INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO
MESTRADO EM ENGENHARIA INFORMÁTICA

Sistema de Recomendação no Contexto Educacional

RITA MENDES DE AZEVEDO
Junho de 2025

POLITÉCNICO
DO PORTO

ISEP INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

Sistema de Recomendação no Contexto Educacional

Rita Azevedo

Dissertação para obtenção do Grau de Mestre em Engenharia Informática, Área de Especialização em Sistemas de Informação e Conhecimento

Orientador: Luiz Faria
Co-orientador: Constantino Martins

Porto, 28 de junho de 2025

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. As exceções estão explicitamente reconhecidas na secção "Considerações éticas" do primeiro capítulo. Esta secção também declara como as ferramentas de IA foram utilizadas e para que finalidade.

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

ISEP, Porto, 28 de junho de 2025

Resumo

Com a constante evolução tecnológica, as escolas portuguesas têm vindo a transitar progressivamente dos manuais escolares em papel para versões digitais. Esta transição exige uma crescente adaptação dos alunos ao uso de ferramentas tecnológicas no seu processo de aprendizagem. No entanto, esta transição nem sempre é acompanhada por soluções eficazes que respondam às necessidades específicas de cada estudante, tornando essencial o desenvolvimento de plataformas que promovam a personalização do ensino.

Neste enquadramento, o presente projeto teve como objetivo principal investigar de que forma um sistema de recomendação personalizado pode apoiar e facilitar o processo de aprendizagem no contexto do ensino escolar. Para tal, procedeu-se à realização de uma revisão do estado da arte sobre sistemas de recomendação aplicados ao ensino. Adicionalmente, procurou-se explorar o potencial dos Large Language Models (LLMs) em contextos educativos e analisar criticamente as plataformas educacionais existentes.

A metodologia adotada envolveu uma abordagem exploratória, sustentada por pesquisa bibliográfica em bases científicas. Na seleção dos artigos, consideraram-se como critérios critérios a atualidade, a relevância e o foco na aplicação de sistemas de recomendação em contextos educacionais.

No âmbito do projeto, foi desenvolvida uma plataforma educacional que disponibiliza conteúdos educativos, integrando um sistema de recomendação personalizado, um assistente virtual de apoio ao estudo entre outras funcionalidades complementares.

Os resultados obtidos demonstraram que a plataforma desenvolvida foi bem recebida pelos participantes. Estes utilizadores beneficiaram de recomendações personalizadas de conteúdo e de um assistente virtual capaz de apoiar o estudo através de comandos personalizados. A análise das métricas indicou um bom desempenho do sistema de recomendação, revelando a sua capacidade de adaptação aos interesses e dificuldades dos estudantes.

Conclui-se que uma plataforma educacional gratuita, equipada com funcionalidades inteligentes e personalizadas, pode representar uma mais-valia significativa no apoio ao ensino digital. Com base nos dados recolhidos, foram ainda implementadas melhorias e adicionadas novas funcionalidades, contribuindo para o desenvolvimento de futuras soluções tecnológicas mais eficazes e centradas no aluno.

Palavras-chave: Educação em Portugal, Recursos Educacionais Digitais, Sistemas de Recomendação, Large Language Models (LLM), Inteligência Artificial no Ensino

Abstract

With the constant technological evolution, Portuguese schools have been progressively transitioning from paper textbooks to digital versions. This transition requires increasing adaptation from students to the use of technological tools in their learning process. However, this transition is not always accompanied by effective solutions that respond to the specific needs of each student, making it essential to develop platforms that promote personalized teaching.

In this context, the main objective of this project was to investigate how a personalized recommendation system can support and facilitate the learning process in the school education context. To this end, a state-of-the-art review on recommendation systems applied to education was conducted. Additionally, the potential of Large Language Models (LLMs) in educational contexts was explored and existing educational platforms were critically analyzed.

The adopted methodology involved an exploratory approach, supported by bibliographic research in scientific databases. In the selection of articles, the criteria considered were timeliness, relevance, and focus on the application of recommendation systems in educational contexts.

Within the scope of the project, an educational platform was developed that provides educational content, integrating a personalized recommendation system, a virtual study support assistant, and other complementary functionalities.

The results obtained demonstrated that the developed platform was well received by the participants. These users benefited from personalized content recommendations and a virtual assistant capable of supporting study through personalized commands. The analysis of metrics indicated good performance of the recommendation system, revealing its ability to adapt to students' interests and difficulties.

It is concluded that a free educational platform, equipped with intelligent and personalized functionalities, can represent a significant added value in supporting digital education. Based on the collected data, improvements were implemented and new functionalities were added, contributing to the development of more effective and student-centered future technological solutions.

Conteúdo

Lista de Figuras **xiii**

Lista de Tabelas **xv**

Lista de Código **xvii**

Lista de Acrónimos **xix**

1 Introdução **1**
1.1 Enquadramento e Contextualização . . . 1
1.2 Proposta de Solução . . . 2
1.3 Problema . . . 2
1.4 Objetivos . . . 2
1.5 Motivação . . . 3
1.6 Metodologia . . . 4
1.6.1 Métodos . . . 4
1.6.2 Processo . . . 4
1.6.3 Questões . . . 4
1.7 Ética e Segurança . . . 5
1.7.1 Considerações Éticas . . . 5
1.7.2 Segurança . . . 6
1.8 Estrutura . . . 6

2 Estado da Arte **9**
2.1 Sistema Educacional Português . . . 9
2.1.1 Graus de Escolaridade e Exames Nacionais . . . 9
2.1.2 Programme for International Student Assessment (PISA) . . . 10
2.1.3 Entidades Responsáveis pelo Ensino Escolar . . . 11
2.1.4 Ensino Superior . . . 11
2.2 Conteúdo Educacional Português . . . 11
2.3 Sistemas de Recomendação . . . 12
2.3.1 Recomendação por filtragem colaborativa . . . 13
Filtragem Colaborativa baseada em modelo . . . 13
Filtragem Colaborativa baseada em memória . . . 13
Clustering . . . 14
2.3.2 Recomendação baseada em conteúdo . . . 15
2.3.3 Recomendação baseada em conhecimento . . . 15
Baseado em tabelas . . . 15
Baseado em restrições . . . 16
Baseado em casos . . . 16
2.3.4 Recomendação baseada em sessão . . . 16

2.3.5 Recomendação Híbrida . . . 16
2.3.6 Características e Desafios dos Métodos de Recomendação . . . 20
Sistemas de Recomendação Educacionais . . . 21
2.4 Large Language Models (LLMs) . . . 24
2.4.1 Componentes de um LLM . . . 25
2.4.2 Desafios . . . 26
2.4.3 Retrieval-Augmented Generation (RAG) . . . 28
2.4.4 Grafos de Conhecimento . . . 29
2.4.5 Integração de Grafos de Conhecimento e RAG no Ensino . . . 30
2.4.6 Utilização em Chatbots . . . 30
2.4.7 Utilização em Sistemas de Recomendação . . . 31
2.5 Custos de um LLM . . . 32
2.5.1 Monetário . . . 33
2.5.2 Hardware . . . 33
2.6 Fontes de dados de conteúdo educacional . . . 33
2.7 Plataformas de Ensino em Portugal . . . 33
2.8 Sumário . . . 35

3 Análise e Desenho do Sistema **37**
3.1 Âmbito do Projeto . . . 37
3.2 Atores e Casos de Uso . . . 38
3.3 Requisitos . . . 40
3.3.1 Requisitos Funcionais . . . 40
3.3.2 Requisitos Não funcionais . . . 43
3.3.3 Requisitos de Segurança . . . 44
3.3.4 Requisitos de Interface e Usabilidade . . . 46
3.3.5 Requisitos de Integração . . . 47
3.4 Modelo de Domínio . . . 48
3.5 Arquitetura do Sistema . . . 49
3.5.1 Arquitetura de Alto Nível do Sistema . . . 49
3.5.2 Arquitetura do *Frontend* . . . 51
3.5.3 Arquitetura do *Backend* . . . 53
3.5.4 Arquitetura dos *Serviços Externos* . . . 54
3.5.5 Diagrama Entidade Relação (ER) . . . 55
3.5.6 Processo de Recomendação do Sistema . . . 56
3.6 Sumário . . . 58

4 Implementação **59**
4.1 Produto . . . 59
4.2 Descrição e Organização . . . 59
4.3 Conjunto de Dados de Conteúdos Educacionais . . . 60
4.4 Controlo de Acesso e Autenticação . . . 63
4.4.1 Controlo de Acesso . . . 63
4.4.2 Autenticação . . . 64
4.5 Gestão de conteúdo curtido pelos utilizadores . . . 66
4.6 Recomendações Personalizadas . . . 66
4.7 Desenvolvimento do Assistente Virtual . . . 68
4.7.1 Configuração do Sistema . . . 68
Análise de Modelos . . . 68

Integração dos Serviços Externos . . . 68
4.7.2 Comandos de Resposta Personalizados . . . 70
4.7.3 Interação com o Assistente Virtual Tuga . . . 71
4.8 Resolução de Exercícios . . . 74
4.9 Módulos . . . 74
4.9.1 Visão Geral . . . 74
4.9.2 Registo . . . 75
4.9.3 Login e Logout . . . 76
4.9.4 Assistente Virtual . . . 77
4.9.5 Histórico . . . 78
4.9.6 Página de Recomendações . . . 78
4.9.7 Conteúdos Educacionais . . . 79
4.9.8 Resolver Exercícios . . . 81
4.10 Sumário . . . 82

5 Caso de Estudo **83**
5.1 Propósito . . . 83
5.2 Metodologia . . . 83
5.3 Procedimento . . . 84
5.4 Recolha de dados . . . 85
5.5 Sumário . . . 86

6 Resultados e Discussão **87**
6.1 Caracterização da Amostra . . . 87
6.2 Avaliação ao Assistente Virtual Tuga . . . 87
6.2.1 Análise das Interações com o Tuga . . . 87
6.2.2 Análise do *Feedback* Sobre o Tuga . . . 88
6.3 Análise do Conteúdo e Recomendações . . . 90
6.3.1 Análise do *Feedback* Sobre o Conteúdo e Sistema de Recomendação 90
6.3.2 Métricas de Avaliação . . . 91
6.3.3 Melhorias Finais ao Sistema de Recomendação . . . 92
6.4 Interface da Plataforma . . . 93
6.5 Validação da Utilização da Plataforma . . . 95
6.6 Sumário . . . 96

7 Conclusões **99**
7.1 Síntese . . . 99
7.2 Objetivos Cumpridos . . . 100
7.3 Resultados Alcançados . . . 100
7.4 Limitações . . . 101
7.5 Trabalho Futuro . . . 102

Bibliografia **105**

Apêndice A Apêndice A - Planeamento do Projeto **115**
A.1 Riscos do Projeto . . . 115
A.2 Competências . . . 115
A.2.1 Competências mais Desenvolvidas . . . 115
A.2.2 Competências a desenvolver . . . 116
A.3 Âmbito do Projeto . . . 116

A.3.1 Principais *Stakeholders* 116
A.3.2 Benefícios do Projeto 117
A.4 *Work Breakdown Structure(WBS)* 117
A.5 Cronograma 119
A.6 Monitorização e Controlo 120

Apêndice B Apêndice B - Requisitos Funcionais Complementares 121

Lista de Figuras

2.1 Métodos Utilizados na Recomendação do Conteúdo, Adaptado de (Martins e Figueiredo 2024) . . . 12
2.2 Método de Recomendação Híbrida Ponderada . . . 17
2.3 Método de Recomendação Híbrida, Alternância . . . 17
2.4 Método de Recomendação Híbrida, Combinação de Diferentes Métodos . . . 18
2.5 Método de Recomendação Híbrida, Fusão de Características . . . 18
2.6 Método de Recomendação Híbrida em Cascata . . . 19
2.7 Método de Recomendação Híbrida, Aumento de Características . . . 19
2.8 Método de Recomendação Híbrida, Meta-Level . . . 20
2.9 Componentes dos LLMs . . . 26
2.10 Retrieval-Augmented Generation (RAG) . . . 29
2.11 Observações a diferentes plataformas e sites portugueses . . . 34

3.1 Diagrama de Casos de Uso Controlo de Acesso . . . 39
3.2 Diagrama de Casos de Uso Utilizador Autenticado . . . 40
3.3 Modelo de Domínio da plataforma O Lama Sabe . . . 48
3.4 Arquitetura do Sistema . . . 49
3.5 Arquitetura de Alto Nível do Sistema . . . 50
3.6 Arquitetura do *container Frontend* . . . 52
3.7 Arquitetura *container Backend* . . . 54
3.8 Vista Lógica do *container* Serviços Externos . . . 55
3.9 Diagrama Entidade Relação (ER) . . . 56
3.10 Fluxo BPMN de Recomendações Personalizadas na Página de Conteúdos . . . 57
3.11 Fluxo BPMN de Recomendações Personalizadas Através de Comandos Personalizados do Assistente Virtual Tuga . . . 58

4.1 Distribuição de Conteúdos por Tipo . . . 61
4.2 Quantidade de Conteúdos por Categoria e Subcategoria . . . 62
4.3 Fluxograma do Controlo de Acesso . . . 64
4.4 Fluxograma do Processo de Autenticação . . . 65
4.5 Fluxograma das Etapas do Sistema de Recomendação . . . 67
4.6 Grafo de Conhecimento . . . 69
4.7 Grafo de Conhecimento Ampliado . . . 70
4.8 Fluxograma de Funcionamento do Tuga . . . 72
4.9 Diagrama de Sequência da Interação com o Tuga . . . 73
4.10 Informação Sobre a Plataforma na Página Visão Geral . . . 75
4.11 Página de Registo . . . 76
4.12 Página Login . . . 76
4.13 Opção *Logout* na Barra de Navegação . . . 77
4.14 *Chat* do Assistente Virtual Tuga . . . 77
4.15 Resposta ao Comando Personalizado . . . 77

4.16 Histórico de Mensagens 78
4.17 Página de Recomendações 79
4.18 Página Livros 80
4.19 Página Documentos 80
4.20 Página Vídeos 81
4.21 Página Exercícios 81
4.22 Resolução de Exercícios 82

6.1 Análise à Relevância da Informação Fornecida pelo Tuga 89
6.2 Análise ao Assistente Virtual Tuga como Apoio ao Estudo 89
6.3 Análise à Utilidade das Recomendações 91
6.4 Análise à Interface da Plataforma 94
6.5 Análise à Utilização Futura da Plataforma 95

A.1 Riscos do Projeto 115
A.2 WBS 118
A.3 Gráfico de *Gantt*, Tarefas 119
A.4 Gráfico de *Gantt*, Cronograma 119

Lista de Tabelas

2.1 Comparação de Sistemas de Recomendação 21
2.2 Comparação de Sistemas de Recomendação 22

3.1 RF1 – Registo do Utilizador 41
3.2 RF2 – Autenticação 41
3.3 RF3 – Terminar Sessão 41
3.4 RF4 – Utilizadores não Autenticados 41
3.5 RF5 – Conteúdo Educacional 41
3.6 RF6 – Resolução de Exercícios 41
3.7 RF7 – Recomendação de Conteúdo 42
3.8 RF8 – Comunicar com o Tuga 42
3.9 RF9 – Comandos de Recomendações Personalizadas 42
3.10 RF10 – Histórico de Mensagens 42
3.11 RNF1 – Desempenho 43
3.12 RNF1.1 – Tempo de Carregamento da Página de Recomendações 43
3.13 RNF1.2 – Tempo de Resposta do Tuga 43
3.14 RNF2 – Envio de Dados para o Tuga 43
3.15 RNF3 – Confiabilidade 43
3.16 RNF3.1 – Tolerância a Falhas 44
3.17 RNF3.2 – Armazenamento dos Dados 44
3.18 RNF4 – Escalabilidade 44
3.19 RS1 – Armazenamento dos Dados Sensíveis dos Utilizadores 44
3.20 RS2 – Controlo de Acesso à Plataforma 44
3.21 RS3 – Segurança das Sessões 45
3.22 RS4 – Invalidar Sessões 45
3.23 RS5 – Encerramento de Conexões à Base de Dados 45
3.24 RS6 – Segurança contra *SQL Injection* 45
3.25 RS7 – Invalidar Sessões 45
3.26 RI1 – Compatibilidade de Navegadores 46
3.27 RI2 – Estabilidade Visual 46
3.28 RI3– Interface Intuitiva 46
3.29 RI4 – Apelo e Conforto Visual 46
3.30 RI4 – Dinâmica do Botão de Envio do Tuga 46
3.31 RI5 – Caixas de Texto Ajustáveis 47
3.32 RIN1 – Integração com Base de Dados Relacional 47
3.33 RIN2 – Integração com *Endpoints* 47
3.34 RIN3 – Apelo e Conforto Visual 47
3.35 RIN4 – Utilização de LLM na Nuvem 47

5.1 Questionário de Opinião 86

6.1 Média Ponderada das Métricas de Avaliação por Número Gostos 92

6.2 Média das Métricas de Avaliação por Número de Gostos Após Melhoria ao Sistema de Recomendação 93

B.1 RF1.1 – Inserir dados pessoais no registo 121
B.2 RF1.2 – Validação do Nome 121
B.3 RF1.3 – Validação do Nome de Utilizador 121
B.4 RF1.4 – Validação da Palavra-passe 121
B.5 RF1.5 – Verificação da Confirmação da Palavra-passe 122
B.6 RF1.6 – Consulta e Aceitação dos Termos e Condições 122
B.7 RF1.7 – Preferência de Recomendação de Conteúdo Variado 122
B.8 RF2.1 – Verificar se o Utilizador Existe 122
B.9 RF2.2 – Verificar a Palavra-passe do Utilizador 122
B.10 RF5.1 – Filtragem de Conteúdo 122
B.11 RF5.2 – Fluxo de Navegação dos Conteúdos 123
B.12 RF5.3 – Gostar de Conteúdos Educacionais 123
B.13 RF5.4 – Visualizar Conteúdos Gostados 123
B.14 RF5.5 – Retirar Gosto de Conteúdos 123
B.15 RF5.6 – Visualizar Conteúdos Curtidos 123
B.16 RF6.1 – Validação da Resposta 123
B.17 RF6.2 – Resposta Incorreta 124
B.18 RF6.3 – Controlo da Validação 124
B.19 RF7.1 – Recomendação Baseada em Conteúdo 124
B.20 RF7.2 – Visualizar Recomendações 124
B.21 RF7.3 – Exclusão de Conteúdos Previamente Gostados 124
B.22 RF7.4 – Recomendação Variada de Conteúdo 125
B.23 RF7.5 – Recomendações para Utilizadores sem Gostos 125
B.24 RF7.6 – Avaliações das Recomendações 125
B.25 RF8.1 – Resposta ao Utilizador 125
B.26 RF8.2 – Texto Multi Linha 125
B.27 RF8.3 – Mensagem Introdutória 125
B.28 RF8.4 – Comando de Ajuda 126
B.29 RF8.5 – Reconhecimento Textual 126
B.30 RF8.6 – Citação de Fontes 126
B.31 RF8.7 – Conformidade Linguística do Tuga 126
B.32 RF8.8 – Controlo de Envio de Mensagens 126

Lista de Código

4.1 LightRAG com Groq e Ollama 69
4.2 Prompt para Validar Resposta 74

Lista de Acrónimos

AGI	Inteligência Geral Artificial.
AISAR	Artificial Intelligence-based Student Assessment and Recommendation.
API	Application Programming Interface.
BPMN	Business Process Model and Notation.
CoT	Chain-of-Thought.
DBSCAN	Density-Based Spatial Clustering of Applications with Noise.
EC	Enhanced Clustering.
ER	Entidade-Relação.
GPU	Graphics Processing Unit.
IA	Inteligência Artificial.
ICL	In-Context Learning.
IDF	Inverse Document Frequency.
IPP	Instituto Politécnico do Porto.
IR	Information Retrieval.
MDP	Markov decision process.
MOOCs	Massive Open Online Courses.
NLP	Natural Language Processing.
OLM	Open Learner Model.
PISA	Programme for International Student Assessment.
RAG	Retrieval Augmentation Generation.
RGPD	Regulamento Geral sobre a Proteção de Dados.
TF	Term Frequency.
WBS	Work Breakdown Structure.

