.

Tabela 28 - Resultados obtidos para as instâncias de 75 pacientes

Cenário	Solver	Deslocação media	Std carga media	Cumprimento TW% média	Tempo(s)
C_C	OR-Tools	659	4.92	100%	60.00
C_C	PyVRP	766	3.62	100%	60.00
C_C	DEAP	766	0.00	9.1%	0.91
R_C	OR-Tools	584	4.93	100%	60.00
R_C	PyVRP	680	3.60	100%	60.00
R_C	DEAP	760	0.00	5.2%	1.12
C_RC	OR-Tools	558	5.08	100%	60.01
C_RC	PyVRP	672	3.50	100%	60.00
C_RC	DEAP	781	0.00	5.2%	1.05

A análise conjunta dos resultados permite retirar várias conclusões. Observa-se que o OR-Tools apresentou o desempenho mais consistente ao longo de todas as instâncias avaliadas. Em todos os cenários garantiu o cumprimento integral das janelas temporais (100%) e, na maioria dos casos, obteve as menores distâncias totais de deslocação.

O PyVRP apresentou igualmente uma taxa de cumprimento de 100% das janelas temporais em todas as instâncias, evidenciando elevada robustez na resolução do problema. No entanto, produziu sistematicamente soluções com maior distância total de deslocação quando comparado com o OR-Tools, embora mantendo um melhor equilíbrio na distribuição da carga de trabalho entre profissionais.

O DEAP distinguiu-se pelo reduzido tempo de execução, inferior a dois segundos em todas as instâncias. Contudo, a qualidade das soluções obtidas revelou-se significativamente inferior, apresentando baixas taxas de cumprimento das janelas temporais e, em diversos casos, concentrando a maioria das visitas numa única rota. Consequentemente, o valor do desvio

padrão da carga igual a zero não representa uma distribuição equilibrada do trabalho, mas sim a inexistência de repartição das visitas pelos diferentes profissionais disponíveis.

À medida que a dimensão das instâncias aumenta, verifica-se um crescimento expectável da distância total percorrida pelos três algoritmos. Apesar desse aumento, o OR-Tools e o PyVRP mantiveram um desempenho estável relativamente ao cumprimento das restrições temporais, enquanto o DEAP evidenciou uma degradação progressiva da qualidade das soluções.

Em síntese, os resultados experimentais indicam que o OR-Tools constitui a solução mais adequada para o problema de otimização de rotas considerado neste trabalho, combinando menores distâncias percorridas com o cumprimento integral das restrições operacionais. O PyVRP apresentou um desempenho competitivo, surgindo como uma alternativa válida, enquanto o DEAP demonstrou maior adequação para fins exploratórios ou de investigação em metaheurísticas do que para utilização como solver principal neste contexto.

As Figuras 13, 14 e 15 apresentam a representação gráfica comparativa dos três solvers para as diferentes instâncias avaliadas, organizadas por métrica de avaliação.

A Figura 13 ilustra a distância total percorrida em cada instância. Observa-se que o OR-Tools obtém consistentemente os valores mais baixos, seguido pelo PyVRP, enquanto o DEAP apresenta um comportamento menos previsível, aproximando-se ou superando o PyVRP nas instâncias de maior dimensão.

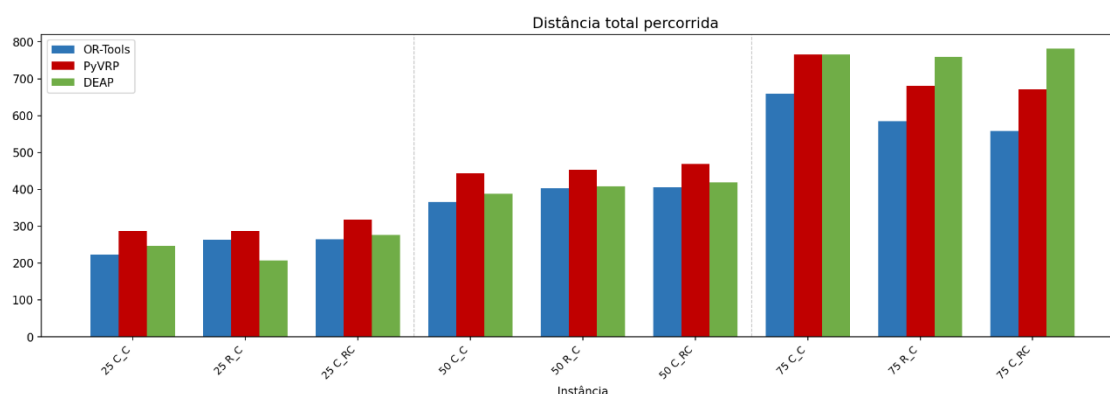


Figura 13 - Distância total percorrida

A Figura 14 apresenta o equilíbrio de carga entre profissionais, medido pelo desvio padrão. O DEAP regista valores de zero em todas as instâncias, não por apresentar melhor equilíbrio, mas por concentrar as visitas numa única rota, conforme referido na Tabela 26. O OR-Tools apresenta os valores mais elevados de desvio padrão, o que, neste contexto, reflete a utilização efetiva de múltiplas rotas com distribuição real de carga.

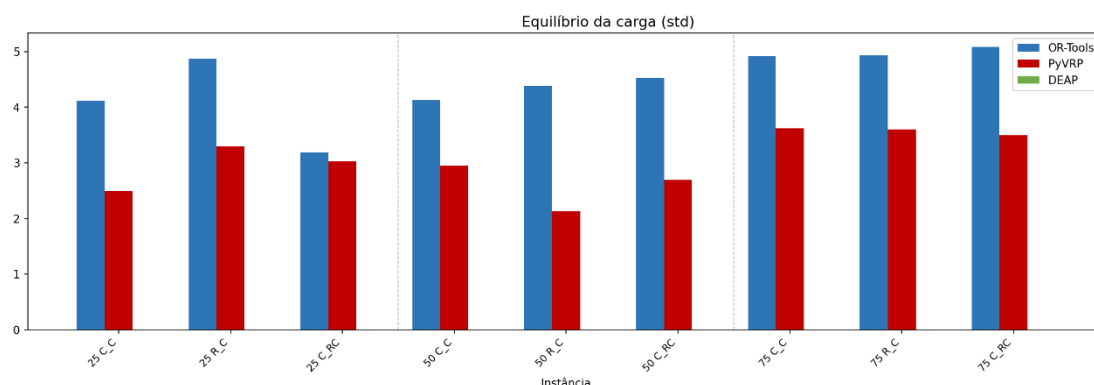


Figura 14 - Equilíbrio de carga (std)

A Figura 15 evidencia a taxa de cumprimento das janelas temporais. O OR-Tools e o PyVRP atingem 100% em todas as instâncias, enquanto o DEAP apresenta taxas substancialmente inferiores, com degradação progressiva à medida que a dimensão do problema aumenta.

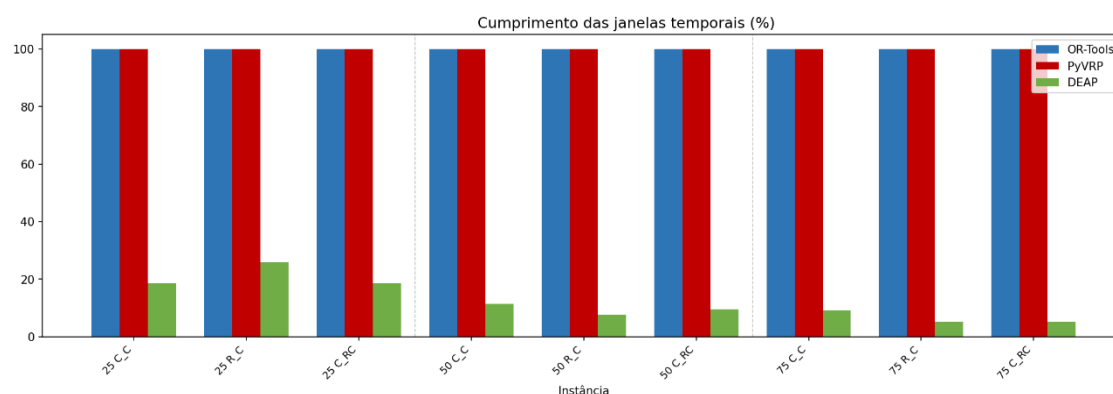


Figura 15 - Cumprimento das janelas temporais (%)

7.2.5 Síntese

Os resultados obtidos permitem concluir que os três solvers apresentam perfis de desempenho distintos, adequados a objetivos operacionais diferentes. Observa-se uma clara hierarquia em termos de robustez e qualidade das soluções, bem como diferenças relevantes ao nível do compromisso entre eficiência, equidade e tempo computacional.

O OR-Tools destaca-se como a solução mais robusta e consistente, assegurando a viabilidade operacional das soluções em todas as instâncias e apresentando o melhor desempenho global no contexto analisado, o que justifica a sua adoção como componente principal do protótipo.

O PyVRP evidencia um perfil mais equilibrado, privilegiando a equidade na distribuição da carga de trabalho, o que pode ser relevante em contextos onde este fator tenha prioridade sobre a minimização do custo total das rotas.

O DEAP, embora eficiente do ponto de vista computacional, revela limitações significativas na qualidade das soluções produzidas quando confrontado com restrições temporais exigentes, o que restringe a sua utilização a contextos exploratórios ou de apoio ao desenvolvimento de metaheurísticas.

Em síntese, a escolha do solver deve ser orientada pelos objetivos operacionais do problema: eficiência e viabilidade operacional favorecem o OR-Tools, equidade favorece o PyVRP, e rapidez computacional favorece o DEAP.

7.3 Avaliação do mecanismo de Re-otimização

7.3.1 Metodologia

A avaliação do mecanismo de *rolling horizon* foi realizada sobre as nove instâncias utilizadas na análise de escalabilidade dos solvers, correspondentes a três dimensões (25, 50 e 75 pacientes) e três configurações espaciais e temporais distintas (C_C, R_C e C_RC). Adicionalmente, foi efetuada uma análise de sensibilidade ao limiar de cancelamento, considerando os valores de 0.3, 0.4 e 0.5.

Para cada combinação de instância e limiar, o modelo XGBoost estimou a probabilidade de cancelamento das visitas agendadas. As visitas cuja probabilidade excedia o limiar definido foram assinaladas como visitas de elevado risco e removidas do plano inicial, desencadeando posteriormente um processo de re-otimização parcial através do OR-Tools.

As características das visitas foram geradas a partir de amostras do conjunto de teste do dataset *Healthcare No-Shows*, seguindo a mesma metodologia adotada durante a fase de implementação. Consequentemente, o perfil de risco identificado depende exclusivamente das características dos pacientes considerados e não da estrutura geográfica ou temporal das instâncias HHCRSP. Por esta razão, para uma mesma dimensão, os resultados obtidos foram idênticos nas três configurações avaliadas (C_C, R_C e C_RC).

7.3.2 Sensibilidade ao limiar de cancelamento

A Tabela 29 apresenta os resultados obtidos para os diferentes limiares de cancelamento considerados.

Tabela 29 - Sensibilidade ao limiar de cancelamento

Dimensão	Limiar	Visitas em risco	% risco	Remanescentes	Re-otimizações
25 pac.	0.3	13	52.0%	12	1

Dimensão	Limiar	Visitas em risco	% risco	Remanescentes	Re-otimizações
25 pac.	0.4	3	12.0%	22	1
25 pac.	0.5	1	4.0%	24	1
50 pac.	0.3	22	44.0%	28	1
50 pac.	0.4	7	14.0%	43	1
50 pac.	0.5	3	6.0%	47	1
75 pac.	0.3	37	49.3%	38	1
75 pac.	0.4	10	13.3%	65	1
75 pac.	0.5	3	4.0%	72	1

Os resultados evidenciam um comportamento consistente do mecanismo de *rolling horizon* em todas as dimensões avaliadas. Em todos os cenários foi desencadeado um único processo de re-otimização, demonstrando que o mecanismo responde adequadamente à identificação de visitas com elevado risco de cancelamento.