Capítulo 1

Introdução

Neste capítulo, são apresentados os principais tópicos relacionados à dissertação. Aborda-se com a dissertação, a definição do problema e o seu contexto, os objetivos a serem alcançados, a motivação subjacente ao trabalho, a metodologia de pesquisa adotada para a análise do estado da arte, bem como a estruturação geral do documento.

1.1 Enquadramento e Contextualização

Este projeto é desenvolvido no âmbito da dissertação do Mestrado em Engenharia Informática, no ramo de Sistemas de Apoio à Decisão, do Instituto Superior de Engenharia do Porto (ISEP). Este consiste no desenvolvimento de um protótipo de sistema de recomendação personalizado, focado nas preferências de estudo dos estudantes.

A utilização de plataformas e recursos digitais no ensino tem vindo a tornar-se cada vez mais presente no quotidiano dos estudantes. A tecnologia evolui constantemente. Consequentemente, muitas instituições de ensino estão progressivamente a substituir os manuais escolares físicos por digitais (Diário de Notícias 2024).

Além dos manuais escolares digitais e dos conteúdos disponibilizados pelas escolas, existe uma vasta quantidade de material disponível na Internet que os estudantes podem utilizar como apoio ao estudo. No entanto, decidir qual conteúdo é mais adequado para estudar pode ser um desafio.

A tomada de decisão está presente no quotidiano, sendo por vezes um processo complexo. Com frequência surgem incertezas sobre o que comprar a seguir, ou encontrar artigos que correspondam ás preferências individuais. No contexto educacional, existe uma ampla diversidade de conteúdos educacionais, permitindo aos estudantes estudar conforme as suas preferências. Entre os conteúdos disponíveis destacam-se fichas de exercícios, livros, vídeos, jogos didáticos, filmes, artigos científicos e outros materiais educativos (Khan Academy 2024; O Bichinho do Saber 2024; Soared 2021).

Os sistemas de recomendação são a solução para ajudar na tomada de decisão. Estes sistemas oferecem sugestões aos utilizadores em diversas áreas (Martins e Figueiredo 2024) . Por exemplo, indicam que filme ou série assistir, sugerem potenciais compras futuras, apresentam notícias selecionadas de acordo com o perfil do utilizador e disponibilizam publicidade personalizada (Martins e Figueiredo 2024). No contexto educativo, estes sistemas possibilitam a recomendação de conteúdos personalizados em função das preferências, dificuldades de cada estudante, perfil do utilizador e outros aspetos, orientando-o nos estudos de forma a otimizar a pesquisa e o acesso ao conteúdo relevante (Shishehchi et al. 2012).

Nem sempre os estudantes encontram resposta às suas questões. Nesses casos, é comum recorrerem à utilização de *chatbots* baseados em Inteligência Artificial (IA), atuando como um complemento ao processo de estudo, auxiliando na obtenção de informações e esclarecimento de dúvidas (Dam et al. 2024).

1.2 Proposta de Solução

O sistema proposto deverá considerar o conteúdo educacional de preferência de cada estudante, fornecendo recomendações adequadas para apoiar o processo de aprendizagem.

Adicionalmente, deverá fornecer uma alternativa no caso do conteúdo disponível não responder às dúvidas e questões dos estudantes.

Como resultado final, será desenvolvido um protótipo funcional a ser validado com um grupo de estudantes, avaliando a eficácia das recomendações e se vão ao encontro do que o estudante pretendia.

1.3 Problema

Atualmente, verifica-se uma ampla diversidade de conteúdos educacionais disponíveis, sendo que cada estudante apresenta um estilo de aprendizagem único. Contudo, a maioria das plataformas educacionais enfrenta limitações significativas, nomeadamente a incapacidade de disponibilizar materiais de ensino adaptados a diferentes estilos de aprendizagem. Além disso, muitas das plataformas mais populares não são de acesso gratuito, representando uma barreira significativa para estudantes com menores recursos financeiros (Escola Mágica 2024; Escola Virtual 2024; Leya Aula Digital 2024). Adicionalmente, as soluções existentes implementadas frequentemente demonstram limitações na capacidade de recomendar eficazmente conteúdos educacionais que atendam às dificuldades dos estudantes.

Caso o sistema não seja capaz de recomendar adequadamente os conteúdos educacionais, existe o risco de não apoiar eficazmente o estudante nas áreas que apresenta maiores dificuldades. Comprometendo o objetivo principal da plataforma, prejudicando o processo de aprendizagem e reduzindo a eficácia do sistema de recomendação.

1.4 Objetivos

A dissertação é guiada pela seguinte questão de investigação para orientar o planeamento e desenvolvimento do projeto:

De que forma um sistema de recomendação personalizado pode apoiar e facilitar o processo de aprendizagem dos estudantes no contexto do ensino escolar?

O principal objetivo consiste no desenvolvimento de um sistema de recomendação personalizado, capaz de sugerir conteúdos aos utilizadores com base nas suas preferências individuais. Este sistema visa potenciar o sucesso no processo de aprendizagem ao longo do percurso escolar obrigatório.

O projeto visa alcançar um conjunto de objetivos, além do objetivo principal,de forma a garantir que decorra de maneira eficaz. Os objetivos seguem-se especificados:

- Realizar uma revisão do estado da arte com o propósito de responder às questões de investigação, permitindo identificar as abordagens mais adequadas para a implementação do sistema de recomendação.
- Desenvolver um protótipo funcional de uma plataforma educacional, dirigida a estudantes do ensino escolar obrigatório, integrada com um sistema de recomendação. Esta deverá ser capaz de fornecer conteúdos educacionais personalizados com base nas preferências de aprendizagem e nas dificuldades identificadas em cada estudante.
- Implementar os métodos e técnicas necessárias que permitam o desenvolvimento de um sistema de recomendação eficaz, com o objetivo de aprimorar as recomendações relevantes aos estudantes.
- Avaliar o desempenho do sistema de recomendação desenvolvido através de um conjunto de métricas, bem como validar a sua eficácia com a participação de estudantes em cenários reais.

1.5 Motivação

Atualmente, observa-se uma transição progressiva dos manuais escolares e conteúdos educacionais do formato físico para o formato digital (Diário de Notícias 2024). Embora esta mudança ainda não esteja implementada em todas as escolas de ensino público em Portugal, observa-se uma tendência crescente de adoção de recursos digitais no ensino escolar. A Internet constitui uma fonte abrangente de informação, disponibilizando uma vasta quantidade de conteúdos educacionais. No entanto, são raras as plataformas e os sites que oferecem uma ampla diversidade de recursos, como vídeos, fichas com resoluções, manuais, entre outros, e que simultaneamente consigam recomendar eficazmente esses conteúdos aos seus utilizadores (Khan Academy 2024; O Bichinho do Saber 2024; Soared 2021).

Apesar da disponibilidade de várias plataformas e sites educacionais, a maioria adota planos de subscrição, o que implica custos adicionais. A atual realidade em Portugal representa uma barreira significativa para famílias com baixas possibilidades económicas, enfrentando dificuldades no acesso a recursos informáticos e na capacidade de investir em serviços pagos, considerados como não essenciais básicos.

Outro desafio está relacionado com a ausência de sistemas de recomendação na maioria das plataformas educacionais portuguesas existentes (Mentes Brilhantes 2024; O Bichinho do Saber 2024). Mesmo que os conteúdos estejam organizados, os estudantes podem sentir-se perdidos ao tentar identificar quais os temas a estudar com base nas suas dificuldades específicas. Este aspeto evidencia uma vantagem dos recursos digitais face ao conteúdo educacional tradicional em papel, pois possibilita a integração de funcionalidades mais dinâmicas e adaptativas.

O desenvolvimento de uma plataforma educacional escolar gratuita que integre um sistema de recomendação personalizado apresenta diversas vantagens. Para além de apoiar os estudantes na melhoria do seu desempenho académico, permite também que se adaptem à nova realidade tecnológica, promovendo o uso preferencial de recursos digitais. Adicionalmente, estas ferramentas potenciam o trabalho autónomo dos estudantes quando o professor não está disponível. Esta iniciativa apresenta um potencial inovador no setor educacional, contribuindo para o avanço das práticas pedagógicas.

1.6 Metodologia

A pesquisa foi conduzida utilizando um conjunto estruturado de ferramentas e métodos, orientados por questões de investigação auxiliares.

1.6.1 Métodos

Na pesquisa, foram utilizadas principalmente as fontes *ResearchGate* e *Google Scholar*, além de *arXiv*, *ACM Digital Library* e *b-on*. Essas fontes já foram utilizadas em pesquisas anteriores e são as mais frequentemente utilizadas, tendo sido eficazes na aquisição do material necessário.

Na seleção dos artigos, foram consideradas as palavras-chave presentes nos títulos e, no caso de exemplos de sistemas de recomendação aplicados ao ensino, analisaram-se também as informações contidas nos resumos. Além disso, optou-se por consultar apenas artigos científicos e de conferências publicados nos últimos 14 anos, a fim de garantir a atualidade e relevância dos conteúdos.

Na aquisição de informações sobre o funcionamento do sistema educacional português, foram consultados os respetivos sites oficiais, da mesma forma que, na análise de plataformas e sites educacionais, foram consultados os endereços URL dessas fontes.

1.6.2 Processo

A pesquisa seguiu uma abordagem exploratória e adaptativa, com a identificação de temas de interesse ao longo do processo. Em algumas situações, dentro dos artigos explorados, procedeu-se à análise de artigos citados pelos autores, expandindo o alcance da pesquisa. Esta abordagem proporcionou uma maior flexibilidade na exploração do conteúdo, permitindo a identificação de informações relevantes a serem incluídas na investigação.

1.6.3 Questões

A pesquisa tem como objetivo responder às seguintes questões, com o objetivo de investigar os diversos tópicos, com maior foco nos métodos utilizados em sistemas de recomendação e modelos de LLM, incluindo a análise de casos já existentes:

- **Q1:** Como está estruturado o sistema educativo português?
- **Q2:** O que é um sistema de recomendação e quais os métodos utilizados na recomendação de conteúdo?
- **Q3:** Que abordagens e ferramentas são utilizadas na recomendação de conteúdo em plataformas educacionais?
- **Q4:** A utilização de IA pode ajudar os estudantes nos estudos?
- **Q5:** De que forma os LLMs podem ser utilizados na construção de sistemas de aprendizagem?

1.7 Ética e Segurança

É fundamental planear e desenvolver o projeto de forma ética, seguindo os códigos de conduta estabelecidos pelo Instituto Politécnico do Porto (IPP).

A segurança dos utilizadores na plataforma a ser desenvolvida deve ser uma prioridade, implementando medidas adequadas.

1.7.1 Considerações Éticas

Durante a execução deste projeto, são seguidas as disposições do Código de Conduta do IPP, em conformidade com o Artigo 6.º **"Deveres dos/as Estudantes"** os seguintes pontos (Instituto Politécnico do Porto 2020):

- **2.8:** Garantir que todas as ideias, frases, parágrafos ou textos de terceiros estão devidamente citados e referenciados, respeitando os direitos autorais e as boas práticas académicas.
- **2.9:** Evitar a submissão de trabalhos já apresentados ou publicados anteriormente, sem comunicação explícita dessa reutilização.
- **2.10:** Garantir que todos os trabalhos realizados são individuais ou, em caso de colaboração, devidamente autorizados, evitando qualquer conluio não permitido.

O Código de Conduta do IPP, referente ao Artigo 8.º **"Declaração de Compromisso"**, reforça a integridade dos trabalhos académicos, enfatizando a sua originalidade e o cumprimento dos padrões éticos, bem como o compromisso com o Código de Conduta Ética do P.PORTO (Instituto Politécnico do Porto 2020).

Na fase de investigação, o projeto foi conduzido em conformidade com o Artigo 10.º **"Boas Práticas nas Atividades de Investigação"**, garantindo o respeito pelas normas éticas e científicas exigidas, assegurando o cumprimento dos direitos de autor e das práticas de conformidade na apresentação dos resultados da investigação (Instituto Politécnico do Porto 2020).

No caso de estudo foram considerados os princípios éticos definidos pela Política de Privacidade e pela Política de Utilização Aceitável das Tecnologias de Informação e Comunicação do Instituto Politécnico do Porto. O armazenamento dos dados pessoais na base de dados, bem como a participação no questionário, foi realizado de forma anónima. Além disso, os utilizadores tiveram de aceitar os Termos e Condições da plataforma antes de concluir o registo de conta.

Durante o desenvolvimento deste projeto, foram utilizadas ferramentas de IA, nomeadamente o OpenArt e o Image Generator, para a criação das imagens dos lamas, com exceção do logótipo (Alotaibi 2024; OpenArt 2022).

Adicionalmente, recorreu-se à utilização à utilização de inteligência artificial para apoiar a implementação da função associada ao modelo *Groq*. Devido a incompatibilidades entre as versões das bibliotecas utilizadas, não foi possível invocar diretamente a função original, sendo, por isso, necessário reescrevê-la como uma alternativa funcional. A inteligência artificial foi igualmente útil na criação de uma lista personalizada de stopwords, gerada com base na análise do conjunto de dados, o que permitiu poupar tempo no processo de exploração dos dados para criação da lista. Para estas tarefas, foi utilizado o assistente *GitHub Copilot* (GitHub, OpenAI 2021).

Salienta-se que, apesar do apoio das ferramentas de IA, a autoria e os resultados apresentados neste projeto são inteiramente da autoria deste trabalho, garantindo a originalidade e a integridade do mesmo.

1.7.2 Segurança

No desenvolvimento do projeto, assegura-se a segurança através da proteção dos dados dos alunos, mantendo a integridade do sistema e a confiança dos utilizadores.

- **Segurança no Desenvolvimento** No desenvolvimento do projeto, serão adotadas as medidas necessárias para garantir a segurança, garantindo a integridade dos dados, assegurando a confiança dos utilizadores na plataforma.
- **Segurança dos dados** A plataforma adotará as medidas necessárias para garantir a conformidade com o Regulamento Geral sobre a Proteção de Dados (RGPD) da união Europeia. As informações sensíveis dos utilizadores serão criptografadas, garantindo a confidencialidade e proteção dos dados. Alguns dados necessários para o tratamento serão recolhidos apenas mediante o consentimento explícito dos utilizadores. Apenas utilizadores autorizados terão acesso a informações pertinentes.

  Durante a realização de testes com um grupo de estudantes, será necessário obter o consentimento prévio dos participantes para a recolha e processamento dos dados, assegurando que todas as etapas do processo de teste estejam em conformidade com as regulamentações de privacidade e proteção de dados.

1.8 Estrutura

A estrutura do documento está dividida em sete capítulos. O presente capítulo apresenta o projeto a ser desenvolvido, bem como a metodologia utilizada para a pesquisa do estado da arte.

O Capítulo 2 aborda o estado da arte, iniciando com uma introdução e organizando o conteúdo em secções dedicadas a cada tema, finalizando com uma conclusão que sintetiza os resultados do estudo.

O Capítulo 3 descreve a análise e o desenho da solução desenvolvida, iniciando com a definição do âmbito do projeto e a identificação dos atores. A partir daí, são especificados os requisitos do sistema, bem como os diferentes modelos que ilustram o funcionamento da solução, finalizando com um sumário ao capítulo.

O Capítulo 4 apresenta os procedimentos de implementação da solução, a plataforma educacional e os seus principais componentes, nomeadamente o assistente virtual e o sistema de recomendação, bem como o conjunto de dados utilizado. O capítulo conclui com um sumário final.

O Capítulo 5 descreve o caso de estudo realizado para avaliação da solução proposta, abordando os objetivos, as etapas envolvidas e a descrição das atividades desenvolvidas em cada fase, finalizando com um sumário ao capítulo.

O Capítulo 6 analisa os resultados obtidos no caso de estudo e respetivas conclusões, incluindo as melhorias introduzidas na solução com base nos dados recolhidos. O capítulo encerra com um sumário final.

Por fim, o Capítulo 7 apresenta as conclusões finais do desenvolvimento do projeto, iniciando com uma síntese. São ainda discutidos os objetivos e resultados alcançados, as principais limitações encontradas ao longo do processo, bem como propostas de trabalhos futuros.

Capítulo 2

Estado da Arte

A evolução das tecnologias no ensino português tem acelerado nos últimos anos. Esta evolução impõe a transição dos manuais físicos para os digitais. Adicionalmente, exige que alunos, professores e encarregados de educação adquiram competências necessárias para utilizar computadores e navegar na Internet de forma segura.

Em cada ano de ensino, são abordados conteúdos educacionais distintos, definidos de acordo com os objetivos curriculares estabelecidos para cada ciclo. Por este motivo, é essencial compreender como está estruturado o sistema educativo português, uma vez que o presente projeto se dedica especificamente ao ensino escolar obrigatório em Portugal.

Apesar da ampla disponibilidade de recursos educativos online, nem todo o material disponível é utilizado no contexto escolar. Assim, é relevante compreender como é feita a seleção dos conteúdos lecionados nas escolas portuguesas, garantindo que os conteúdos estejam alinhados com os objetivos e exigências do ensino obrigatório em Portugal.

O presente estado da arte visa explorar as abordagens mencionadas, com o objetivo de proporcionar uma compreensão aprofundada de cada uma delas, analisar as técnicas utilizadas, e examinar alguns sistemas e arquiteturas de outros autores. Este estudo possibilitará a identificação das abordagens mais adequadas a serem implementadas e propor melhorias nas recomendações existentes. Além disso, discutir-se-á o funcionamento do ensino em Portugal, com destaque o Programme for International Student Assessment (PISA), as entidades responsáveis pelo ensino e será realizada uma análise de algumas plataformas e sites de ensino disponíveis no país.

2.1 Sistema Educacional Português

Nesta secção abordam-se diferentes conteúdos no contexto educacional. Focando-se na explicação do funcionamento do sistema educativo em Portugal, o que é o PISA, quais as entidades responsáveis pela educação e os diferentes níveis de ensino seguidos pelos estudantes no país.

2.1.1 Graus de Escolaridade e Exames Nacionais

Em Portugal o ensino escolar é organizado de acordo com os princípios democráticos estabelecidos pela Constituição da República (Eurydice – Comissão Europeia 2024). O sistema educativo português compreende a Educação Pré-Escolar, a Educação Escolar e a Educação Extra-Escolar (Diário da República 1986). A Educação Escolar está dividido em três níveis de ensino, Ensino Básico, Secundário e Superior (Diário da República 1986).

A educação Pré-Escolar, com um ciclo de frequência opcional dos 3 aos 6 anos de idade, seguido do Ensino Básico composto por três ciclos Sequenciais (Eurydice – Comissão Europeia 2024). O 1º ciclo com duração de 4 anos , o 2º ciclo de 2 anos e o 3º ciclo com uma duração de 3 anos. Apesar do Ensino Básico ser igual para todos os estudantes, algumas escolas oferecem ensino artístico, acrescentando ao currículo do estudante uma formação complementar na área artística (Eurydice – Comissão Europeia 2024).

Depois de concluir o Ensino Básico, o estudante passa para o Ensino Secundário, correspondendo a 3 anos de escolaridade, proporcionando formação em diversas áreas de ensino. Neste âmbito, existem os cursos Científico-Humanísticos, Profissionais, Artísticos Especializados e Científicos-Tecnológicos (Eurydice – Comissão Europeia 2024).

A avaliação é feita ao longo dos anos curriculares através de provas de aferição (Instituto de Avaliação Educativa (IAVE) 2024b). Estas provas realizam-se nos anos intermédios 2º, 5º e 8º dos respetivos ciclos (Instituto de Avaliação Educativa (IAVE) 2024b). O objetivo consiste em aferir os conhecimentos dos estudantes e agir atempadamente na melhoria das aprendizagens (Instituto de Avaliação Educativa (IAVE) 2024b). Os exames finais nacionais ocorrem no final do ensino básico, no 9º ano escolar, e no ensino secundário os estudantes dos cursos científico-humanísticos realizam quatro exames nacionais de conclusão durante os anos letivos 11º e 12º (Eurydice – Comissão Europeia 2024).

2.1.2 Programme for International Student Assessment (PISA)

Além das unidades curriculares, é fundamental considerar as competências desenvolvidas pelos estudantes após atingirem a idade da escolaridade obrigatória, ou seja, 15 anos (Instituto de Avaliação Educativa (IAVE) 2024a). O PISA, desenvolvido pela Organização para a Cooperação e Desenvolvimento Económico(OCDE), tem como objetivo avaliar as competências dos estudantes em enfrentar os desafios da vida quotidiana, bem como em literacia financeira e pensamento crítico (Instituto de Avaliação Educativa (IAVE) 2024a). Este programa também analisa a capacidade dos estudantes de mobilizar competências em leitura, matemática e ciências no contexto prático do dia a dia (Instituto de Avaliação Educativa (IAVE) 2024a).

A avaliação é realizada por meio de um teste informatizado, com diferentes versões que abrangem variados domínios (Instituto de Avaliação Educativa (IAVE) 2024a). Cada item do teste é estruturado de forma a incluir um conteúdo específico, processos cognitivos, um contexto real e um grau de dificuldade previamente estabelecido (Instituto de Avaliação Educativa (IAVE) 2024a).

O programa abrange cerca de 80 países e economias de todo o mundo, com a participação regular de Portugal (Instituto de Avaliação Educativa (IAVE) 2024a). O processo de seleção dos participantes é realizado em duas etapas (Instituto de Avaliação Educativa (IAVE) 2024a). Na primeira etapa, as escolas públicas e privadas são selecionadas aleatoriamente dentro de uma amostra previamente definida (Instituto de Avaliação Educativa (IAVE) 2024a). Na segunda etapa, procede-se à seleção aleatória de estudantes elegíveis dentro das escolas selecionadas (Instituto de Avaliação Educativa (IAVE) 2024a).

2.1.3 Entidades Responsáveis pelo Ensino Escolar

Os professores são os principais responsáveis no ensino escolar. No 1º ciclo do ensino básico a maioria dos professores ensina todas as disciplinas a uma designada turma. Nos seguintes ciclos o mais habitual é haver um professor por disciplina de ensino e um diretor(a) de turma.

Os encarregados de educação recorrem a centros de estudo, onde os estudantes poderão fazer os trabalhos de casa, estudar e por vezes participar em atividades extra curriculares. Na maioria dos casos, os centros de estudo para estudantes do 1.º ciclo funcionam também como um espaço onde os estudantes podem permanecer até que os pais terminem o horário de trabalho. Alguns centros de estudo possuem serviços de apoio a estudantes com cuidados mais especiais de educação (Aementes Brilhantes 2024).

Em áreas mais específicas de estudo, os estudantes e encarregados de educação recorrem a explicadores, com a possibilidade de sessões de explicações individuais ou em grupo (Tutor.pt 2024). Existe também suporte dedicado à preparação para os exames nacionais, centrado no esclarecimento de dúvidas e na capacitação dos estudantes para a realização dos exames.

Para além do ensino presencial, é possível obter explicações à distância. O projeto Sala do Futuro é um exemplo que facilita essa modalidade(Sala do Futuro 2024). A plataforma (Superprof 2024) é uma comunidade internacional de professores para ensino online, conta com profissionais em diversas áreas escolares, bem como em áreas artísticas e desportivas.

Familiares, amigos e colegas constituem uma alternativa ao estudo com profissionais especializados. Alguns encarregados de educação também dispõem de tempo para apoiar os filhos nos estudos. Quando o estudante adquire responsabilidade e autonomia suficientes para estudar de forma independente, é comum recorrer a amigos ou grupos de estudo formados por colegas de turma. Há também estudantes que optam por estudar sozinhos, com o apoio de conteúdos educacionais e dos professores nas escolas.

2.1.4 Ensino Superior

Após a conclusão do ensino secundário, os estudantes podem prosseguir os estudos para o ensino superior, classificado em ensino universitário e ensino politécnico (Direção-Geral do Ensino Superior 2024). Existem três ciclos de estudos de graus licenciado, mestre e doutor, e também um ciclo de estudos superior não conferente ao grau académico, denominado curso técnico superior profissional, correspondente ao ciclo de estudos curto ligado ao 1º ciclo previsto no Quadro de Qualificações do Espaço Europeu do Ensino Superior (Direção-Geral do Ensino Superior 2024).

2.2 Conteúdo Educacional Português

A Direção-Geral da Educação é responsável por planificar as diferentes provas e exames do sistema de ensino português, estabelecer as normas de classificação, bem como identificar o material didático, incluindo os manuais escolares necessários, assegurando as condições adequadas para as avaliações e a certificação (Direção-Geral da Educação 2011).

Os manuais escolares aprovados pela Direção-Geral da Educação são fornecidos por várias editoras especializadas. A *Leya* oferece manuais físicos, como também manuais digitais interativos, testes e exames de preparação nas diversas áreas do ensino escolar (Leya Aula Digital 2024). A Porto Editora e a Areal Editores são também exemplos de editoras que

disponibilizam manuais físicos, digitais e materiais de apoio escolar (Areal Editores 1984; Porto Editora 1994).

Os manuais escolares obrigatórios incluem frequentemente, livros de fichas para apoio ao estudo (Areal Editores 1984; LeYa 2008; Porto Editora 1994). Contudo em alguns casos, professores, centros de estudo e outras entidades elaboram fichas pedagógicas adicionais para auxiliar os alunos na aprendizagem. A DGE também disponibiliza apoio aos professores, orientando-os na organização e planeamento destas atividades (Direção-Geral da Educação (DGE) 2019).

Além de manuais, testes e fichas, existem outros recursos que podem ser utilizados para aprender, sendo também avaliados por especialistas na área da educação. Os jogos são uma dessas opções, incluindo tanto jogos digitais como físicos, como é o caso dos jogos de cartas *SuperTmatik*, também disponíveis online (Soared 2021; SuperTmatik s.d.).

Os filmes, desenhos animados e documentários constituem um complemento no ensino. Em particular os desenhos animados revelam-se especialmente eficazes no ensino primário, dado que nesta faixa etária os alunos demonstram maior interesse em assistir a conteúdos animados do que em ler livros, sobretudo quando ainda não sabem ler (Netflix 2015; Wikipedia contributors 1987).

2.3 Sistemas de Recomendação

Verifica-se uma grande variedade de métodos utilizados nos sistemas de recomendação, sendo os mais conhecidos os baseados em conteúdo, filtragem colaborativa, baseados em conhecimento, baseados em sessão descritos no diagrama da figura2.1 (Lucas et al. 2013). Estes métodos são mais utilizados em *e-commerce*. No entanto podem ser aplicados em outros domínios, incluindo educacional, quando combinados com outras abordagens e métodos específicos existentes na recomendação (Intayoad, Becker e Temdee 2017; Shinde e Kulkarni 2011)

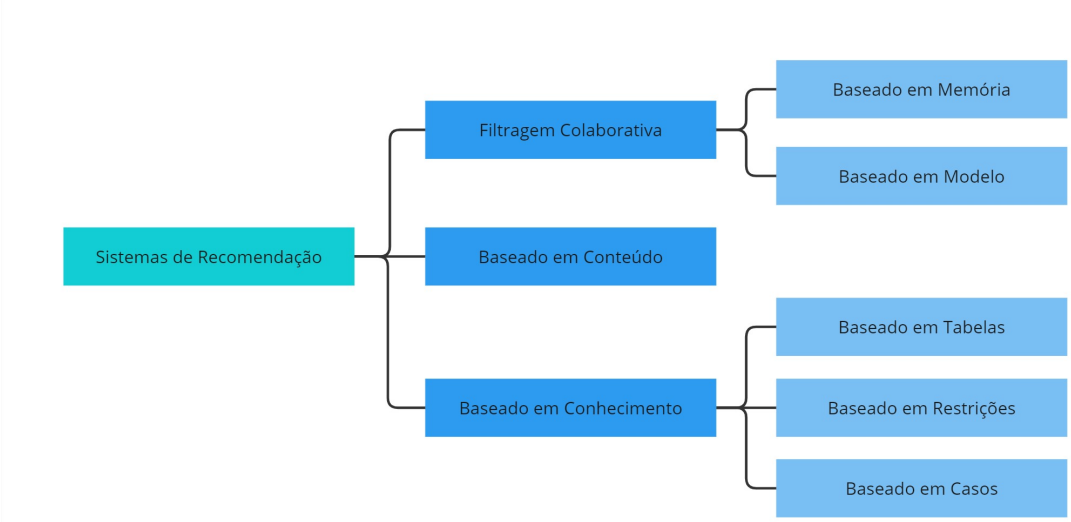

Figura 2.1: Métodos Utilizados na Recomendação do Conteúdo, Adaptado de (Martins e Figueiredo 2024)

2.3.1 Recomendação por filtragem colaborativa

O método filtragem colaborativa modela os interesses do utilizador para identificar similaridades entre utilizadores com interesses comuns, de uma amostra ou *dataset* de utilizadores, com o objetivo de recomendar conteúdos ainda não explorados por um utilizador específico (Barkan et al. 2020; Breese, Heckerman e Kadie 2013).

Existem duas abordagens, baseada em modelo na qual o algoritmo tenta comprimir uma grande base de dados em um modelo, e realiza a recomendação aplicando um mecanismo de inferência ao modelo, e a abordagem baseada em memória, onde o algoritmo carrega a base de dados na memória do sistema e faz recomendações com base nesses dados (Do, D. Nguyen e L. Nguyen 2010).

As abordagens mais comuns de recomendação por filtragem colaborativa incluem *clustering*, classificação, modelos latentes, Markov decision process (MDP)) e fatorização de matrizes (Do, D. Nguyen e L. Nguyen 2010). No contexto educacional, o *clustering* destaca-se como uma das ferramentas mais adequadas para fornecer recomendações, especialmente quando utilizado em conjunto com algoritmos apropriados como suplemento para melhorar o processo de *clustering* (Bagunaid, Chilamkurti e Veeraraghavan 2022; Bhaskaran e Marappan 2023).

Filtragem Colaborativa baseada em modelo

A abordagem compila as preferências dos utilizadores em modelos descritivos de utilizadores, itens ou classificações (*ratings*), com o objetivo de fazer previsões e recomendar itens personalizados. (Pennock et al. 2013). Para isso utiliza *feedback* explícito, por exemplo classificações de 1 a 5 estrelas, ou *feedback* implícito, as interações do utilizador com o sistema de recomendação, que inclui por exemplo a análise de comportamentos do utilizador, o tempo gasto a visualizar determinado artigo (Q. Zhao et al. 2019).

Na recomendação por filtragem colaborativa baseada em modelos, podem ser utilizados o agrupamento bayesian e as redes *bayesianas*, que são modelos probabilísticos (Breese, Heckerman e Kadie 2013).

Filtragem Colaborativa baseada em memória

Consiste em calcular a distância entre utilizadores ou *itens*, utilizando uma média ponderada das classificações, com o objetivo de calcular os valores de preferência dos utilizadores em relação aos itens recomendados (**eferenceCFModelMemory_1**).

Existem várias formas de calcular a distância, sendo as mais utilizadas a correlação de *Pearson* (2.1), normalmente utilizada para escalas de 1 a 5, e a similaridade do cosseno, aplicada em dados binários (**eferenceCFModelMemory_1**).

A correlação de *Pearson* (2.1) quantifica a associação entre variáveis, geralmente contínuas, e a sua escala varia de -1 a +1, sendo que 0 indica a ausência de associação entre as variáveis. (Schober, Boer e Schwarte 2018).

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{2.1}$$

A similaridade do cosseno modela o texto como vetores de termos, onde cada termo corresponde a uma dimensão de um vetor numa matriz (Rahutomo, Kitasuka e Aritsugi 2014). Um documento pode ser representado por um vetor **d** (2.2), e uma consulta *(query)* pode ser representada por um vetor **q** (2.3) (Rahutomo, Kitasuka e Aritsugi 2014). Os pesos w_{d_i} e w_{q_i} são números reais que indicam a frequência ou relevância de cada termo no documento e na consulta (Rahutomo, Kitasuka e Aritsugi 2014). Cada dimensão do vetor corresponde a um termo presente no documento ou na consulta (Rahutomo, Kitasuka e Aritsugi 2014).

$$\vec{d} = (w_{d_0}, w_{d_1}, \ldots, w_{d_k}) \tag{2.2}$$

$$\vec{q} = (w_{q_0}, w_{q_1}, \ldots, w_{q_k}) \tag{2.3}$$

A similaridade do cosseno (2.4) calcula a similaridade entre os dois vetores (Rahutomo, Kitasuka e Aritsugi 2014).

$$\text{Sim}(\vec{q}, \vec{d}) = \frac{\vec{q} \cdot \vec{d}}{|\vec{q}||\vec{d}|} = \frac{\sum_{k=1}^{t} w_{q_k} \times w_{d_k}}{\sqrt{\sum_{k=1}^{t} (w_{q_k})^2} \cdot \sqrt{\sum_{k=1}^{t} (w_{d_k})^2}} \tag{2.4}$$

Clustering

O processo de *clustering* consiste em dois passos (Do, D. Nguyen e L. Nguyen 2010). Inicialmente os utilizadores são organizados em grupos, denominados *clusters*, com base na similaridade dos seus gostos (Do, D. Nguyen e L. Nguyen 2010) . Para calcular a distância entre os utilizadores e identificar essas similaridades, utilizam-se métricas como a distância de *Minkowski* (2.5), Euclidiana (2.6) ou *Manhattan* (2.7), sendo que quanto menor a distância entre os utilizadores, maior a sua similaridade (Do, D. Nguyen e L. Nguyen 2010).

$$\text{distance}_{\text{Minkowski}}(\mathbf{u}_1, \mathbf{u}_2) = \left(\sum_j |r_{1j} - r_{2j}|^q \right)^{\frac{1}{q}} \tag{2.5}$$

$$\text{distance}_{\text{Euclidian}}(\mathbf{u}_1, \mathbf{u}_2) = \sqrt{\sum_j (r_{1j} - r_{2j})^2} \tag{2.6}$$

$$\text{distance}_{\text{Manhattan}}(\mathbf{u}_1, \mathbf{u}_2) = \sum_j |r_{1j} - r_{2j}| \tag{2.7}$$

No segundo passo, os utilizadores são agrupados nos *clusters*, utilizando técnicas como *k-centroid* e *k-means*, sendo esta última a mais amplamente utilizada (Do, D. Nguyen e L. Nguyen 2010; Júnior 2004)."

2.3.2 Recomendação baseada em conteúdo

Em comparação com a filtragem colaborativa, a recomendação baseada em conteúdo focase na análise do conteúdo que o utilizador demonstrou interesse (Umair Javed et al. 2021). Esta abordagem é amplamente utilizada em recomendações baseadas em texto, como documentos, descrições e títulos de produtos, incluindo exemplos como filmes, artigos e outros conteúdos. (Umair Javed et al. 2021).

Nesta abordagem, são utilizados modelos probabilísticos para realizar as recomendações, árvores de decisão ou redes neuronais (Ishak et al. 2023). Além disso, algoritmos de *machine learning* podem ser aplicados para inferir as preferências dos utilizadores com base em exemplos, utilizando descrições ricas em características do conteúdo analisado (Umair Javed et al. 2021).

O Term Frequency (TF) e o Inverse Document Frequency (IDF) são utilizados como unidades de informação e servem para determinar a relevância relativa de um registo, artigo, notícia, filme, entre outros conteúdos (U. Javed et al. 2021). O peso TF-IDF é elevado quando um termo aparece com frequência em um documento e é raro no conjunto de documentos (Umair Javed et al. 2021)

Por exemplo, no contexto do *e-learning*, pode ser utilizado para indicar os recursos que o estudante deve visitar (U. Javed et al. 2021). Segundo Umair Javed e restantes autores, consiste em prescrever diferentes endereços com conteúdo, criar o perfil ontológico do estudante ao extrair os seus interesses a partir do perfil do utilizador e desta forma, construir o sistema baseado em ontologia (U. Javed et al. 2021).

Uma ontologia é é uma inovação emergente para a representação semântica da informação (U. Javed et al. 2021). Consiste num modelo conceptual que pode ser utilizado para descrever o domínio de discurso, modelando o conjunto de conceitos e relações (U. Javed et al. 2021).