O limiar de 0.3 revelou-se particularmente conservador, classificando aproximadamente metade das visitas como estando em situação de elevado risco. Embora esta abordagem maximize a capacidade de antecipação de potenciais cancelamentos, conduz simultaneamente à remoção de uma proporção significativa das visitas inicialmente planeadas, aumentando a perturbação operacional e reduzindo a utilização dos recursos disponíveis.

Por outro lado, o limiar de 0.5 apresentou um comportamento mais permissivo, identificando apenas entre 4% e 6% das visitas como estando em situação de elevado risco. Apesar de minimizar alterações ao plano operacional, este valor poderá reduzir a capacidade preventiva do sistema ao excluir visitas cuja probabilidade de cancelamento seja moderada, mas ainda relevante.

O limiar intermédio de 0.4 apresentou o comportamento mais equilibrado, identificando aproximadamente 12% a 14% das visitas como estando em situação de elevado risco. Este resultado sugere que o valor adotado durante a implementação constitui um compromisso adequado entre estabilidade operacional e capacidade de antecipação de cancelamentos, permitindo intervir apenas sobre um subconjunto limitado de visitas potencialmente problemáticas.

A Figura 16 apresenta a evolução da percentagem de visitas classificadas como de elevado risco para os diferentes limiares considerados, bem como o número de re-otimizações realizadas em cada cenário.

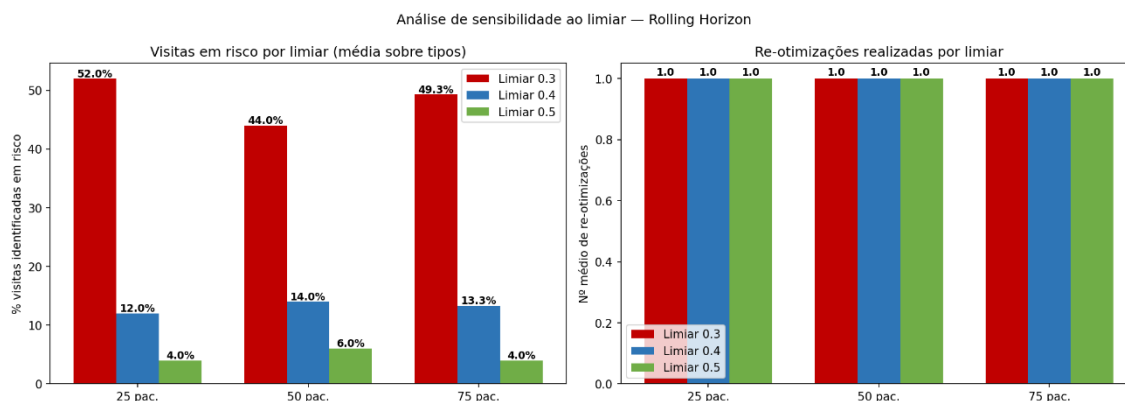


Figura 16 - Sensibilidade ao limiar de cancelamento

7.3.3 Robustez do mecanismo

Um dos resultados mais relevantes desta avaliação é a consistência do comportamento do mecanismo de *rolling horizon* ao longo de todas as combinações analisadas. Nos 27 cenários avaliados (9 instâncias × 3 limiares), foi desencadeado exatamente um processo de re-otimização sempre que foram identificadas visitas de elevado risco, independentemente da dimensão ou da configuração da instância.

Este comportamento evidencia que o mecanismo mantém a sua funcionalidade à medida que a dimensão do problema aumenta, reproduzindo nas instâncias de 25, 50 e 75 pacientes o padrão de funcionamento previamente observado durante a fase de implementação com instâncias de 10 pacientes.

Verificou-se igualmente que o tempo de re-otimização atingiu sistematicamente o limite temporal configurado de 60 segundos. Este resultado indica que o OR-Tools utilizou integralmente o tempo disponível para exploração do espaço de soluções, fenómeno expectável em problemas de otimização combinatoria com restrições temporais. Apesar desta limitação temporal, todas as soluções produzidas permaneceram viáveis e respeitaram as restrições impostas pelas instâncias, demonstrando a adequação da abordagem para os cenários considerados.

7.3.4 Síntese

Os resultados obtidos demonstram que o mecanismo de *rolling horizon* opera de forma consistente em instâncias de diferentes dimensões, sendo capaz de identificar visitas de elevado risco e desencadear processos de re-otimização seletiva sempre que necessário.

A análise de sensibilidade ao limiar de cancelamento permitiu ainda avaliar o impacto deste parâmetro no comportamento do sistema. Os resultados sugerem que o valor de 0.4 constitui um compromisso operacional razoável entre estabilidade do plano e capacidade de intervenção preventiva. Limiares inferiores conduzem à identificação de um número excessivo de visitas de elevado risco, aumentando significativamente o número de alterações ao plano inicial, enquanto limiares superiores reduzem a capacidade preventiva do sistema ao restringirem a detecção de visitas potencialmente problemáticas.

Em conjunto, estes resultados validam o funcionamento do mecanismo de re-otimização proposto e demonstram a sua capacidade para apoiar o reajustamento dinâmico dos planos de visita perante potenciais cancelamentos, contribuindo para uma gestão mais flexível dos serviços domiciliários.

7.4 Avaliação Comparativa Final

7.4.1 Comparação de solvers em instâncias com maior escala

A avaliação comparativa final integrou os resultados obtidos nos três estudos de escalabilidade realizados. Os resultados detalhados para cada dimensão e tipo de instância foram apresentados nas Figuras 14, 15 e 16, permitindo identificar padrões consistentes no desempenho dos três solvers avaliados.

De forma transversal a todas as dimensões e configurações testadas, o OR-Tools apresentou o melhor desempenho global, combinando o menor tempo de deslocação com o cumprimento integral das janelas temporais. Estes resultados confirmam a adequação deste solver como solução principal para o protótipo desenvolvido.