2.3.3 Recomendação baseada em conhecimento

Sistemas de recomendação baseados em conhecimento podem ser considerados um complemento à filtragem colaborativa e à recomendação baseada em conteúdo, evitando problemas de *cold start* (Burke 2000; Lorenzi e Ricci 2005; Towle e Quinn 2000).

Suportam a determinação de recomendações específicas em domínios complexos e de alto envolvimento (Uta et al. 2024). São capazes de considerar restrições, como serviços financeiros de alto risco, e fornecer explicações sobre as recomendações, mesmo em situações onde não seja possível identificar uma solução (Uta et al. 2024).

Existem três tipos de sistemas baseados em conhecimento, os baseados em tabelas, os baseados em restrições e os baseados em casos (Uta et al. 2024).

Baseado em tabelas

Os sistemas baseados em tabelas podem ser aplicados quando o conjunto de itens oferecidos é limitado, ou seja, quando o espaço de itens é relativamente pequeno, o que é frequente, por exemplo em recomendações de câmaras digitais ou serviços financeiros (Felfernig, Friedrich et al. 2006; Felfernig, Isak et al. 2007). Ao utilizar uma representação do conhecimento baseada em tabelas, podem ser executadas consultas à base de dados para identificar um

conjunto de candidatos a recomendações que atendam às preferências do utilizador (Felfernig, Friedrich et al. 2006; Felfernig, Le e Lubos 2023).

Baseado em restrições

Baseia-se na ideia de que o conhecimento de recomendação é representado em termos de um conjunto de variáveis e um conjunto correspondente de restrições (Uta et al. 2024). Os conjuntos de itens podem ser representados de forma intencional com base em um conjunto de restrições específicas do domínio (Uta et al. 2024).

Estas representações de conhecimento são especialmente úteis quando o espaço de soluções *(itens)* se torna intratável, ou seja, quando definir e gerir todas as alternativas é extremamente ineficiente e suscetível a erros, e as consultas de pesquisa associadas se tornam ineficazes e, no mínimo impraticáveis em contextos interativos (Falkner, Felfernig e Haag 2011).

Baseado em casos

A recomendação baseada em casos utiliza uma representação rica em conhecimento do domínio dos itens, e daí ser classificada como um tipo de recomendação baseada em conhecimento (Burke 2000; Khan e A. Hoffmann 2003; Lorenzi e Ricci 2005).

Num cenário básico, o conjunto de itens pode ser considerado como a base de casos, e as recomendações são determinadas ao identificar os itens dessa base que melhor correspondem às preferências do utilizador (Lorenzi e Ricci 2005).

2.3.4 Recomendação baseada em sessão

A recomendação baseada em sessão foca-se em aprender as preferências dos utilizadores das sessões associadas e criadas durante o processo de consumo de conteúdo com o objetivo de fornecer recomendações mais oportunas e precisas (Panagiotakis e Papadakis 2022).

A maioria das abordagens para recomendações baseadas em sessões utiliza aprendizagem sequencial, embora seja mais fácil de implementar o processo de mineração pode ser computacionalmente exigente (Cao et al. 2025).

Para implementar de forma mais sofisticada a recomendação baseada em sessões, é recomendável adotar abordagens de modelagem sequencial, geralmente baseadas em Markov Chain (MC), Reinforcement Learning (RL) e Markov Decision Processes (MDP), ou RNN(Cao et al. 2025; Hosseinzadeh Aghdam et al. 2015; Soh et al. 2017; Tavakol e Brefeld 2014).

2.3.5 Recomendação Híbrida

O objetivo da recomendação híbrida é combinar dois ou mais métodos de recomendação, de forma a superar as limitações de cada método individual e proporcionar resultados mais eficazes na recomendação (Martins e Figueiredo 2024). Um exemplo é utilizar a filtragem colaborativa para identificar os tópicos de estudo mais populares entre os estudantes e recorrer à recomendação baseada em conhecimento para sugerir conteúdos em que os estudantes apresentem maiores dificuldades.

Existem diversas formas de implementar um sistema de recomendação híbrido, dependendo da finalidade para a qual o sistema foi concebido e dos métodos utilizados. Alguns dos

métodos de recomendação híbrida incluem recomendação ponderada, alternância entre métodos de recomendação, combinação de métodos, fusão de características, abordagem em cascata, aumento de características e meta-level (Çano e Morisio 2017).

O método de recomendação híbrida ponderada 2.2 combina diferentes métodos de recomendação, atribuindo pontuações a cada conteúdo recomendado e integrando os resultados de cada método em uso através de uma fórmula linear, criando uma lista ou previsão das recomendações (Çano e Morisio 2017).

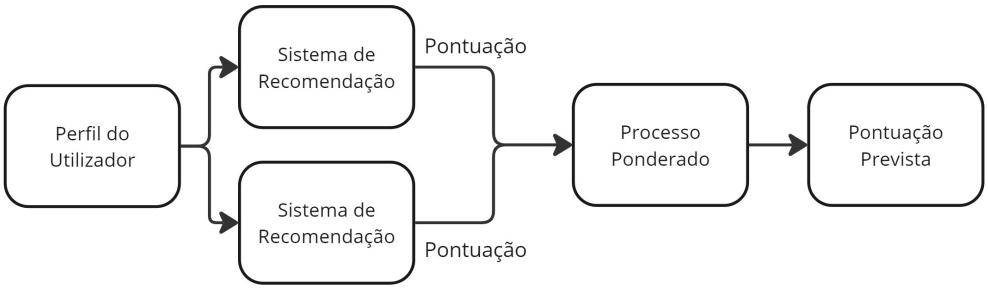

Figura 2.2: Método de Recomendação Híbrida Ponderada
(Chiang 2021; Martins e Figueiredo 2024)

O método de recomendação híbrida que utiliza alternância entre técnicas 2.3 baseia-se na situação corrente de alternar entre métodos de recomendação (Çano e Morisio 2017). Utilizando o exemplo anterior dos estudantes, poderia-se usar a filtragem colaborativa quando o sistema ainda não tem muita informação sobre o utilizador e, à medida que adquire mais dados, mudar para a recomendação baseada em conhecimento.

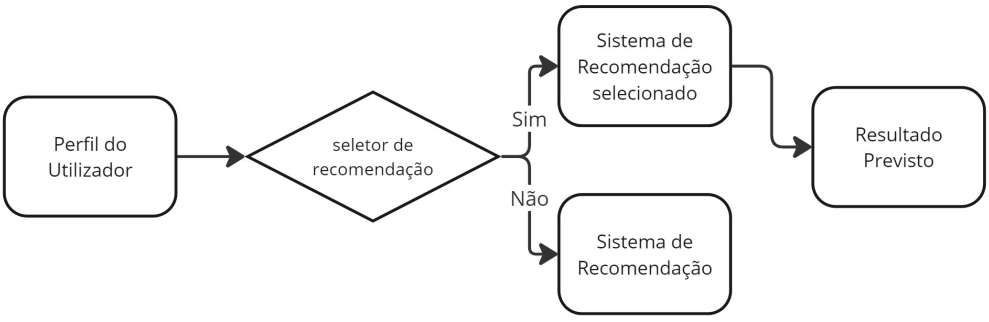

Figura 2.3: Método de Recomendação Híbrida, Alternância
(Chiang 2021; Martins e Figueiredo 2024)

O método de recomendação híbrida que combina diferentes métodos 2.4 é utilizado quando é necessário fazer um grande número de recomendações, evitando o problema associado à falta de conhecimento sobre o utilizador (Çano e Morisio 2017). No caso de um estudante

ser novo na plataforma, esse método oferece uma solução, uma vez que não há necessidade de questionar o utilizador, o que ajuda a contornar o problema de *cold start* (Lucas et al. 2013). Utilizando o exemplo de filtragem colaborativa e recomendação baseada em conhecimento, caso não haja nenhum conhecimento prévio sobre o utilizador, a lista desse método não retornaria resultados, exibindo apenas os resultados da filtragem colaborativa. Caso contrário, mostraria os resultados de ambos os métodos.

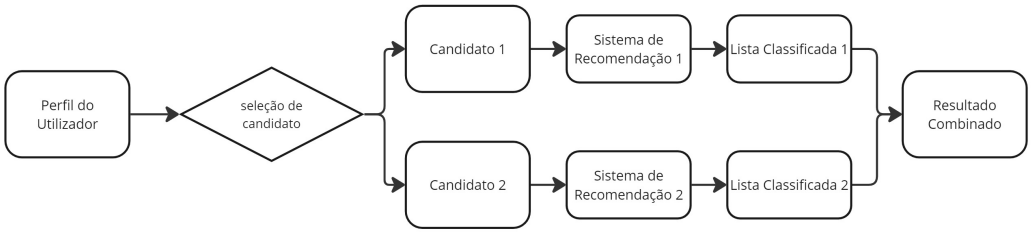

Figura 2.4: Método de Recomendação Híbrida, Combinação de Diferentes Métodos
(Chiang 2021; Martins e Figueiredo 2024)

O método de recomendação híbrida que utiliza fusão de características 2.5 combina todas as características em um único algoritmo, utilizando um dos métodos de recomendação como saída, atuando como fornecedor adicional de dados de características (Çano e Morisio 2017).

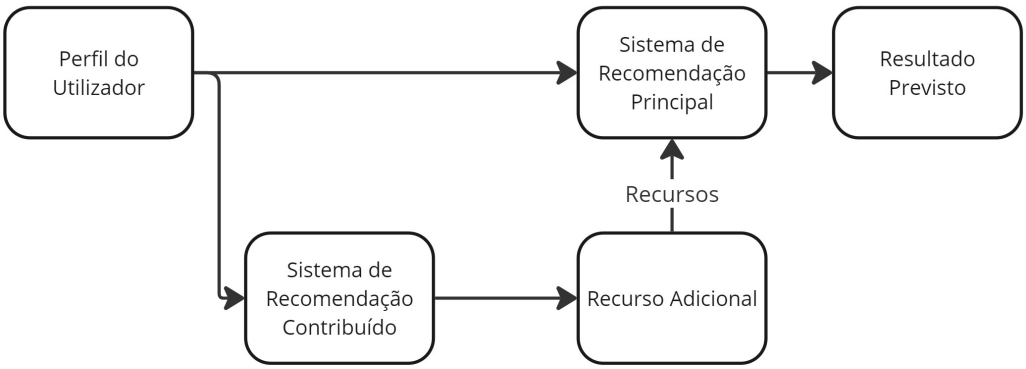

Figura 2.5: Método de Recomendação Híbrida, Fusão de Características
(Chiang 2021; Martins e Figueiredo 2024)

O método de recomendação híbrida em cascata utiliza a lista de resultados obtidos no primeiro método de recomendação e refina a recomendação no próximo método, filtrando os resultados do método anterior (Çano e Morisio 2017). Esta abordagem impede a aplicação do segundo método a itens que foram indubitavelmente eliminados pelo primeiro método de filtragem (Çano e Morisio 2017). No contexto educacional, o primeiro método poderia recomendar tópicos nos quais o estudante apresenta dificuldades, enquanto que o segundo

filtraria, mostrando apenas o conteúdo educacional de maior preferência, por exemplo mais vídeos explicativos do que exercícios.

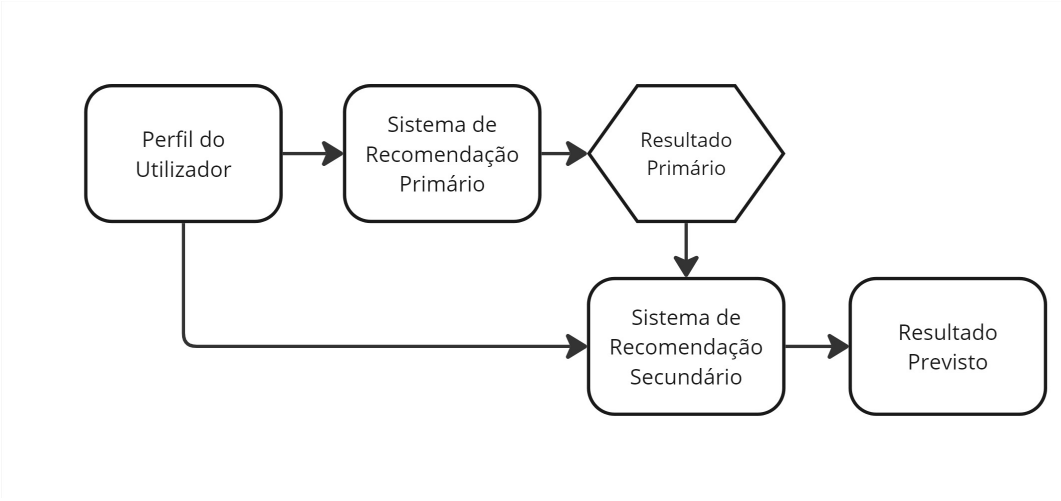

Figura 2.6: Método de Recomendação Híbrida em Cascata
(Chiang 2021; Martins e Figueiredo 2024)

O método de recomendação híbrida de aumento de características 2.7 utiliza o primeiro método de recomendação para criar uma classificação ou pontuação de um item, que é integrado como entrada no segunda método de recomendação (Çano e Morisio 2017).

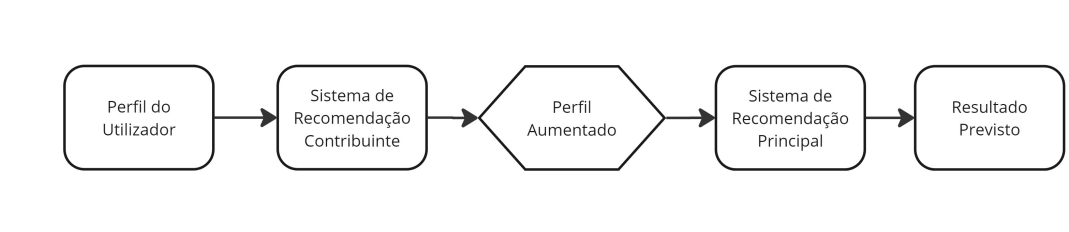

Figura 2.7: Método de Recomendação Híbrida, Aumento de Características
(Chiang 2021; Martins e Figueiredo 2024)

O método de recomendação híbrido meta-level 2.8 é semelhante ao método de aumento de características, no qual o primeiro método tipicamente baseado em conteúdo, é responsável pela construção de modelos de representação de *itens* (Çano e Morisio 2017). Esses modelos são então utilizados como entrada para o segundo método, geralmente a filtragem colaborativa, com o objetivo de associar os *itens* aos perfis dos utilizadores (Çano e Morisio 2017).

Figura 2.8: Método de Recomendação Híbrida, Meta-Level

2.3.6 Características e Desafios dos Métodos de Recomendação

Os métodos mais utilizados para recomendar conteúdo também possuem desafios e desvantagens na implementação.

A filtragem colaborativa destaca-se pela capacidade de fornecer recomendações eficazes sem exigir um conhecimento detalhado sobre os itens recomendados, sendo ainda um dos métodos que alcançam os melhores resultados de acurácia (Phalle e Bhushan 2024).

No entanto, analisa grandes volumes de dados, consumindo muitos recursos do sistema e exigindo um tempo considerável para o processamento (Do, D. Nguyen e L. Nguyen 2010). Além disso, pode gerar recomendações imprevisíveis, uma vez que mesmo que um item seja comum entre vários utilizadores, ele pode não ser compatível com o perfil do utilizador que recebeu a recomendação (Do, D. Nguyen e L. Nguyen 2010). Quando há escassez de conhecimento sobre os utilizadores, o problema de *cold start* pode tornar a implementação deste método desafiadora (Phalle e Bhushan 2024).

A recomendação baseada em conteúdo é mais eficaz e fácil de implementar em comparação com a filtragem colaborativa, destacando-se pela sua transparência, independência e capacidade de recomendar itens não classificados (U. Javed et al. 2021). Além disso, oferece recomendações personalizadas ao utilizador, uma vez que se foca exclusivamente no conteúdo do item ou nos interesses do próprio utilizador (U. Javed et al. 2021). Enfrenta problemas de *cold start*, menos graves do que na filtragem colaborativa, contudo, os algoritmos podem ser conservadores, tendendo a recomendar categorias familiares e evitando explorar novas opções potencialmente interessantes (U. Javed et al. 2021).

A recomendação baseada no conhecimento pode ser utilizada para superar algumas das limitações da filtragem colaborativa e da recomendação baseada em conteúdo, uma vez que não exige conjuntos de dados extensivos para fornecer recomendações, sendo também menos suscetível a problemas de *cold start* (Bouraga et al. 2014). Contudo, apresenta uma limitação significativa na construção da base de conhecimento, que requer um conhecimento aprofundado do domínio e experiência considerável em representação de conhecimento (Bouraga et al. 2014).

Na recomendação baseada em sessão, as recomendações são fornecidas mais rapidamente (Jannach, Quadrana e Cremonesi 2022). No entanto quando se observa uma sequência muito curta de interações durante a sessão em curso, as preferências de longo prazo dos utilizadores não estão disponíveis e a intenção do utilizador pode variar de sessão para sessão (Jannach, Quadrana e Cremonesi 2022).

Sistemas de Recomendação Educacionais

As Tabelas 2.1 2.2 apresentam os sistemas e a arquitetura propostos por outros autores, destacando os elementos utilizados na implementação, os resultados obtidos e as direções sugeridas para trabalhos futuros.

Tabela 2.1: Comparação de Sistemas de Recomendação

Autores	**Sistema/Abordagem**	**Características/Descrição**	**Resultados/Conclusões**
Abdi, Khosravi, Sadiq e Gasevic 2020	Sistema de recomendação de conteúdo educacional baseado em Open Learner Model (OLM) integrado no sistema RiPPLE	Duas interfaces de estudo, uma não complementada que fornece um conjunto de filtros, e uma complementada com o OLM	A interface complementada aumentou o envolvimento, compreensão e confiança dos estudantes, mas apresentou dificuldades de navegação para alguns dos estudantes
Shishehchi et al. 2012	Sistema de recomendação que sugere materiais de aprendizagem ao estudante	Baseado em conhecimento e ontologias. O sistema analisa o estilo de aprendizagem do utilizador e avalia o seu conhecimento prévio	A abordagem só é validade teoricamente
Bouihi e Bahaj 2019	Arquitetura clássica de aplicação web de 3 camadas, com a adição de uma camada semântica *Semantic Web Rule Language (SWLR)* para recomendar conteúdo educacional	A camada semântica contém dois subsistemas semânticos, um baseado em ontologia para modelar o conteúdo de aprendizagem e o contexto do aluno, e outro baseado em regras *Semantic Web Rule Language (SWLR)* como um método de recomendação e filtragem	Apenas mencionam os elementos utilizados na arquitetura proposta, sem apresentação validações práticas ou testes
Obeid et al. 2018	Recomendação baseada em ontologia de cursos superiores e universidades, a estudantes do ensino secundário	Sistema híbrido que utiliza ontologias criadas no *Protégé* para modelar o conhecimento do domínio e técnicas de *machine learning* para agrupar e criar modelos de perfis de estudantes graduados	Não mencionam a realização de testes, propondo para trabalhos futuros a utilização de *machine learning* para analisar os perfis de estudantes graduados, com o objetivo de propor os N melhores cursos e universidades relacionados a um novo estudante com base no perfil.