O PyVRP evidenciou um comportamento consistente ao longo de todas as instâncias avaliadas, produzindo soluções com melhor equilíbrio de carga entre profissionais, embora à custa de tempos de deslocação superiores aos obtidos pelo OR-Tools. Este resultado evidencia o compromisso entre eficiência operacional e distribuição equitativa da carga de trabalho.

Por sua vez, o DEAP apresentou uma degradação progressiva do cumprimento das janelas temporais à medida que a dimensão das instâncias aumentou. Apesar dos reduzidos tempos de execução observados, as elevadas taxas de violação das restrições temporais limitam a sua aplicabilidade em contextos operacionais reais, sendo mais adequado como referência experimental ou para cenários simplificados sem restrições temporais críticas.

7.4.2 Avaliação do pipeline completo

A avaliação do pipeline integrado comparou o plano gerado pelo OR-Tools sem re-otimização (*baseline*) com o plano ajustado após aplicação do mecanismo de *rolling horizon*, utilizando um

limiar de cancelamento de 0.4. A comparação foi realizada para as instâncias de 25, 50 e 75 pacientes, sendo os resultados apresentados na Tabela 30.

Tabela 30 - Comparação entre plano base e plano re-otimizado

Dimensão	Tipo	Plano	Visitas	Deslocação	Std carga	TW%
25 pac.	C_C	Baseline	25	223	4.12	100%
25 pac.	C_C	Re-otimizado	22	215	5.28	100%
25 pac.	R_C	Baseline	25	263	4.87	100%
25 pac.	R_C	Re-otimizado	22	224	4.88	100%
25 pac.	C_RC	Baseline	25	265	3.19	100%
25 pac.	C_RC	Re-otimizado	22	256	3.72	100%
50 pac.	C_C	Baseline	50	366	4.13	100%
50 pac.	C_C	Re-otimizado	43	342	4.48	100%
50 pac.	R_C	Baseline	50	403	4.39	100%
50 pac.	R_C	Re-otimizado	43	348	4.18	100%
50 pac.	C_RC	Baseline	50	406	4.53	100%

Dimensão	Tipo	Plano	Visitas	Deslocação	Std carga	TW%
50 pac.	C_RC	Re-otimizado	43	352	4.22	100%
75 pac.	C_C	Baseline	75	659	4.92	100%
75 pac.	C_C	Re-otimizado	65	594	4.79	100%
75 pac.	R_C	Baseline	75	584	4.93	100%
75 pac.	R_C	Re-otimizado	65	536	4.88	100%
75 pac.	C_RC	Baseline	75	558	5.08	100%
75 pac.	C_RC	Re-otimizado	65	514	5.23	100%

Os resultados demonstram que o mecanismo de rolling horizon produziu uma redução consistente do tempo total de deslocação em todas as instâncias avaliadas. As reduções por instância variaram entre 3,4% (25 C_RC) e 14,8% (25 R_C), com as instâncias de 50 e 75 pacientes a registarem reduções entre 6,6% e 13,5% e entre 7,9% e 9,9%, respetivamente. Esta melhoria resulta da remoção das visitas identificadas como de elevado risco de cancelamento e da subsequente re-otimização das rotas para os pacientes remanescentes.

Verifica-se igualmente que a diminuição do tempo de deslocação foi inferior à redução observada no número de visitas — de 25 para 22 pacientes nas instâncias de 25 pac., de 50 para 43 nas de 50 pac., e de 75 para 65 nas de 75 pac. Este comportamento é esperado, uma vez que a remoção de pacientes não elimina proporcionalmente os percursos necessários para garantir a cobertura dos restantes. A reorganização das rotas continua dependente da distribuição

geográfica das visitas remanescentes, podendo exigir deslocações significativas mesmo após a redução da carga de trabalho.

A Figura 17 apresenta a distância total percorrida por instância, evidenciando as reduções obtidas após re-otimização em todos os cenários.

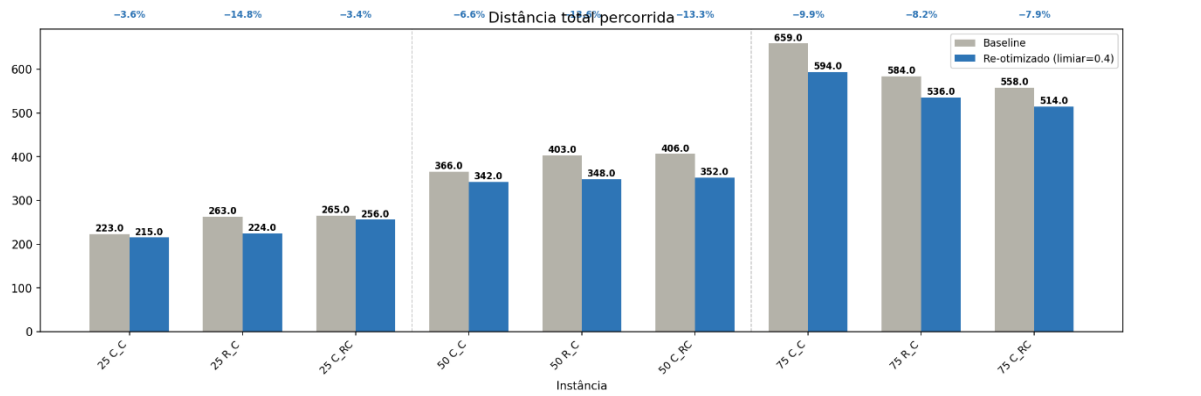


Figura 17 - Distância total percorrida: Baseline vs. Re-otimizado

A Figura 18 mostra que o equilíbrio de carga se manteve globalmente estável, sem alterações expressivas face ao plano original.