Tabela 2.2: Comparação de Sistemas de Recomendação

Autores	**Sistema/Arquitetura**	**Abordagens implementadas**	**Resultados/Conclusões**
Rabahallah, Mahdaoui e Azouaou 2018	Sistema para recomendar Massive Open Online Courses (MOOCs) a alunos online	Seguiu as etapas: Criação de ontologias para representar o conhecimento do aluno e dos MOOCs; Cálculo de similaridades com base nas ontologias e classificações, bem como a realização de previsões para o aprendiz-alvo; Seleção dos N melhores MOOCs através do motor de recomendação de filtragem colaborativa baseado em memória	Não mencionam a realização de testes, propondo para trabalhos futuros a integração de outras técnicas, como o *Support Vector Machine (SVM)*
Mbaye 2018	Sistema de recomendação para recomendar materiais de estudo aos alunos	Utilização de ontologias para representar o conhecimento do aprendiz, recomendação baseado em fitragem colaborativa e algoritmo de decisão para descobrir padrões de aprendizagem	Não mencionam a realização de testes
Bagunaid, Chilamkurti e Veeraraghavan 2022	Sistema de recomendação inteligente para assistência académica aos estudantes (Artificial Intelligence-based Student Assessment and Recommendation (AISAR))	Cálculo de pontuações, *clustering* e previsão de desempenho, além de fornecer recomendações personalizadas	Utilizaram métricas obtendo resultados superiores aos sistemas de recomendação existentes que utilizam *machine learning*. Concluíram que o *clustering* pode ser uma solução viável para prever o desempenho de estudantes. No entanto, apenas avalia o desempenho de grupos de estudantes, e não de cada estudante.
Bhaskaran e Marappan 2023	Sistema híbrido para resolver os problemas dos métodos de recomendação atuais, permitindo que os alunos utilizem o material de ensino organizado em cursos adequados	*Modifed One Source Denoising (MOSD)* ,*Modifed Anarchic Society Optimization (MASO)*, *Sequencial Generalizado Aprimorado (EGSP)*, *clustering* com utilização de Density-Based Spatial Clustering of Applications with Noise (DBSCAN) e Enhanced Clustering (EC)	Avaliado utilizando conjuntos de dados de diversas áreas e métricas. Os autores concluíram que o desempenho do sistema é superior ao de outros métodos de recomendação.

O sistema Ripple usado no estudo por (Abdi, Khosravi, Sadiq e Gasevic 2020) é um sistema adaptativo de aprendizagem a nível de curso, denominado RiPPLE, que utiliza *crowdsourcing* para o desenvolvimento de atividades de aprendizagem (Khosravi, Kitto e Joseph 2019). O RiPPLE gera um modelo de aprendizagem com base no sistema de classificação Elo para estimar o estado de conhecimento de um estudante (Abdi, Khosravi, Sadiq e Gašević 2019).

As regras utilizadas por (Bouihi e Bahaj 2019) foram divididas em quatro categorias. *Learning History Rules(LHR)*, utilizadas para recomendar ao aluno materiais que sejam coerentes com o seu histórico de aprendizagem. *Learning Performance Rules (LPR)*, aplicadas para recomendar objetos de aprendizagem que estejam alinhados com o *feedback* e classificações de desempenho do aprendiz. *Learning Social Network Rules (LSNR*), que são usadas para recomendar objetos de aprendizagem com base nas interações sociais e relacionamentos do aprendiz. *Learning Pathway Rules(LPR)*, empregadas para determinar a ordem dos materiais de aprendizagem recomendados em um percurso de aprendizagem.

No sistema desenvolvido por (Bagunaid, Chilamkurti e Veeraraghavan 2022) as pontuações são estimadas utilizando *Recurrent Neural Networks (RNN)*, que integram as etapas de pré-processamento, extração de características e estimativa de pontuações. Para o *clustering*, utilizaram a distância de *Mahalanobis*, com a aplicação do algoritmo DBSCAN para agrupar estudantes com base no seu desempenho, medido em termos de pontuações. A identificação do desempenho de cada estudante é realizada através de um processo *threshold-based MapReduce*, que categoriza o desempenho individualmente. As recomendações para os estudantes são criadas por um algoritmo de aprendizagem por reforço conhecido como *State–action–reward–state–action (SARSA)*.

No sistema desenvolvido por (Bhaskaran e Marappan 2023), a técnica *Modifed One Source Denoising (MOSD)* foi aplicada para pré-processar o conjunto de dados dos alunos, a estratégia *Modifed Anarchic Society Optimization (MASO)* utilizada para melhorar as medições de desempenho e a estratégia de Padrão *Sequencial Generalizado Aprimorado (EGSP)* aplicada para minerar padrões sequenciais de aprendizagem, avaliando assim os hábitos e interesses dos alunos. Dependendo dos diferentes modelos de métodos de aprendizagem e do comportamento dos alunos, a agrupação é realizada utilizando o algoritmo de *clustering* mais amplamente utilizado, o DBSCAN aprimorado e EC para ajudar na identificação de *clusters* de diferentes tamanhos e formas.

Em comparação com o sistema de recomendação proposto por Wala Bagunaid e demais autores (Bagunaid, Chilamkurti e Veeraraghavan 2022), o modelo desenvolvido por S. Bhaskaran e Raja Marappan (Bhaskaran e Marappan 2023) utiliza o EC como uma abordagem complementar ao algoritmo DBSCAN, aprimorando ainda mais o processo de agrupamento.

Os autores que avaliaram o sistema, utilizaram métricas como *Mean Absolute Error(MAE)*, acurácia, tempo de processamento de consultas, *ranking score*, *recall* e precisão e taxa de falsos positivos, taxa de falsos negativos(Bagunaid, Chilamkurti e Veeraraghavan 2022; Bhaskaran e Marappan 2023).

2.4 Large Language Models (LLMs)

Os LLMs são modelos de aprendizagem profunda capazes de reconhecer, extrair, resumir, prever e criar texto com base no conhecimento adquirido durante o treino em conjuntos de dados de grande dimensão (Chockalingam et al. 2023).

Os LLMs estabeleceram-se como ferramentas versáteis para uma ampla variedade de tarefas, incluindo tradução automática, resposta a perguntas e geração de texto (Hadi et al. 2023). Destacam-se pela utilização do efeito de escala na ampliação de sua capacidade, posicionando-se como ferramentas de uso geral para a resolução de tarefas diversas (W. X. Zhao et al. 2024).

O maior avanço nos LLMs ocorreu com a introdução da arquitetura *Transformer*, baseada no mecanismo de autoatenção, o que possibilitou a paralelização e o processamento eficiente de relações entre elementos distantes no texto (Vaswani et al. 2017). Esta arquitetura estabeleceu as bases para modelos como o GPT da OpenAI e o BERT do Google (Hadi et al. 2023; Vaswani et al. 2017).

Treinar LLMs exige grande processamento de dados e a execução de treino paralelo distribuído, o que requer algoritmos de treino distribuído para a aprendizagem dos parâmetros da rede neuronal, frequentemente utilizando de forma conjunta diversas estratégias de paralelização (W. X. Zhao et al. 2024).

Os LLMs podem ser utilizados em diversas áreas e para múltiplos propósitos. Wayne Xin Zhao e demais autores (W. X. Zhao et al. 2024) resumiram as áreas de aplicação mais comuns e as direções de investigação associadas. As áreas de domínio mais frequentes incluem saúde, finanças, investigação científica, direito e educação. Na educação os LLMs, oferecem oportunidades para aprimorar as experiências de aprendizagem, apoiar os professores e desenvolver conteúdo educacional (W. X. Zhao et al. 2024). Para os estudantes, os LLMs podem fornecer materiais de estudo personalizados e questões práticas com base em informações sobre desempenho e preferências fornecidas, criando experiências de aprendizagem individualizadas (Dai et al. 2023; W. X. Zhao et al. 2024).

No campo das direções de investigação, as abordagens clássicas concentram-se em tarefas de Natural Language Processing (NLP), modelos de (Information Retrieval (IR)) e sistemas de recomendação (W. X. Zhao et al. 2024). Em relação a capacidades avançadas, destacam-se os LLMs multimodais e aqueles enriquecidos com grafos de conhecimento (W. X. Zhao et al. 2024). Além disso, emergem novas linhas de investigação, como o desenvolvimento de agentes baseados em LLMs e a aplicação de LLMs para fins de avaliação (W. X. Zhao et al. 2024).

2.4.1 Componentes de um LLM

Um LLM requer componentes essenciais para garantir um funcionamento eficiente (Naveed et al. 2024). Entre esses componentes, destacam-se alguns deles 2.9 (Naveed et al. 2024):

- **Tokenização**: É uma etapa essencial de pré-processamento no treino de LLMs, que consiste em dividir o texto em unidades indivisíveis denominadas *tokens*. (Kudo 2018; Sennrich, Haddow e Birch 2016; Webster e Kit 1992).
- **Codificação Posicional**: Para capturar a informação de posição no texto, foram introduzidos os *encodings* posicionais, como os *embeddings* absolutos, relativos ou aprendidos (Vaswani et al. 2017). Dentro dos *encodings* relativos, Alibi e RoPE são dois tipos amplamente utilizados em LLMs (Naveed et al. 2024).
- **Atenção em LLMs**: A atenção atribui pesos aos *tokens* de entrada com base na sua importância entre palavras, permitindo que o modelo dê mais ênfase aos *tokens* relevantes (Vaswani et al. 2017).
- **Funções de Ativação**: As funções de ativação desempenham um papel crucial nas capacidades de ajuste de curva das redes neurais (Hornik, Stinchcombe e White 1989). As funções de ativação utilizadas em LLMs incluem a *rectified Linear Unit (ReLU)*, a *Gaussian Error Linear Unit (GeLU)* e variantes como a *Gated Linear Unit (GLU)* (Naveed et al. 2024).
- **Normalização de Camadas**: A normalização de camadas contribui para uma convergência mais rápida (Vaswani et al. 2017). Os LLMs utilizam normalização pré-camada, aplicando-a antes da *Multi-Head Attention (MHA)* (Baevski e Auli 2018; Naveed et al. 2024).
- **Pré-processamento de Dados**: As técnicas de pré-processamento de dados em LLMs incluem filtragem de qualidade ,utilizando abordagens baseadas em classificadores ou heurísticas, de-duplicação de dados para evitar a memorização e redução de privacidade para eliminar informações pessoais sensíveis dos dados de treinamento (Naveed et al. 2024).

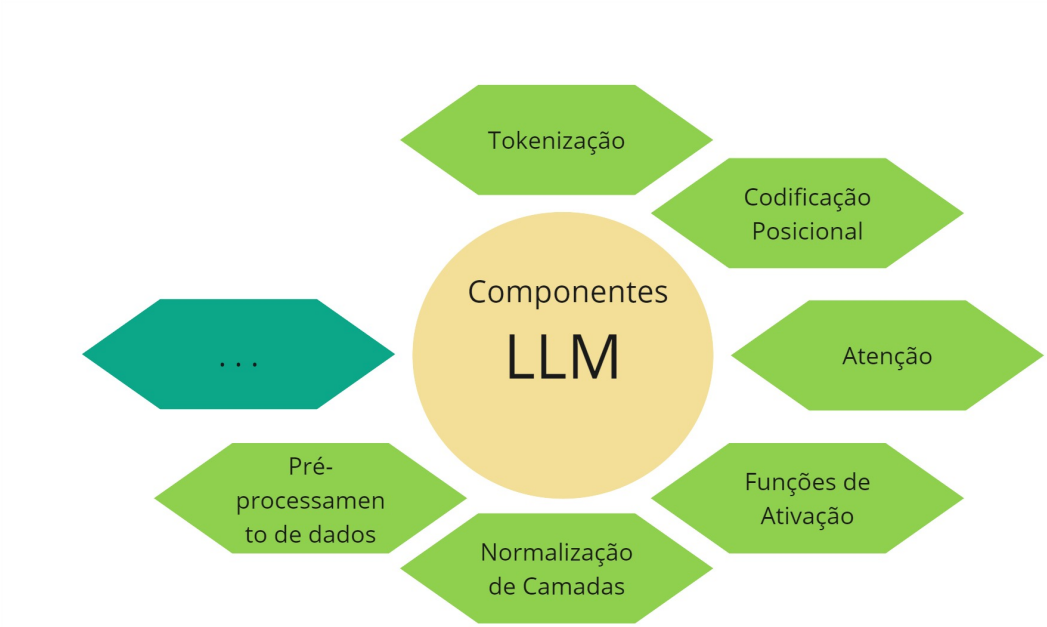

Figura 2.9: Componentes dos LLMs

2.4.2 Desafios

Os LLMs enfrentam um conjunto de desafios significativos que não podem ser ignorados. Para abordar essas limitações de forma eficaz, é fundamental reconhecê-las (Naveed et al. 2024).

- **Custos:** Não só em LLMs, é um problema enfrentado por vários programadores nas áreas de *machine learning* e IA. Os LLMs exigem elevados investimentos em custos de produção, recursos computacionais significativos e consumo intensivo de energia durante o processo de treino em larga escala (Naveed et al. 2024). Uma forma de reduzir o custo computacional consiste na diminuição do conjunto de dados utilizado no treino. Caso isso não seja viável, o treino pode ser realizado de forma escalonada, em etapas progressivas, ou por meio da utilização de infraestrutura em nuvem. O *Amazon SageMaker* (Amazon Web Services 2024a) é um desses exemplos, oferecendo recursos otimizados para treino de modelos de larga escala.

- **Controlo, Segurança e Privacidade nos LLMs:** Os LLMs podem por vezes, criar conteúdo inadequado, prejudicial ou não verídico, seja por acidente ou devido a *prompts* específicos (Naveed et al. 2024). Além disso, estão sujeitos a riscos como a exposição de informações pessoais, criação de conteúdos antiéticos, ataques de *backdoor*, *jailbreaks*, injeção de *prompts* e envenenamento de dados (Das, Amini e Wu 2024; Naveed et al. 2024). A Microsoft oferece o serviço *Azure Content Moderator* (Microsoft 2024), uma solução projetada para detetar conteúdos prejudiciais criados por utilizadores ou por IA em aplicações e serviços. Este serviço realiza uma análise minuciosa de conteúdos utilizando Application Programming Interface (API), identificando linguagem inadequada, imagens explícitas e outros elementos potencialmente sensíveis (Microsoft 2024). Também fornece ferramentas para personalização de filtros e suporte para revisões humanas, ajudando a garantir que os conteúdos estejam em

conformidade com padrões éticos e legais (Microsoft 2024). Adicionalmente, a implementação de políticas de segurança, como a conformidade com o RGPD no contexto da União Europeia é outra abordagem. Essas práticas, combinadas e a implementação de mecanismos de defesa (Das, Amini e Wu 2024), garantem a segurança e confiança nos LLMs.

- **Viés e Justiça:** Os LLMs podem herdar e amplificar os vieses sociais presentes nos seus dados de treino (Naveed et al. 2024). Esses vieses podem manifestar-se nas saídas do modelo, resultando em potenciais questões éticas e desafios relacionados à justiça (Bender et al. 2021).

- **Overfitting**: Os LLMs são suscetíveis ao *overfitting*, aprendendo padrões ruidosos e específicos nos dados de treino extensivos, o que pode resultar em respostas ilógicas (Naveed et al. 2024; C. Zhang et al. 2021). Para mitigar esse problema, é essencial equilibrar memorização e generalização, pois excesso de memorização pode tornar o modelo inflexível e incapaz de lidar com novos (Tänzer, Ruder e Rei 2021). A memorização permite ao modelo reter detalhes específicos dos dados de treino, garantindo respostas precisas a perguntas concretas, enquanto que a generalização capacita o modelo de inferir e responder a novas entradas (Naveed et al. 2024).

- **Raciocínio e Planeamento:** Algumas tarefas, aparentemente simples e facilmente executadas por humanos, ainda estão além das capacidades dos atuais LLMs (Naveed et al. 2024). Isso ocorre porque esses modelos baseiam-se principalmente na previsão da próxima palavra com base na probabilidade dos dados de treino, não sendo projetados para garantir habilidades de raciocínio lógico e estruturado de forma sólida (Naveed et al. 2024; Valmeekam et al. 2022).A implementação de regras ajuda os LLMs a atuar de forma eficaz em situações específicas que não possuem capacidade de raciocínio, resultando em resultados mais precisos e eficazes.

- **Alucinações:** As alucinações em LLMs ocorrem quando o modelo produz respostas que embora pareçam plausíveis, são incorretas e não se baseiam nas informações fornecidas (Y. Zhang et al. 2023). Existem três tipos principais de alucinações (Naveed et al. 2024). A alucinação contraditória com a entrada, quando o conteúdo gerado pelo modelo se desvia da informação original dada pelo utilizador (Naveed et al. 2024). A alucinação contraditória com o contexto, que ocorre quando o modelo cria respostas que contrariam informações previamente geradas por ele mesmo (Naveed et al. 2024). A alucinação factual, onde o modelo produz respostas que não coincidem com o conhecimento factual amplamente reconhecido (Naveed et al. 2024). Uma solução que pode ser implementada para ultrapassar esse problema em chatbots baseados em LLM é a utilização de engenharia de *prompts*. (Dam et al. 2024; Wang et al. 2023)

- **Conhecimento Limitado:** A informação adquirida durante o pré-treinamento é limitada e pode tornar-se obsoleta com o tempo (Naveed et al. 2024). O uso de *pipeline* de aumento de recuperação permite gerar respostas precisas, no entanto, os modelos pré-treinados não são projetados para trabalhar com Retrieval Augmentation Generation (RAG), o que limita a capacidade do modelo de aceder informações atualizadas de forma direta (Borgeaud et al. 2022; Brown et al. 2020; Ram et al. 2023; Saravia 2022; Touvron et al. 2023).

- **Esquecimento Catastrófico:** Quando os LLMs são pré-treinados com grandes conjuntos de dados e posteriormente ajustados *(fine-tuned)* para um domínio específico,

há uma redução nos recursos de treino (Naveed et al. 2024). No entanto, eles enfrentam problemas como a adaptação ao domínio e o esquecimento catastrófico, que dificultam a retenção do conhecimento original ao aprender novas tarefas (Naveed et al. 2024).