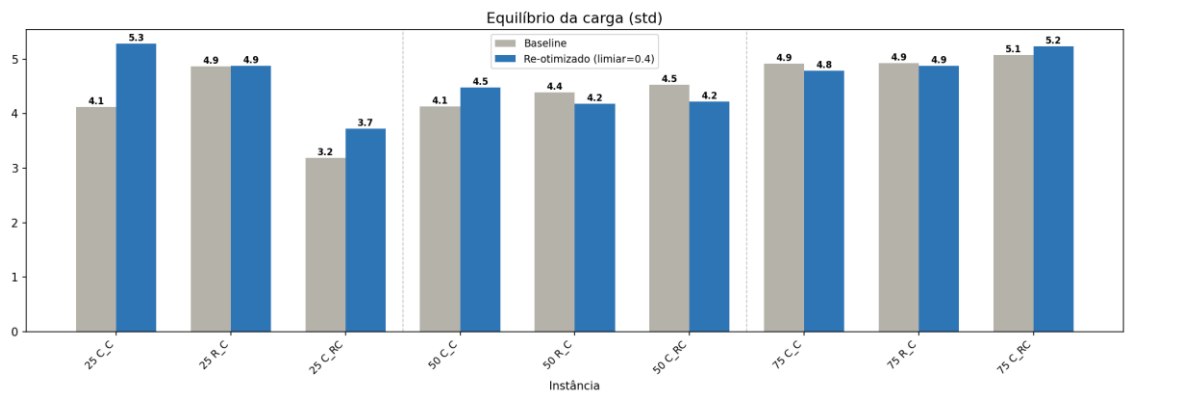


Figura 18 - Equilíbrio da carga (std): Baseline vs. Re-otimizado

A Figura 19 confirma que o cumprimento das janelas temporais permaneceu integral em todas as instâncias, demonstrando que a redução de deslocação não comprometeu a viabilidade operacional das soluções produzidas.

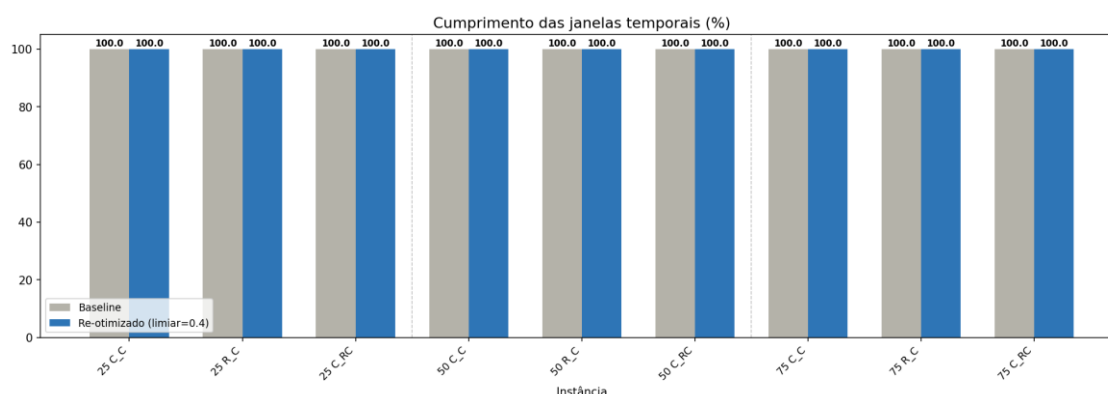


Figura 19 - Cumprimento das janelas temporais (%): Baseline vs. Re-otimizado

7.4.3 Síntese

Os resultados obtidos sugerem que a integração entre o modelo preditivo de cancelamentos e o mecanismo de rolling horizon permite adaptar dinamicamente os planos de visitas domiciliárias sem comprometer a viabilidade operacional das soluções geradas. Em todas as instâncias avaliadas, observou-se uma redução no tempo total de deslocação após a re-otimização, mantendo simultaneamente o cumprimento integral das janelas temporais.

A análise de sensibilidade realizada na Secção 7.3 confirmou que o limiar de 0.4 representa o compromisso mais equilibrado entre a identificação de visitas com elevada probabilidade de cancelamento e a preservação da estabilidade do plano operacional. Valores inferiores tendem a desencadear intervenções excessivas, enquanto valores superiores reduzem a capacidade de antecipação do sistema.

Em conjunto, os resultados da avaliação demonstram que o protótipo desenvolvido é capaz de combinar previsão de cancelamentos, otimização de rotas e re-otimização dinâmica num único processo integrado, contribuindo para a redução de deslocações desnecessárias e para uma utilização mais eficiente dos recursos disponíveis nos serviços de apoio domiciliário.

7.5 Discussão dos Resultados

A avaliação experimental realizada permitiu analisar o comportamento do protótipo em instâncias de diferentes dimensões, desde 10 até 75 pacientes, abrangendo diferentes distribuições geográficas e configurações de janelas temporais. Os resultados obtidos evidenciam a robustez da arquitetura proposta e permitem identificar as principais vantagens e limitações das abordagens avaliadas.

7.5.1 Desempenho dos solvers de otimização

O OR-Tools revelou o melhor desempenho global, combinando o cumprimento integral das janelas temporais com os menores tempos de deslocação nas diferentes dimensões analisadas. Estes resultados confirmam a adequação da estratégia baseada em Routing Solver com Guided Local Search para problemas de planeamento domiciliário com restrições temporais. Apesar de atingir sistematicamente o limite temporal definido, o solver produziu soluções viáveis e operacionalmente adequadas em todos os cenários avaliados.

O PyVRP apresentou um comportamento consistente ao longo de todas as instâncias, evidenciando um compromisso entre eficiência e equidade na distribuição do trabalho. Embora os tempos de deslocação tenham sido superiores aos obtidos pelo OR-Tools, o equilíbrio de carga entre profissionais foi sistematicamente melhor, sugerindo que esta abordagem poderá ser particularmente interessante em contextos onde a distribuição equitativa das visitas constitui um objetivo organizacional relevante.

O DEAP revelou limitações significativas no tratamento de restrições temporais. A degradação progressiva da taxa de cumprimento das janelas temporais com o aumento da dimensão das instâncias demonstra que a representação adotada não é adequada para problemas HHCRSP de dimensão realista. Ainda assim, os tempos de execução reduzidos sugerem potencial para futuras extensões ou para utilização em cenários exploratórios com restrições menos exigentes.