- **Interpretabilidade e Explicabilidade:** A natureza "caixa-preta"dos LLMs representa um desafio na compreensão dos seus processos de tomada de decisão, o que é fundamental para garantir uma aceitação mais ampla e confiante, especialmente em domínios sensíveis (Naveed et al. 2024).
- **Processamento em Tempo Real:** Os LLMs frequentemente possuem centenas de camadas e milhões de parâmetros, o que dificulta o processamento em tempo real devido às altas exigências computacionais e ao armazenamento limitado de pesos nas plataformas de hardware, especialmente em ambientes de computação na borda *(edge computing)* (Naveed et al. 2024).
- **Dependências de Longo Prazo:** Embora os LLMs tenham mostrado progresso significativo na compreensão e criação de texto, frequentemente enfrentam dificuldades em preservar o contexto e lidar com dependências de longo prazo, especialmente em conversas complexas de múltiplos turnos ou em documentos longos. Esta limitação pode resultar em respostas incoerentes ou irrelevantes (Naveed et al. 2024).

2.4.3 Retrieval-Augmented Generation (RAG)

O RAG representa uma estrutura técnica que integra sistemas de busca de informação com modelos de síntese textual, facilitando a consulta de repositórios de conhecimento externos durante a produção de respostas (P. Lewis et al. 2020). É uma das soluções para ultrapassar diversos desafios nos LLMs, como alucinações, conhecimento desatualizado, processos de raciocínio não transparentes e de difícil rastreamento (Yunfan Gao et al. 2024). Além disso, é o processo de otimizar a saída de um grande modelo de linguagem, de modo a garantir referência a uma base de conhecimento confiável fora das fontes de dados de treino antes de criar uma resposta (Amazon Web Services 2024b).

O RAG reduz efetivamente o problema da geração de conteúdo factualmente incorreto, ao mesmo tempo que proporciona melhores informações para os LLMs responderem a tarefas mais complexas e exigentes em termos de conhecimento durante a fase de inferência, o que tem levado a um desenvolvimento rápido dos estudos sobre RAG (Yunfan Gao et al. 2024).

A imagem 2.10 descreve a tecnologia de pesquisa RAG, a sua evolução e utilização com o GPT, estando dividida nas em três fases, pré-treinamento, ajuste fino e inferência (Yunfan Gao et al. 2024).

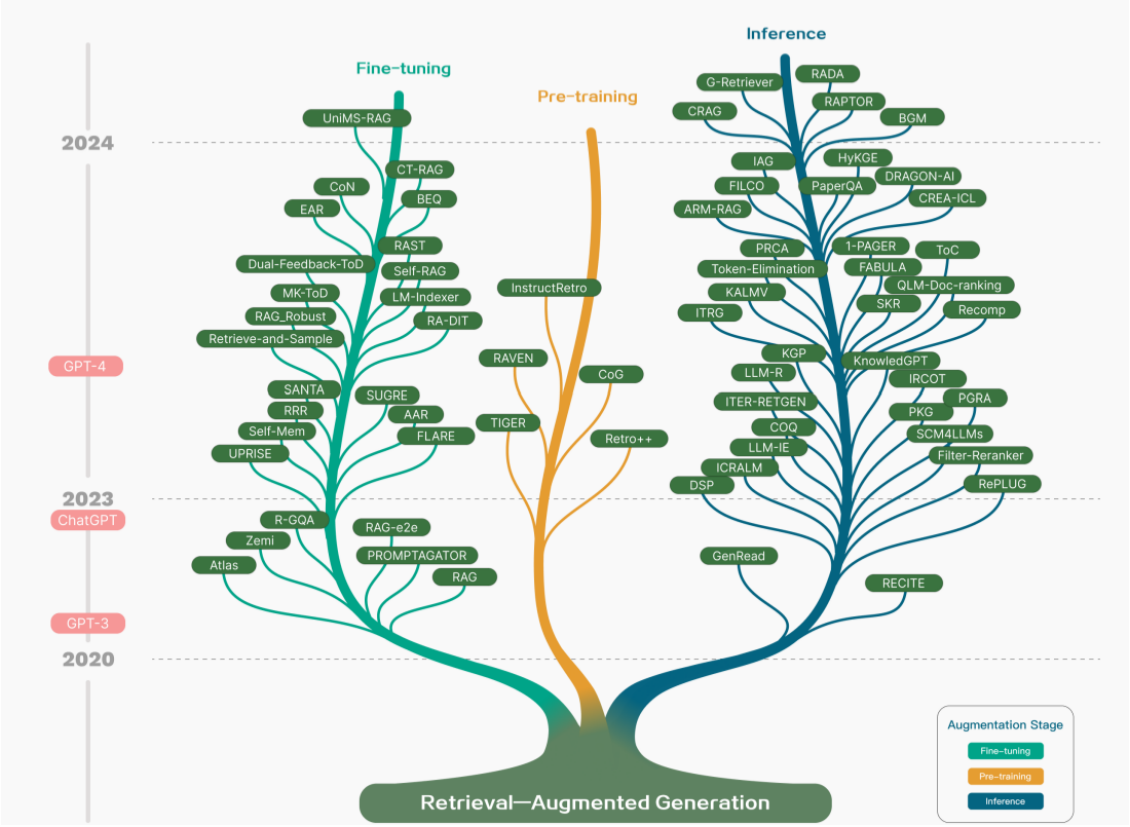

Figura 2.10: Retrieval-Augmented Generation (RAG)
(Yunfan Gao et al. 2024)

O paradigma de pesquisa RAG está em constante evolução e pode ser categorizado em três estágios, RAG Ingênuo, RAG Avançado e RAG Modular (Yunfan Gao et al. 2024). Embora a abordagem RAG seja eficiente em termos de custo e apresente desempenho superior ao dos LLMs nativos, ela também apresenta limitações (Yunfan Gao et al. 2024).

2.4.4 Grafos de Conhecimento

Os grafos de conhecimento permitem organizar a informação e gerir o conhecimento a partir de grandes volumes de dados, aprimorando a qualidade dos serviços de informação e oferecendo aos utilizadores serviços mais inteligentes (X. Chen, Jia e Xiang 2020). Têm sido amplamente utilizados em sistemas de pergunta-resposta, motores de busca e sistemas de recomendação, com o objetivo de identificar erros e inferir novas conclusões a partir dos dados já existentes (X. Chen, Jia e Xiang 2020).

A utilização em LLMs contribui para ultrapassar problemas relacionados com a precisão factual e as alucinações, proporcionando uma estrutura de representação de conhecimento, com entidades codificadas e as suas relações num formato legível por máquina (Bui et al. 2024). Nos grafos de conhecimento, os nós representam as entidades, enquanto que as arestas indicam as relações entre elas. Cada nó e aresta deve estar associado a uma propriedade ou atributo, fornecendo detalhes adicionais ou metadados (Carta et al. 2023).

Os LLMs também podem ser utilizados para criar grafos de conhecimento, analisando e criando texto em grande escala (Carta et al. 2023). Este processo auxilia na extração de entidades, relações, desambiguação ou inferência textual (Carta et al. 2023).

Também existe os chamados índices baseados em grafos de conhecimento, que podem ser utilizados em sistemas RAG (Yunfan Gao et al. 2024). Uma vantagem dessa abordagem é a transformação do processo de recuperação de informações em instruções compreensíveis pelos LLMs, o que aprimora a precisão na recuperação do conhecimento e permite que o LLM crie respostas coerentes, melhorando a eficiência geral do sistema RAG (Yunfan Gao et al. 2024).

2.4.5 Integração de Grafos de Conhecimento e RAG no Ensino

Após a análise dos RAGs e grafos de conhecimento em LLMs, concluo que estas abordagens podem ser aplicadas no contexto educacional de várias formas.

Os grafos de conhecimento ajudam na organização da informação, podendo ser utilizados para monitorizar o progresso do estudante e ajustar o conteúdo recomendado com base no seu desempenho e no histórico de estudos, representando uma abordagem alternativa ao *clustering*.

A combinação de ambas as abordagens em um LLM pode ser utilizada para desenvolver um sistema inteligente, no qual é fornecido ao estudante um conjunto de perguntas pré-definidas, acompanhadas das respetivas respostas, promovendo uma aprendizagem mais personalizada e eficaz.

2.4.6 Utilização em Chatbots

A utilização de LLMs teve um impacto significativo, inaugurando uma nova era impulsionada por IA (Dam et al. 2024). O surgimento do *ChatGPT* e do *GPT-4* promove uma reavaliação das possibilidades da Inteligência Geral Artificial (AGI) (W. X. Zhao et al. 2024). AGI é a tarefa de construir sistemas computacionais capazes de compreender ou aprender qualquer tarefa intelectual que um ser humano consiga, normalmente considerado uma IA fraca ou estreita, projetada para realizar tarefas como reconhecimento facial ou jogar xadrez (Chaccour et al. 2024).

Mais recentemente, o ChatGPT implementou um mecanismo de utilização de *plugins* externos, que podem ser comparados com os dos LLMs (W. X. Zhao et al. 2024). Este mecanismo tem o potencial de expandir consideravelmente o alcance das capacidades desses modelos (W. X. Zhao et al. 2024).

Além do ChatGPT ser o LLM mais mencionado, o Google revelou o BARD, uma interface de utilizador pública para IA generativa colaborativa, que passou a ser designada recentemente como *Gemini* (Dam et al. 2024; Wikipedia 2023). O *Bing Chat*, introduzido pela *Microsoft* e alimentado pelo motor *GPT-4*, e os modelos *Claude* e *Claude 2*, que se destacam pelas suas capacidades de raciocínio abrangentes e respostas mais seguras, são outros exemplos mais conhecidos (Lozic e Stular 2023; Ortiz 2023). Por outro lado, o *Ernie*, treinado com uma vasta gama de dados, incluindo triliões de páginas da web, bilhões de dados de voz, dados de pesquisa, dados de imagem e um grafo de conhecimento com 550 bilhões de factos, representa uma realização impressionante (Dam et al. 2024; Wikipedia 2023; Yang 2023).

Na área do ensino, os *chatbots* baseados em LLMs possuem um grande potencial em diversos níveis de educação, podendo desempenhar uma variedade de papéis de apoio no ensino escolar, graduação e pós-graduação, aprimorando a experiência de aprendizagem em todos esses níveis educacionais (Dam et al. 2024). Também podem ser aplicados na escrita académica, atuando como assistentes para instrutores, auxiliando na elaboração do currículo ao fornecer esboços estruturados. Podem oferecer exercícios de gramática, discussões interativas, *feedback* instantâneo e muitas outras abordagens inovadoras no contexto educacional (Dam et al. 2024).

Chatwattana, Yangthisarn e Tabubpha 2024 propuseram um sistema de recomendação com um *chatbot* baseado em IA, cujo principal objetivo era desenvolver a capacidade de analisar consultas e oferecer informações valiosas e precisas sobre sugestões educacionais na *College of Industrial Technology*, da *King Mongkut's University of Technology North Bangkok, Tailândia*. Os autores obtiveram resultados positivos, demonstrando que o sistema foi eficaz ao analisar as perguntas dos usuários e fornecer respostas de forma eficiente.

2.4.7 Utilização em Sistemas de Recomendação

De acordo com os autores (Z. Zhao et al. 2024), existem três abordagens principais para desenvolver LLMs em tarefas de recomendação, das quais pré-treino, ajuste fino e utilização de *prompts*.

O pré-treino consiste no treino de LLMs com um vasto conjunto de dados não rotulados e diversificados, permitindo que os modelos adquiram uma compreensão ampla de diversos aspetos linguísticos, incluindo gramática, sintaxe, semântica e até raciocínio baseado em senso comum (Z. Zhao et al. 2024).

O ajuste fino é um passo essencial para a aplicação de LLMs pré-treinados em tarefas específicas (Z. Zhao et al. 2024). Nos sistemas de recomendação, é fundamental que os LLMs compreendam conhecimentos mais aprofundados do domínio (Z. Zhao et al. 2024). Este paradigma envolve treinar o modelo pré-treinado com conjuntos de dados específicos da tarefa, que incluem interações entre utilizadores e itens, bem como informações adicionais sobre os utilizadores e itens (Z. Zhao et al. 2024).

A utilização de *prompts* é o paradigma mais recente para adaptar os LLMs a tarefas específicas, recorrendo a *prompts* personalizados para cada tarefa (Z. Zhao et al. 2024). Esta abordagem é particularmente útil em cenários onde são necessários grandes conjuntos de dados específicos e atualizações de parâmetros dispendiosas (Z. Zhao et al. 2024). Uma investigação recente tem explorado ativamente o uso de *prompts* para melhorar o desempenho dos LLMs em sistemas de recomendação (Z. Zhao et al. 2024). Técnicas avançadas, como o In-Context Learning (ICL) e o Chain-of-Thought (CoT), têm sido cada vez mais analisadas, com o objetivo de criar *prompts* eficazes para diferentes tarefas de recomendação (Z. Zhao et al. 2024). No entanto, variações nos *prompts* podem resultar em mudanças significativas nos resultados do modelo, o que é abordado por meio de *prompt engineering*, que envolve a criação de consultas em linguagem natural para guiar as respostas dos LLMs de forma eficaz (Naveed et al. 2024; Saravia 2022; Webson e Pavlick 2021).

A utilização de LLMs em sistemas de recomendação também pode ser observada através de *chatbots*, que auxiliam o aluno quando este não encontra resposta no conteúdo disponível onde o sistema de recomendação está implementado.

Outra abordagem é a utilização dos *chatbots* para melhorar as recomendações. Y. Gao e demais autores (Y. Gao et al. 2023) propuseram o *Chat-Rec*, que utiliza o *ChatGPT* para interagir com os utilizadores por meio de conversas, refinando os conjuntos candidatos gerados pelo sistema de recomendação para a recomendação de filmes e obtendo assim uma maior precisão nos resultados.

Os LLMs também são úteis na previsão das classificações atribuídas pelos utilizadores aos *itens*, ao analisar as interações e preferências dos mesmos (Kang et al. 2023; Zhiyuli et al. 2023). Nesse contexto, foi apresentado um paradigma denominado P5, que propõe a conversão de diversos formatos de dados de recomendação em sequências de linguagem natural, mapeando as análises de utilizadores e itens em índices (Geng et al. 2022).

No âmbito educacional, essa abordagem pode ser utilizada para recolher informações sobre os tipos de conteúdos educativos que os estudantes preferem, bem como tópicos ou matérias específicas, facilitando, assim, a recomendação personalizada.

Sein Kim e restantes autores Kim et al. 2024 propuseram o sistema *A-LLMRec*, que aproveita diretamente o conhecimento colaborativo contido num sistema de recomendação pré-treinado de última geração. Este sistema permite que qualquer sistema de recomendação que utiliza filtragem colaborativa seja integrado, o que implica que serviços que utilizam os seus próprios modelos de recomendação podem prontamente aproveitar o poder do LLM (Kim et al. 2024). O desenvolvimento envolveu o alinhamento dos *embeddings* de itens de um sistema de recomendação baseado em filtragem colaborativa congelado com as informações textuais associadas, de forma a capturar tanto o conhecimento colaborativo quanto o textual, utilizando um modelo *Sentence-BERT* (SBERT) (Kim et al. 2024). Em seguida, os *embeddings* combinados foram projetados no espaço de *tokens* do LLM, permitindo uma integração fluida no modelo de recomendação (Kim et al. 2024). O sistema demonstrou ser eficaz em diversos cenários, como *cold start* e *warm start*, superando sistemas de recomendação tradicionais com filtragem colaborativa, oferecendo uma abordagem inovadora e as vantagens de não ser necessário ajustar o sistema de recomendação baseado em filtragem colaborativa ou os LLMs, além de reduzir os requisitos computacionais (Kim et al. 2024).

2.5 Custos de um LLM

Existem diferentes modelos, cada um com várias versões para atender a computadores com diferentes capacidades computacionais. No entanto, as versões destinadas a dispositivos com menor poder de processamento são mais limitadas e podem não oferecer todas as funcionalidades desejadas pelos utilizadores.

O Ollama disponibiliza gratuitamente uma ampla variedade de modelos, desde opções para computadores com menor capacidade computacional (ex.: TinyLlama) até modelos que exigem um maior poder de processamento (ex.: Llama3.2) (Meta 2024b; P. Zhang et al. 2024). Os modelos mais leves apresentam limitações, e o grande tamanho dos modelos impõe desafios significativos de inferência, seja para dispositivos com recursos limitados ou com hardware avançado (Meta 2024b; Yuan et al. 2024).

O treino desses modelos envolve um custo elevado, sendo viável realiza-lo apenas uma única vez (Narayanan et al. 2021; Smith et al. 2022). O processo apresenta dois grandes desafios, os parâmetros do modelo já não podem ser armazenados inteiramente na memória principal, mesmo dos GPUs mais avançadas e mesmo que o modelo caiba em um único GPU, o elevado

número de operações computacionais necessárias pode resultar em tempos de treino longos (Narayanan et al. 2021).

2.5.1 Monetário

A definição dos hiperparâmetros, do número de GPUs, da paralelização intra e inter-nó, dos mecanismos de comunicação, do tamanho dos parâmetros do modelo LLM e do número total de tokens para treino torna-se um fator crítico não só para otimizar o tempo de treino do modelo, mas, mais importante ainda, para reduzir o custo monetário associado a esse processo (Bang et al. 2024).

2.5.2 Hardware

Diversos componentes de hardware podem influenciar o desempenho e a execução de LLM. Entre eles, destacam-se o número de núcleos do processador *Central Processing Unit(CPU)*, a capacidade da memória *Random Access Memory(RAM)*, processador Graphics Processing Unit (GPU) e as características do armazenamento, incluindo o tipo de disco *Hard Disk Drive(HDD)* ou *Solid State Drive(SSD)* e o seu espaço disponível (Rodrigues, Vasconcelos e Maciel 2020).

Quando é necessário executar múltiplos modelos, normalmente é atribuída uma GPU por modelo para garantir que as exigências computacionais sejam atendidas sem interferências (Zhu et al. 2024). No entanto, essa abordagem não considera a natureza dinâmica das necessidades de investigação, onde a popularidade e a utilização dos modelos podem variar significativamente (Zhu et al. 2024). Além disso, a demanda computacional tende a aumentar ao longo do tempo, e é importante ter em conta o tempo de resposta do modelo quando a rapidez é um fator critico (Baek, Lee e Seok 2024; Zhu et al. 2024).

Zhu et al. 2024 realizaram um estudo concluindo que a utilização de partição *Multi Instance GPU (MIG)* pode reduzir o número de GPUs necessárias, proporcionando uma alocação mais eficiente dos recursos computacionais e diminuição de custos de energia.

2.6 Fontes de dados de conteúdo educacional

2.7 Plataformas de Ensino em Portugal

Para compreender a estrutura das plataformas e sites de ensino em Portugal e os seus principais conteúdos, foram analisadas algumas das plataformas educacionais mais conhecidas, *Khan Academy*, Escola Virtual, Aula Digital, Escola Mágica e Matemática.pt (Escola Mágica 2024; Escola Virtual 2024; Khan Academy 2024; Leya Aula Digital 2024; Matematica.pt 2024). Estas plataformas foram identificadas através de pesquisas direcionadas a conteúdos educacionais ou por estarem associadas a editoras de manuais escolares, como é o caso da Aula Digital, pertencente à editora *Leya* (Leya Aula Digital 2024).

Também foram analisados alguns sites menos populares, como Mentes Brilhantes e O Bichinho do Saber, com o objetivo de comparar os recursos disponíveis em plataformas de maior e menor popularidade, permitindo uma avaliação mais abrangente das soluções existentes(Mentes Brilhantes 2024; O Bichinho do Saber 2024).

Os critérios avaliados alinham-se com os objetivos do projeto em desenvolvimento. Nesse contexto, foram analisados aspetos como a gratuitidade das plataformas, a existência de

recomendações personalizadas, a diversidade de conteúdos abrangendo diferentes ciclos de ensino, bem como a disponibilização de soluções associadas a conteúdos educacionais, como exames e exercícios. Adicionalmente, avaliou-se a presença de funcionalidades de análise de estatísticas de aprendizagem, uma característica destacada particularmente na *Khan Academy* (Khan Academy 2024).

			Ciclos de Ensino Escolar						
	Gratuito	**Recomendações personalizadas**	**1º Ciclo**	**2º e 3ºCiclo**	**Secundário**	**Disciplinas Diversificadas**	**Conteudo diversificado**	**Soluções**	**Estatisticas de Aprendizagem**
Khan Academy	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim
escola virtual	Não	Não verificado	Sim	Sim	Sim	Sim	Não verificado	Não verificado	Não verificado
auladigital	Não	Não verificado	Não	Sim	Sim	Sim	Sim	Sim	Não verificado
Escola Mágica	Não	Não verificado	Sim	Sim	Não	Sim	Sim	Não verificado	Não verificado
matematica.pt	Parcilamente	Não verificado	Sim	Sim	Sim	Não	Sim	Sim	Não Verificado
Mentes Brilhantes	Sim	Não	Sim	Sim	Sim	Não	Não	Sim	Não
O bichinho do saber	Sim	Não	Sim	Sim	10º e 11º	Sim	Sim	Não	Não

Figura 2.11: Observações a diferentes plataformas e sites portugueses

Após a análise dos critérios de cada site/plataforma, verifiquei que apenas o *Khan Academy* atende a todos os critérios, sendo a única das plataformas "maiores"a oferecer acesso totalmente gratuito entre as observadas (Khan Academy 2024). Devido a essa limitação, alguns critérios não puderam ser avaliados em determinadas plataformas ou sites, sendo assinalados como "Não verificado"na tabela.

Todos os sites e plataformas analisados apresentam conteúdo educacional abrangendo mais de um ciclo de ensino e diversas áreas, com exceção do matematica.pt e Mentes Brilhantes, que se concentram exclusivamente na matemática (Matematica.pt 2024; Mentes Brilhantes 2024). Além disso, o Mentes Brilhantes carece de diversidade no conteúdo, oferecendo apenas fichas de treino (Mentes Brilhantes 2024).

Dentro dos critérios possíveis de analisar, concluí que o *Khan Academy* é a melhor opção, fornecendo adicionalmente funcionalidade de criação de turmas de estudo, adição de estudantes e turmas através de um código (Khan Academy 2024).

Para utilizadores que têm menos possibilidades financeiras, os sites menos populares analisados apresentam algumas vantagens, embora limitados. Estes tendem a focar-se numa única disciplina, com pouco conteúdo diversificado ou mesmo ausência de soluções, como é o caso do O Bichinho do Saber, mas torna-se uma solução alternativa (O Bichinho do Saber 2024).

2.8 Sumário

A revisão do estado da arte apresentada respondeu às questões de investigação formuladas. Esta análise abrangeu os aspetos fundamentais dos sistemas de recomendação, do sistema educativo português, das plataformas educacionais e das tecnologias emergentes como os LLMs.

Existem diversas técnicas de recomendação, sendo a filtragem colaborativa uma das mais comuns, utilizada para identificar utilizadores com interesses semelhantes. A recomendação baseada em conteúdo foca-se nos itens que o utilizador demonstrou interesse (Umair Javed et al. 2021), enquanto a abordagem baseada em conhecimento fornece explicações sobre as recomendações feitas (Uta et al. 2024).

Cada uma destes técnicas apresenta vantagens e desvantagens. Contudo, estas podem ser mitigadas através da implementação de sistemas híbridos, que combinam as qualidades de cada abordagem para maximizar a sua eficácia, entre outras técnicas dependendo do problema enfrentado.

Foram analisados vários sistemas de recomendação no contexto educacional, bem como uma arquitetura sugerida no contexto de *e-learning* e outros meios de aprendizagem. No entanto, alguns dos sistemas analisados não mencionaram a realização de testes sobre os sistemas desenvolvidos (Bouihi e Bahaj 2019; Mbaye 2018; Obeid et al. 2018; Shishehchi et al. 2012).

Durante a análise destacou-se a abordagem de Solmaz Abdi e demais autores (Abdi, Khosravi, Sadiq e Gasevic 2020) que utilizaram gráficos para ilustrar a progressão dos alunos nos seus estudos. No entanto, alguns alunos consideraram difícil a navegação nesses gráficos. Uma desvantagem dessa abordagem é que, caso os alunos regridam na aprendizagem e sintam que não estão a conseguir acompanhar os colegas nos estudos, podem sentir-se inseguros e ansiosos com o progresso de aprendizagem.

Vários autores utilizaram ontologias para modelar o conhecimento do domínio e o conhecimento dos alunos (Mbaye 2018; Obeid et al. 2018; Rabahallah, Mahdaoui e Azouaou 2018).

Outra abordagem eficaz demonstrada foi o uso de *clustering*, particularmente com o algoritmo DBSCAN em conjunto com o algoritmo EC como auxilio (Bagunaid, Chilamkurti e Veeraraghavan 2022).

Além de muitos autores terem utilizado abordagens inovadoras, que se afastam das tradicionais, alguns recorreram a abordagens híbridas que incluíram uma dessas abordagens tradicionais, combinando filtragem colaborativa ou recomendação baseada em conhecimento (Bouihi e Bahaj 2019; Rabahallah, Mahdaoui e Azouaou 2018).

Além destas abordagens, os autores exploraram uma ampla gama de técnicas, apresentando uma grande variedade de algoritmos, métodos e ferramentas utilizados no desenvolvimento de sistemas de recomendação. Havendo mais ainda por explorar analisando mais abordagens utilizadas por outros autores não mencionados neste estado da arte.

Os LLM podem auxiliar os estudantes, oferecendo materiais de estudo personalizados e questões práticas, ao analisar os estilos de aprendizagem, o desempenho e as preferências, criando assim experiências de aprendizagem individualizadas (Dai et al. 2023; W. X. Zhao et al. 2024).

Embora os LLM enfrentem diversos desafios, é possível superá-los principlamente através de RAG, grafos de conhecimento. Os LLM podem ser utilizados para criar grafos de conhecimento, analisando e gerando texto em grande escala (Carta et al. 2023). Existem também índices baseados em grafos de conhecimento, que podem ser empregados em sistemas RAG (Yunfan Gao et al. 2024).

Além disso os LLM são igualmente amplamente utilizados em *chatbots*, marcando o início de uma nova era impulsionada pela IA (Dam et al. 2024). Esses *chatbots* também podem ser aplicados na área de ensino e na escrita académica (Dam et al. 2024).

A utilização de LLM também se estende à previsão de classificações atribuídas pelos utilizadores aos itens e à realização de recomendações em três etapas, pré-treinamento, ajuste fino e uso de *prompts* (Kang et al. 2023; Z. Zhao et al. 2024; Zhiyuli et al. 2023). No entanto também se encontra a utilização de *chatbots* AI como auxilio aos sistemas de recomendação.

A análise das plataformas permitiu identificar algumas das plataformas e sites de ensino já existentes em Portugal. Dentro dos critérios estabelecidos, destacou-se a *Khan Academy*, que se caracteriza por ser gratuita, possuir um sistema de recomendação e oferecer conteúdo diversificado (Khan Academy 2024).

Com isto, pode-se concluir que a combinação das técnicas tradicionais constitui apenas uma base de possibilidades para implementar um sistema de recomendação no contexto educacional. Os LLM, grafos de conhecimento e *chatbots* têm demonstrado ser abordagens complementares eficazes. Após a análise das plataformas existentes e do funcionamento do ensino em Portugal, é possível afirmar que estas tecnologias oferecem soluções inovadoras para a personalização da aprendizagem e o aprimoramento da experiência educativa, havendo ainda diversas possibilidades de implementação, especialmente quando comparadas com as outras plataformas e sites de ensino em Portugal.

Capítulo 3

Análise e Desenho do Sistema

O processo de Análise e Desenho do Sistema da plataforma educacional é fundamental para definir o âmbito do projeto, identificar os principais atores envolvidos e os papéis que estes desempenham nos diferentes casos de uso. Para além disso, são definidos os requisitos e elaborados diversos diagramas e modelos que descrevem o funcionamento da plataforma, o seu comportamento face às diferentes possibilidades de interação, bem como a lógica de desenvolvimento. Este processo assegura que a solução proposta cumpre os objetivos principais do projeto.

3.1 Âmbito do Projeto

A plataforma O Lama Sabe tem em consideração os diferentes estilos de aprendizagem dos alunos. Para tal, disponibiliza conteúdos e materiais educacionais diversificados e adaptados às suas necessidades. Ao oferecer recomendações personalizadas, o sistema reduz o tempo que o aluno gastaria a procurar materiais que se adaptam ás suas preferências, orientando-o diretamente para os tópicos de estudo relevantes e áreas de dificuldade identificadas.

A integração de um assistente virtual nomeado de Tuga, auxilia os alunos em situações onde não consigam encontrar respostas para as suas dúvidas ou problemas e outros apoios necessários, promovendo um suporte contínuo e acessível.

Um *design* intuitivo permite aos utilizadores consultar facilmente os conteúdos e compreender de forma natural como utilizar a plataforma e as diversas funcionalidades. Além disso, uma interface visualmente apelativa contribui para aumentar o interesse e a motivação dos utilizadores em explorar a plataforma.

As medidas de segurança garantem a proteção e privacidade dos dados dos utilizadores, promovendo uma utilização segura e reforçando a confiança na utilização do sistema.

Adicionalmente, a plataforma é de uso gratuito, considerando que muitas das principais plataformas educativas portuguesas analisadas no estudo do estado da arte são parcialmente gratuitas ou pagas.

3.2 Atores e Casos de Uso

Os atores assumem um papel fundamental na definição das interações entre os utilizadores e o sistema. Na plataforma O Lama Sabe, o principal ator identificado é o Utilizador, uma pessoa que recorre à plataforma com o objetivo de adquirir conhecimento e tirar partido dos conteúdos e funcionalidades disponibilizados.

O Utilizador assume dois estados, não autenticado, quando acede à plataforma sem efetuar *login*, e autenticado, após fazer *login* com sucesso. O Utilizador não autenticado tem acesso limitado, estando apenas habilitado a realizar um conjunto restrito de casos de uso.

Como atores secundários temos o assistente virtual Tuga e o sistema de recomendação. O Tuga fala com o utilizador, responde às suas perguntas e sugere conteúdos com base nos comandos que este escreve em interação com o sistema de recomendação.

Os casos de uso descrevem as funcionalidades essenciais que a plataforma O Lama Sabe disponibiliza aos utilizadores, detalhando as interações entre os mesmos e o sistema. A seguir, são apresentados os casos de uso identificados para a plataforma, dos quais também representados nos diagramas de casos de uso ilustrados nas figuras 3.1 3.2:

- **UC1 - Ler sobre a plataforma:** Os utilizadores podem consultar informações gerais sobre a plataforma, incluindo a sua finalidade e funcionalidades disponibilizadas.
- **UC2 - Visualizar Conteúdos preferidos dos utilizadores:** Os utilizadores ao ler sobre a plataforma veem os conteúdos que foram mais apreciados pelos outros utilizadores;
- **UC3 - Registo:** O utilizador não autenticado cria uma conta na plataforma;
- **UC4 - Login:** O utilizador não autenticado autentica-se na plataforma para ter acesso aos seus recursos e passa a ser um utilizador autenticado;
- **UC5 - Logout:** O utilizador autenticado termina a sua sessão na plataforma, encerrando o acesso à conta e passa a ser um utilizador não autenticado.
- **UC6 - Visualizar conteúdos educacionais:** O utilizador autenticado vê os conteúdos educacionais disponíveis;
- **UC7 - Filtrar por categoria:** O utilizador autenticado pode filtrar os conteúdos educacionais ao selecionar uma ou mais categorias de interesse.
- **UC8 - Resolver Exercícios:** O utilizador autenticado pode resolver exercícios incluídos nas fichas de exercícios, correspondentes aos conteúdos educacionais.
- **UC9 - Gostar de um conteúdo:** O utilizador autenticado dá "gosto"a um conteúdo educacional que lhe agradou;
- **UC10 - Visualizar conteúdo que gostou:** O utilizador autenticado vê os conteúdos aos quais deu "gosto";
- **UC11 - Retirar o gosto do conteúdo:** O utilizador autenticado retira o "gosto"de um conteúdo educacional que previamente marcou como "gosto";
- **UC12 - Recomendações:** O sistema de recomendação sugere conteúdos ao utilizador autenticado com base nos gostos do utilizador;
- **UC13 - Enviar mensagem ao Tuga:** O utilizador autenticado envia uma mensagem ao Tuga através do *chat*;

- **UC14 - Enviar mensagem ao utilizador:** O Tuga responde à mensagem do utilizador autenticado através do *chat*;
- **UC15 - Pedir ajuda ao Tuga:** Se o utilizador autenticado pedir ajuda, o Tuga fornece uma mensagem com os comandos personalizados disponíveis;
- **UC16 - Pedir recomendações ao Tuga:** O Tuga fornece ao utilizador autenticado recomendações personalizadas do sistema de recomendação;
- **UC17 - Visualizar o histórico de mensagens:** O Utilizador autenticado pode consultar o histórico de mensagens trocadas com o Tuga;