7.5.2 Desempenho do mecanismo de rolling horizon

O mecanismo de re-otimização dinâmica demonstrou um comportamento estável e consistente ao longo de todas as instâncias avaliadas. A identificação de visitas com elevada probabilidade de cancelamento permitiu reduzir o esforço de deslocação necessário para executar os planos remanescentes, preservando simultaneamente o cumprimento integral das janelas temporais.

A análise de sensibilidade realizada ao limiar de cancelamento confirmou que o valor de 0.4 representa um compromisso adequado entre capacidade de deteção e estabilidade operacional. Valores inferiores conduzem a intervenções excessivamente agressivas, enquanto valores superiores reduzem a capacidade do sistema para antecipar potenciais cancelamentos.

7.5.3 Desempenho do modelo preditivo

O classificador de cancelamentos obteve um desempenho moderado, mas consistente com estudos que utilizam dados genéricos de marcações médicas. Embora o valor de AUC-ROC obtido demonstre capacidade preditiva superior à aleatoriedade, os resultados sugerem que a incorporação de variáveis específicas do contexto de apoio domiciliário poderá melhorar significativamente a qualidade das previsões. Ainda assim, o objetivo principal desta investigação consistia em validar a integração entre previsão de cancelamentos e re-otimização dinâmica, objetivo que foi alcançado com sucesso.

7.6 Limitações e Trabalho Futuro

7.6.1 Limitações identificadas

A avaliação experimental permitiu identificar um conjunto de limitações do protótipo que devem ser consideradas na interpretação dos resultados obtidos.

A primeira limitação prende-se com o modelo preditivo de cancelamentos. O classificador XGBoost foi desenvolvido com o dataset Healthcare No-Shows, que não foi recolhido especificamente para o contexto de apoio domiciliário. A ausência de variáveis operacionais relevantes, como o histórico de visitas anteriores, a severidade clínica do paciente ou a relação de continuidade com o profissional de saúde — limita o desempenho preditivo do modelo. O AUC-ROC obtido (0.605) indica capacidade discriminativa superior à aleatoriedade, mas insuficiente para uma classificação de alta precisão.

A segunda limitação refere-se ao mecanismo de re-otimização parcial. A implementação atual remove as visitas em risco da instância e re-otimiza sobre os nós remanescentes sem considerar explicitamente o equilíbrio de carga entre profissionais. Em instâncias com elevado número de visitas em risco, esta abordagem pode concentrar as visitas remanescentes num subconjunto de profissionais, gerando desequilíbrios operacionais. Esta limitação foi documentada na fase de implementação e mantém-se relevante na avaliação com instâncias maiores.

A terceira limitação diz respeito à ausência de restrições de competências na re-otimização parcial. A instância reduzida utilizada durante a re-otimização não preserva as restrições de competências dos profissionais presentes na instância original, tratando todos os recursos como equivalentes. Em cenários operacionais reais, nem todos os profissionais têm competências para realizar todos os tipos de serviço, pelo que esta simplificação pode gerar planos tecnicamente inviáveis.

A quarta limitação relaciona-se com o DEAP e a arquitetura de rota única. A implementação do algoritmo genético considera uma única rota sequencial para todos os pacientes, o que inviabiliza a sua utilização direta em instâncias com múltiplos profissionais e restrições temporais. Esta simplificação foi assumida explicitamente no desenvolvimento do protótipo e limita a comparabilidade do DEAP com os restantes solvers em termos de qualidade de solução.

7.6.2 Trabalho futuro

Com base nas limitações identificadas e nos resultados obtidos, identificam-se as seguintes linhas de trabalho futuro para o desenvolvimento do protótipo.

A recolha ou construção de um dataset específico de apoio domiciliário permitiria ajustar o modelo preditivo com variáveis operacionais relevantes, nomeadamente histórico de visitas, diagnósticos clínicos, perfil de continuidade assistencial e características dos profissionais. Este

dataset permitiria potencialmente aumentar o AUC-ROC do classificador para valores superiores a 0.70, tornando o mecanismo de rolling horizon mais preciso na identificação de visitas em risco.

A integração explícita das restrições de competências dos profissionais no processo de otimização constitui a extensão mais direta e prioritária ao protótipo atual. O parser desenvolvido já carrega integralmente as matrizes de competências e de requisitos de serviço presentes nas instâncias HHCRSP, e foi implementada uma função auxiliar que determina os profissionais elegíveis para cada visita com base nas suas qualificações. A incorporação desta informação no solver OR-Tools permitiria restringir automaticamente a atribuição de visitas a profissionais sem as competências necessárias, garantindo a viabilidade técnica dos planos gerados. Em paralelo, a substituição da duração média de serviço pela duração específica por profissional e tipo de serviço permitiria modelar com maior fidelidade o tempo efetivo de atendimento, aproximando o protótipo de condições operacionais reais.

A integração explícita das restrições de competências na instância reduzida utilizada durante a re-otimização parcial garantiria que os planos ajustados são tecnicamente viáveis, respeitando as qualificações de cada profissional para cada tipo de serviço.

A reformulação do DEAP para suporte a múltiplas rotas, através de uma representação genética baseada em partições de pacientes por profissional e incorporação das restrições temporais na função de avaliação, permitiria uma comparação mais justa com o OR-Tools e o PyVRP, e poderia explorar o potencial do seu tempo de execução reduzido em cenários de re-otimização rápida.

Por fim, a avaliação utilizando dados reais provenientes de organizações de apoio domiciliário e a extensão da avaliação a instâncias de 100 ou mais pacientes, com aumento do limite temporal dos solvers, permitiria caracterizar com maior detalhe o comportamento de escalabilidade do protótipo e identificar eventuais pontos de degradação de qualidade em instâncias de dimensão próxima de cenários operacionais reais.

8 Conclusão

8.1 Síntese do trabalho realizado

Esta dissertação demonstrou a viabilidade de integrar modelos preditivos de aprendizagem automática e algoritmos de otimização de rotas num sistema unificado de apoio à decisão para serviços de apoio ao domicílio. O objetivo central consistiu no desenvolvimento de um protótipo capaz de combinar previsão de cancelamentos, planeamento de rotas e re-otimização dinâmica num único pipeline.

O sistema desenvolvido baseia-se em dados públicos provenientes do Healthcare No-Shows dataset, para a componente preditiva, e do HHCRSP 2021, para a componente de otimização de rotas. A abordagem proposta permite estabelecer a ligação entre a previsão de comportamento dos utentes e a geração de planos operacionais de visitas domiciliárias.

A validação experimental foi realizada através de instâncias de dimensão crescente (10 a 75 pacientes), contemplando diferentes configurações de distribuição geográfica e estrutura temporal, permitindo avaliar o comportamento do sistema em cenários progressivamente mais complexos.

8.2 Resposta aos objetivos

Os objetivos definidos na Secção 1.3 foram globalmente alcançados, evidenciando a viabilidade da abordagem proposta para integração de previsão e otimização em contexto de apoio domiciliário.

O objetivo relativo ao desenvolvimento de modelos preditivos para antecipar padrões de procura foi atingido através da implementação de um classificador de cancelamentos baseado em XGBoost e de um modelo de previsão de procura diária baseado em média móvel. Ambos os modelos demonstraram capacidade preditiva acima do nível de referência, embora condicionados pela natureza e disponibilidade dos dados utilizados.

O objetivo relacionado com a implementação de algoritmos de otimização de rotas e alocação de recursos foi igualmente cumprido, tendo sido avaliados três solvers distintos (OR-Tools, PyVRP e DEAP) em instâncias de dimensão crescente. Os resultados demonstraram elevada robustez do OR-Tools e PyVRP na satisfação das restrições temporais, bem como diferentes compromissos entre eficiência e distribuição de carga.

O objetivo de integração entre previsão e planeamento foi concretizado através do desenvolvimento de um pipeline unificado, incluindo um mecanismo de rolling horizon, que permitiu incorporar informação preditiva no processo de geração e ajuste de planos operacionais.

O objetivo relativo à demonstração de ganhos operacionais em cenários simulados foi parcialmente alcançado. Embora se tenham observado reduções consistentes no tempo total de deslocação após re-otimização, a magnitude dos ganhos variou entre instâncias e ficou abaixo do intervalo inicialmente esperado para alguns cenários, o que pode ser explicado pelas características do dataset utilizado e pelas limitações do modelo preditivo.

A análise das subperguntas de investigação permite ainda confirmar que o sistema desenvolvido responde adequadamente às questões propostas, nomeadamente na integração de mecanismos preditivos no planeamento, na comparação de estratégias de otimização e na articulação entre previsão e re-otimização dinâmica.

8.3 Contributos

O principal contributo desta dissertação consiste no desenvolvimento e validação de um pipeline integrado que combina modelos preditivos de cancelamentos com algoritmos de otimização de rotas VRPTW e um mecanismo de rolling horizon para re-otimização adaptativa em contexto de apoio domiciliário. Ao contrário de abordagens tradicionais, que tratam previsão e planeamento como problemas independentes, o sistema desenvolvido permite que a informação preditiva influencie diretamente a geração e atualização dos planos operacionais, fechando o ciclo entre previsão e decisão.

Um segundo contributo é a implementação de uma estrutura de avaliação unificada para comparação de solvers VRPTW com métricas heterogéneas, permitindo a avaliação consistente de abordagens baseadas em diferentes funções objetivo e representações internas.

Um terceiro contributo é a validação empírica do impacto da integração entre previsão e otimização no desempenho do sistema, demonstrando que a incorporação de informação preditiva no processo de planeamento permite ajustar dinamicamente os planos operacionais mantendo a viabilidade das soluções geradas.

8.4 Limitações

As principais limitações do trabalho foram documentadas em detalhe na Secção 7.6.1 e decorrem essencialmente de três fatores.

Em primeiro lugar, a utilização de datasets públicos não especificamente concebidos para o contexto de apoio ao domicílio limita a representatividade dos padrões de procura e afeta a generalização do modelo preditivo para cenários reais.

Em segundo lugar, a simplificação da arquitetura do DEAP para suporte a uma única rota impede a sua comparação direta com os restantes solvers, condicionando a análise de desempenho relativo entre abordagens.

Em terceiro lugar, não foi considerada a modelação explícita das competências profissionais no processo de otimização e re-otimização, o que restringe a expressividade do modelo no que respeita à alocação de recursos humanos em contextos reais de prestação de cuidados.

Estas limitações são inerentes ao âmbito de um protótipo académico e não comprometem a validade das conclusões obtidas.

8.5 Trabalho futuro

As linhas de trabalho futuro mais relevantes identificadas são a recolha ou construção de um dataset específico de apoio domiciliário para melhoria do modelo preditivo, a implementação de um mecanismo de distribuição equilibrada de carga na re-otimização parcial, a integração explícita das restrições de competências na instância reduzida, e a reformulação do DEAP para suporte a múltiplas rotas com restrições temporais. A extensão da avaliação a instâncias de 100 ou mais pacientes permitiria adicionalmente caracterizar com maior detalhe o comportamento de escalabilidade do protótipo em condições próximas de cenários operacionais reais.

Referências

- Algendi, A., Urrutia, S., Hvattum, L. M. & Helgheim, B. I. (2024) 'Home Healthcare Staffing, Routing, and Scheduling Problem With Multiple Shifts and Emergency Considerations', Networks. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/net.22260> (Acedido: 1 dezembro 2024).
- Data Science Central (2025) CRISP-DM: The Standard for Data Mining and Knowledge Discovery. Disponível em: <https://www.datascience-pm.com/crisp-dm-2/> (Acedido: 14 dezembro 2025).
- Decerle, J., Grunder, O., Hassani, E. & Barakat, O. (2018) 'A memetic algorithm for a home health care routing and scheduling problem', Operations Research for Health Care. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2211692317300735> (Acedido: 1 outubro 2025).
- Fikar, C. & Hirsch, P. (2017) 'Home health care routing and scheduling: A review', Computers & Operations Research. Disponível em: <https://doi.org/10.1016/j.cor.2016.07.019> (Acedido: 1 dezembro 2025).
- Instituto Nacional de Estatística (2025) Projeções de população residente em Portugal. Disponível em: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_bo ui=707257343&DESTAQUESmodo=2 (Acedido: 10 outubro 2025).
- Jesus, P. S. (2023) Rotas dinâmicas, um caso de estudo nos serviços urgentes de prestação de cuidados respiratórios domiciliários [Dissertação de Mestrado, ISCTE — Instituto Universitário de Lisboa]. Disponível em: https://repositorio.iscte-iul.pt/bitstream/10071/30063/1/master_pedro_seoane_jesus.pdf (Acedido: 10 dezembro 2025).
- Mendes, P. A. S. (2016) Desenho e Planeamento de Rotas de um Serviço de Apoio Domiciliário [Dissertação de Mestrado, ISCTE — Instituto Universitário de Lisboa]. Disponível em: <https://repositorio.iscte-iul.pt/bitstream/10071/13395/1/Paulo%20Vers%C3%A3o%20Final.pdf> (Acedido: 8 dezembro 2025).
- Oliveira, J. R. O. de (2015) 'Abordagens heurísticas de solução para Problemas de Roteamento de Veículos (PRV)', Revista Espacios, 36(13). Disponível em: <https://www.revistaespacios.com/a15v36n13/15361303.html> (Acedido: 6 outubro 2025).
- Oliveira, P. A. M. P. C. (2022) VRP dinâmico e de janelas de tempo aplicado ao planeamento de transportes [Dissertação de Mestrado, Instituto Politécnico do Porto]. Disponível em:

<https://recipp.ipp.pt/bitstreams/ca9ba8b5-f675-4d34-ab61-0c234bc0e9aa/download>
(Acedido: 4 janeiro 2026).

Pereira, J. M. S. de (2022) Uma Abordagem à Resolução do Capacitated Vehicle Routing Problem Utilizando uma Meta-Heurística Baseada em Recozimento Quântico [Dissertação de Mestrado, Universidade do Minho]. Disponível em:
<https://repositorium.uminho.pt/server/api/core/bitstreams/2551e3e9-6604-40d8-a4b3-02af91150838/content> (Acedido: 5 outubro 2025).

Sales, R. D. F. (2022) Desenvolvimento e Aplicação de uma Meta-Heurística ao Home Health Care Problem [Dissertação de Mestrado, Universidade Nova de Lisboa]. Disponível em:
https://run.unl.pt/bitstream/10362/158714/1/Sales_2022.pdf (Acedido: 8 dezembro 2025).

Toth, P. & Vigo, D. (2014) Vehicle Routing: Problems, Methods, and Applications. SIAM. Disponível em: <https://epubs.siam.org/doi/book/10.1137/1.9781611973594> (Acedido: 6 janeiro 2026).

World Health Organization (2021) Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva: WHO. Disponível em:
<https://www.who.int/publications/i/item/9789240029200> (Acedido: 17 janeiro 2026).

Afkummer, A. et al. (2021) *HHCRSP Dataset 2021*. GitHub. Disponível em:
<https://github.com/afkummer/hhcrsp-dataset-2021>
(Acedido: 20 fevereiro 2026).

Jiao Zhao (2022) *Bi-objective Home Health Care Dataset*. figshare. Disponível em:
<http://doi.org/10.6084/m9.figshare.21339072>
(Acedido: 20 fevereiro 2026).

Yucel, E., Cebeci, E. e Koc, C. (2023) *Data instances of the heterogeneous fleet electric home health care routing and scheduling problem with synchronized demand*. Mendeley Data. Disponível em: <https://doi.org/10.17632/4dxm7tvmc9.1>
(Acedido: 20 janeiro 2026).

Arroba, J. (2016) *Medical Appointment No-Shows Dataset*. Kaggle. Disponível em: <https://www.kaggle.com/datasets/joniarroba/noshowappointments>
(Acedido: 20 fevereiro 2026).

González Galtier, C. (2024) *Medical Appointment Scheduling Dataset*. Kaggle. Disponível em: <https://www.kaggle.com/datasets/carogonzalezgaltier/medical-appointment-scheduling-system>
(Acedido: 20 fevereiro 2026).

- Centers for Medicare & Medicaid Services (2025) *Home Health Services Dataset*.
Disponível em: <https://data.cms.gov/provider-data/dataset/6jpm-sxkc>
(Acedido: 20 fevereiro 2026).
- Chen, T. e Guestrin, C. (2016) 'XGBoost: A scalable tree boosting system', *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.
Disponível em: <https://dl.acm.org/doi/10.1145/2939672.2939785>
(Acedido: 12 março 2026).
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. e Liu, T.-Y. (2017) 'LightGBM: A highly efficient gradient boosting decision tree', *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, pp. 3146–3154.
Disponível em: <https://papers.nips.cc/paper/6907-lightgbm-a-highly-efficient-gradient-boosting-decision-tree>
(Acedido: 16 março 2026).
- Wouda, N. A., Lan, L. e Kool, W. (2024) 'PyVRP: A high-performance VRP solver package', *INFORMS Journal on Computing*, 36(4), pp. 943–955.
Disponível em: <https://doi.org/10.1287/ijoc.2023.0055>
(Acedido: 20 março 2026).
- Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M. e Gagné, C. (2012) 'DEAP: Evolutionary algorithms made easy', *Journal of Machine Learning Research*, 13, pp. 2171–2175.
Disponível em: <https://jmlr.org/papers/v13/fortin12a.html>
(Acedido: 22 março 2026).
- Shwartz-Ziv, R. e Armon, A. (2022) 'Tabular data: Deep learning is not all you need', *Information Fusion*, 81, pp. 84–90
Disponível em: <https://doi.org/10.1016/j.inffus.2021.11.011>
(Acedido: 27 março 2026).
- Perron, L. e Furnon, V. (2023) *OR-Tools* (versão 9.15). Google.
Disponível em: <https://developers.google.com/optimization>
(Acedido: 27 março 2026).
- Lundberg, S. M. e Lee, S.-I. (2017) 'A unified approach to interpreting model predictions', *Advances in Neural Information Processing Systems 30 (NIPS 2017)*. Disponível em <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
(Acedido: 14 Abril 2026).