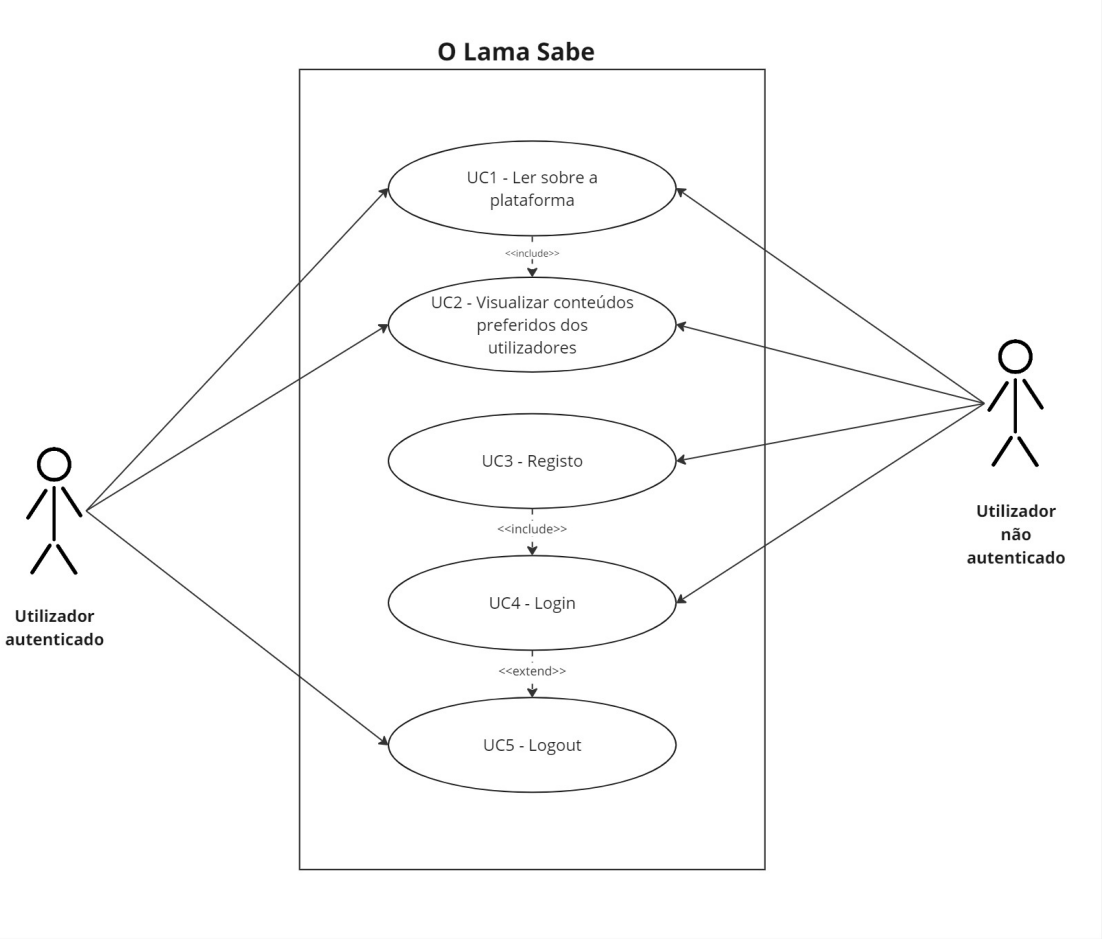

Figura 3.1: Diagrama de Casos de Uso Controlo de Acesso

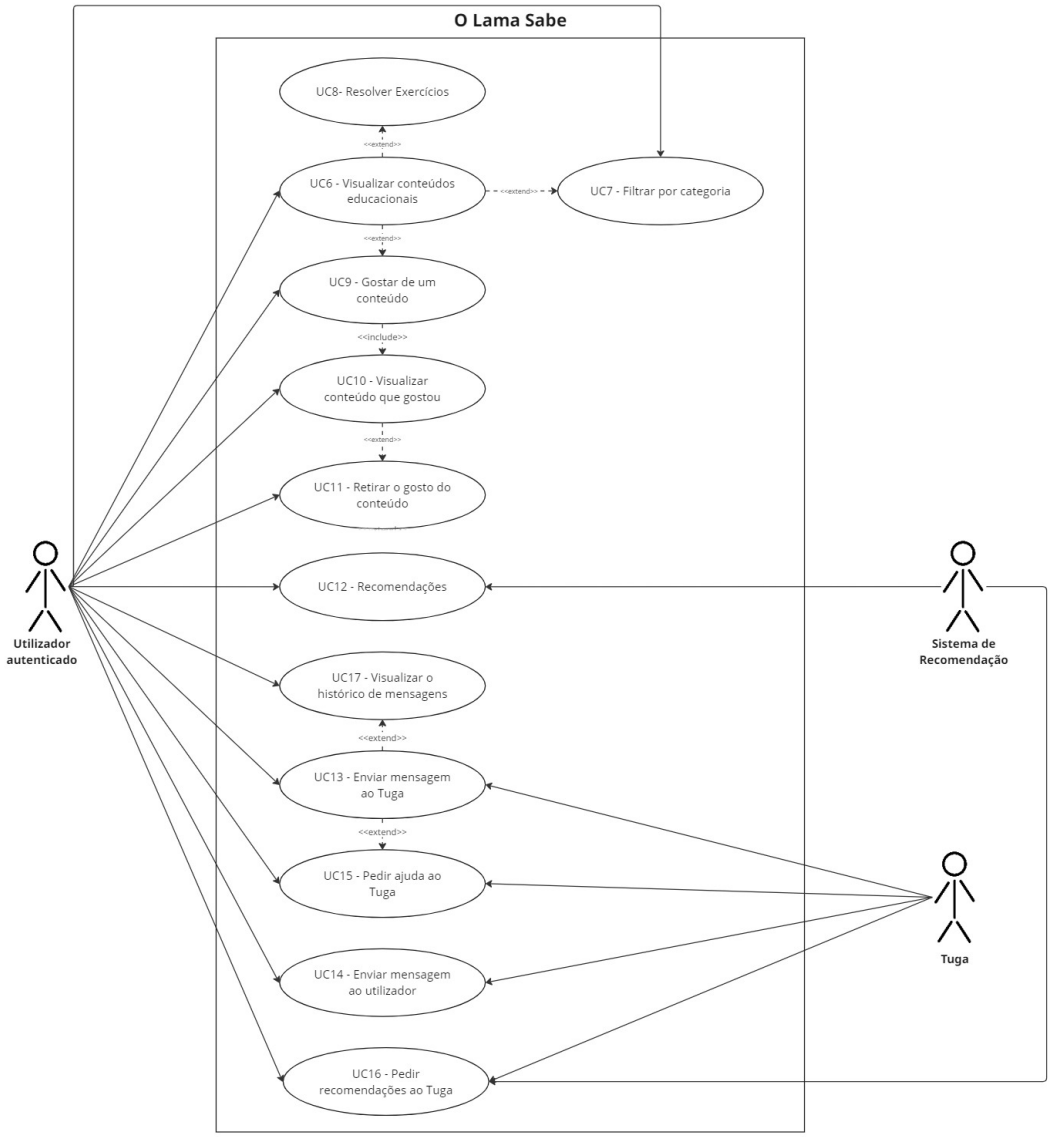

Figura 3.2: Diagrama de Casos de Uso Utilizador Autenticado

3.3 Requisitos

A plataforma foi desenvolvida com base em requisitos previamente definidos que visam responder às necessidades dos utilizadores e *stakeholders*. Cada requisito está documentado numa tabela que inclui um identificador único, título, descrição e nível de prioridade. A implementação dos requisitos foi realizada de acordo com a respetiva prioridade.

3.3.1 Requisitos Funcionais

Os requisitos funcionais definem as diversas funcionalidades que a plataforma deve disponibilizar. Nesta subsecção são apresentados os principais requisitos funcionais do sistema, sendo que os restantes podem ser consultados no Anexo B.

RF1	**Registo do Utilizador**
Descrição	O sistema deve permitir que os utilizadores se registem para obter acesso aos conteúdos educacionais e funcionalidades disponibilizadas pela plataforma.
Prioridade	Alta

Tabela 3.1: RF1 – Registo do Utilizador

RF2	**Autenticação**
Descrição	O sistema deve permitir que o utilizador se autentique na plataforma após concluir o registo com sucesso, introduzindo as credenciais de acesso, *username* e palavra-passe.
Prioridade	Alta

Tabela 3.2: RF2 – Autenticação

RF3	**Terminar Sessão**
Descrição	O sistema deve permitir ao utilizador terminar sessão (*logout*) de forma a impedir acessos não autorizados, especialmente quando o acesso for feito a partir de um dispositivo que não seja de uso pessoal.
Prioridade	Alta

Tabela 3.3: RF3 – Terminar Sessão

RF4	**Utilizadores não Autenticados**
Descrição	Os utilizadores não autenticados apenas têm acesso à página "visão geral"onde contem informação sobre a plataforma, o assistente virtual e os conteúdos preferidos dos utilizadores, no caso de o utilizador tentar aceder aos outros conteúdos é redirecionado para a página de autenticação
Prioridade	Alta

Tabela 3.4: RF4 – Utilizadores não Autenticados

RF5	**Conteúdo Educacional**
Descrição	O sistema deve exibir os diferentes conteúdos educacionais por tipo de conteúdo, nomeadamente, vídeos, livros, documentos, exercícios.
Prioridade	Alta

Tabela 3.5: RF5 – Conteúdo Educacional

RF6	**Resolução de Exercícios**
Descrição	O sistema deve disponibilizar ao utilizador exercícios na própria plataforma, das fichas de exercícios disponibilizadas, com o objetivo de reforçar a aprendizagem e auxiliar na compreensão dos erros cometidos durante o processo de resolução.
Prioridade	Média

Tabela 3.6: RF6 – Resolução de Exercícios

RF7	**Recomendação de Conteúdo**
Descrição	O sistema deve gerar recomendações personalizadas com base nos conteúdos que o utilizador já gostou.
Prioridade	Alta

Tabela 3.7: RF7 – Recomendação de Conteúdo

RF8	**Comunicar com o Tuga**
Descrição	O sistema deve permitir que o utilizador envie mensagens ao assistente virtual Tuga através de um *chat* de mensagens, utilizando um botão de envio ou tecla *enter* para enviar as mensagens.
Prioridade	Alta

Tabela 3.8: RF8 – Comunicar com o Tuga

RF9	**Comandos de Recomendações Personalizadas**
Descrição	O assistente virtual Tuga deve disponibilizar comandos personalizados que permitam ao utilizador solicitar recomendações de conteúdo, apresentando até 10 conteúdos recomendados. O utilizador pode optar por visualizar todos os conteúdos ou por tipo específico, como por exemplo "recomenda vídeos".
Prioridade	Alta

Tabela 3.9: RF9 – Comandos de Recomendações Personalizadas

RF10	**Histórico de Mensagens**
Descrição	O sistema deve permitir ao utilizador visualizar o histórico de mensagens trocadas com o assistente virtual Tuga, incluindo a data e hora de cada mensagem.
Prioridade	Média

Tabela 3.10: RF10 – Histórico de Mensagens

3.3.2 Requisitos Não funcionais

Os requisitos não funcionais descrevem as características de qualidade da plataforma com que as funcionalidades devem ser implementadas.

RNF1	**Desempenho**
Descrição	O sistema deve manter um desempenho consistente e eficiente, respondendo em curtos períodos de tempo.
Prioridade	Alta

Tabela 3.11: RNF1 – Desempenho

RNF1.1	**Tempo de Carregamento da Página de Recomendações**
Descrição	O sistema deve ser capaz de carregar a página de recomendações em tempo útil, de forma a garantir uma navegação fluida e proporcionar uma experiência de utilização satisfatória.
Prioridade	Alta

Tabela 3.12: RNF1.1 – Tempo de Carregamento da Página de Recomendações

RNF1.2	**Tempo de Resposta do Tuga**
Descrição	O sistema deve assegurar que o assistente virtual Tuga responda ao utilizador e apresente as recomendações em tempo útil, com tempos de resposta reduzidos, garantindo uma experiência fluida ao utilizador.
Prioridade	Alta

Tabela 3.13: RNF1.2 – Tempo de Resposta do Tuga

RNF2	**Envio de Dados para o Tuga**
Descrição	O sistema deve realizar a indexação dos documentos e da informação utilizada pelo assistente virtual Tuga no momento da execução da aplicação, apenas caso ainda não tenham sido previamente carregados.
Prioridade	Alta

Tabela 3.14: RNF2 – Envio de Dados para o Tuga

RNF3	**Confiabilidade**
Descrição	O sistema deve assegurar o funcionamento contínuo, mesmo na presença de falhas ou erros.
Prioridade	Alta

Tabela 3.15: RNF3 – Confiabilidade

RNF3.1	**Tolerância a Falhas**
Descrição	O sistema deve funcionar corretamente, mesmo que o assistente virtual Tuga ou outro recurso da plataforma falhem.
Prioridade	Alta

Tabela 3.16: RNF3.1 – Tolerância a Falhas

RNF3.2	**Armazenamento dos Dados**
Descrição	O sistema deve garantir o armazenamento consistente dos dados do utilizador, incluindo histórico de mensagens.
Prioridade	Alta

Tabela 3.17: RNF3.2 – Armazenamento dos Dados

RNF4	**Escalabilidade**
Descrição	O sistema deve ser capaz de suportar o aumento da quantidade de utilizadores e volume de dados, sem comprometer o desempenho da plataforma.
Prioridade	Alta

Tabela 3.18: RNF4 – Escalabilidade

3.3.3 Requisitos de Segurança

Os requisitos de segurança asseguram a proteção dos dados dos utilizadores e a integridade do sistema.

RS1	**Armazenamento dos Dados Sensíveis dos Utilizadores**
Descrição	Os dados sensíveis dos utilizadores devem ser armazenadas de forma segura, utilizando técnicas de *hashing*, para garantir a proteção contra ataques.
Prioridade	Alta

Tabela 3.19: RS1 – Armazenamento dos Dados Sensíveis dos Utilizadores

RS2	**Controlo de Acesso à Plataforma**
Descrição	O sistema deve incluir um controlador de autenticação que assegure o acesso exclusivo a utilizadores autenticados às páginas autorizadas. Caso utilizadores não autenticados tentem aceder a essas páginas, deverão ser redirecionados para a página de *login*.
Prioridade	Alta

Tabela 3.20: RS2 – Controlo de Acesso à Plataforma

RS3	**Segurança das Sessões**
Descrição	O sistema deve criar e gerir as sessões dos utilizadores de forma segura através da utilização de *tokens* de autenticação.
Prioridade	Alta

Tabela 3.21: RS3 – Segurança das Sessões

RS4	**Invalidar Sessões**
Descrição	O sistema deve invalidar os *tokens* das sessões dos utilizadores ao realizar o *logout*.
Prioridade	Alta

Tabela 3.22: RS4 – Invalidar Sessões

RS5	**Encerramento de Conexões à Base de Dados**
Descrição	O sistema deve encerrar as conexões à base de dados após a sua utilização, mesmo em caso de ocorrência de erros, com o objetivo de prevenir fuga de recursos que possam comprometer a segurança.
Prioridade	Alta

Tabela 3.23: RS5 – Encerramento de Conexões à Base de Dados

RS6	**Segurança contra SQL Injection**
Descrição	O sistema deve utilizar consultas parametrizadas, denominadas *placeholders*, nas interações com a base de dados, com o objetivo de prevenir ataques de *SQL Injection*.
Prioridade	Alta

Tabela 3.24: RS6 – Segurança contra *SQL Injection*

RS7	**Conformidade com RGPD**
Descrição	O sistema deve estar em conformidade com o RGPD, assegurando o armazenamento, tratamento e eliminação dos dados pessoais dos utilizadores.
Prioridade	Alta

Tabela 3.25: RS7 – Invalidar Sessões

3.3.4 Requisitos de Interface e Usabilidade

Os requisitos de interface e usabilidade referem-se aos elementos visuais, funcionais e experiencial da interação entre o utilizador e a plataforma. Estes requisitos visam garantir uma interface intuitiva, acessível, eficiente e confortável de utilizar, promovendo uma experiência de uso satisfatória e fluida.

RI1	**Compatibilidade de Navegadores**
Descrição	O sistema deve ser compatível com os diferentes navegadores, por exemplo, *Chrome* e *Firefox*.
Prioridade	Alta

Tabela 3.26: RI1 – Compatibilidade de Navegadores

RI2	**Estabilidade Visual**
Descrição	O sistema deve adaptar-se automaticamente a diferentes tamanhos e resoluções de ecrã, mantendo o aspeto visual e comportamento funcional da interface.
Prioridade	Alta

Tabela 3.27: RI2 – Estabilidade Visual

RI3	**Interface Intuitiva**
Descrição	A plataforma deve proporcionar uma experiência de navegação intuitiva, assegurando acessibilidade e consistência visual entre as diferentes páginas e funcionalidades do sistema.
Prioridade	Alta

Tabela 3.28: RI3– Interface Intuitiva

RI4	**Apelo e Conforto Visual**
Descrição	A plataforma deve ser visualmente atrativa, utilizando uma paleta de cores adequada, de forma a captar a atenção do utilizador e proporcionar conforto na sua utilização.
Prioridade	Alta

Tabela 3.29: RI4 – Apelo e Conforto Visual

RI4	**Dinâmica do Botão de Envio do Tuga**
Descrição	O sistema deve alterar o ícone do botão de envio no *chat* do assistente virtual Tuga de 'enviar' para 'a processar' até o Tuga responder ao utilizador, de forma a informar visualmente o utilizador sobre o estado da interação.
Prioridade	Média

Tabela 3.30: RI4 – Dinâmica do Botão de Envio do Tuga

RI5	Caixas de Texto Ajustáveis
Descrição	As caixas de texto da plataforma devem permitir o seu redimensionamento, possibilitando ao utilizador ajustar o tamanho conforme necessário, tanto ao escrever mensagens para o assistente virtual Tuga, como na resolução de exercícios.
Prioridade	Média

Tabela 3.31: RI5 – Caixas de Texto Ajustáveis

3.3.5 Requisitos de Integração

Os requisitos de integração especificam como o sistema interage com serviços externos, incluindo os utilizados para o desenvolvimento do assistente virtual Tuga.

RIN1	Integração com Base de Dados Relacional
Descrição	O sistema deve integrar-se com uma base de dados relacional para assegurar a gestão, armazenamento e controlo dos dados da aplicação.
Prioridade	Alta

Tabela 3.32: RIN1 – Integração com Base de Dados Relacional

RIN2	Integração com Endpoints
Descrição	O sistema deve integrar *endpoints* HTTP para permitir a comunicação com serviços externos, assegurando a troca segura de dados e a autenticação adequada.
Prioridade	Alta

Tabela 3.33: RIN2 – Integração com *Endpoints*

RIN3	Integração com RAG
Descrição	O sistema deve utilizar um serviço de RAG para consultar informações provenientes de fontes locais e externas, incluindo o carregamento de dados no caso de fontes locais.
Prioridade	Alta

Tabela 3.34: RIN3 – Apelo e Conforto Visual

RIN4	Utilização de LLM na Nuvem
Descrição	O sistema deve integrar um LLM na nuvem, permitindo superar as limitações de hardware local e possibilitando o uso de modelos de maior capacidade.
Prioridade	Alta

Tabela 3.35: RIN4 – Utilização de LLM na Nuvem

3.4 Modelo de Domínio

O modelo de domínio ilustrado na figura 3.3 descreve os principais conceitos da plataforma e serve de base para compreender a lógica do desenvolvimento do sistema.

Este modelo é constituído por várias entidades e relações, sendo o Utilizador a entidade principal.

Um Utilizador pode ter associadas uma ou várias Sessões, correspondentes aos diferentes acessos à plataforma.

Na interação com o assistente virtual Tuga, o utilizador pode não enviar nenhuma mensagem, ou pode enviar várias.

O Utilizador pode não ter Preferência por nenhum Conteúdo, ou pode ter Preferência por vários. Um Conteúdo pode não receber Preferências de nenhum Utilizador, ou pode ser preferido por múltiplos Utilizadores.

O Utilizador pode não receber nenhuma ou múltiplas Recomendações. As Recomendações são geradas com base nas Preferências, se o utilizador não tiver preferência por nenhum Conteúdo, não serão geradas Recomendações mas sim exibido as Preferências dos utilizadores.

Cada Recomendação pode conter zero ou mais Conteúdos recomendados. Num raro cenário em que o utilizador já tenha demonstrado preferência de todos os conteúdos disponíveis, o sistema de recomendação poderá não ter novos Conteúdos para sugerir.

Para efeitos de testes ao sistema, cada Recomendação é avaliada com recurso a métricas de avaliação específicas.

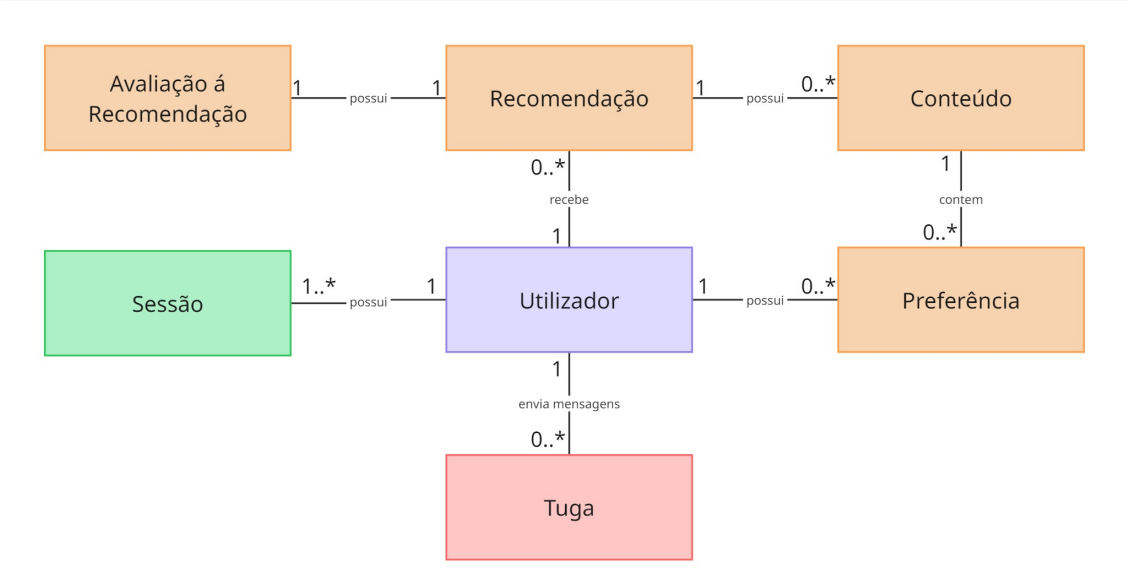

Figura 3.3: Modelo de Domínio da plataforma O Lama Sabe

3.5 Arquitetura do Sistema

A arquitetura do sistema representada na Figura 3.4, segundo o modelo C4 de nível 1, apresenta uma visão geral do sistema como um todo. Neste nível, são identificados os principais intervenientes externos, o utilizador, e o sistema de software, a plataforma educacional O Lama Sabe, bem como as interações entre eles.

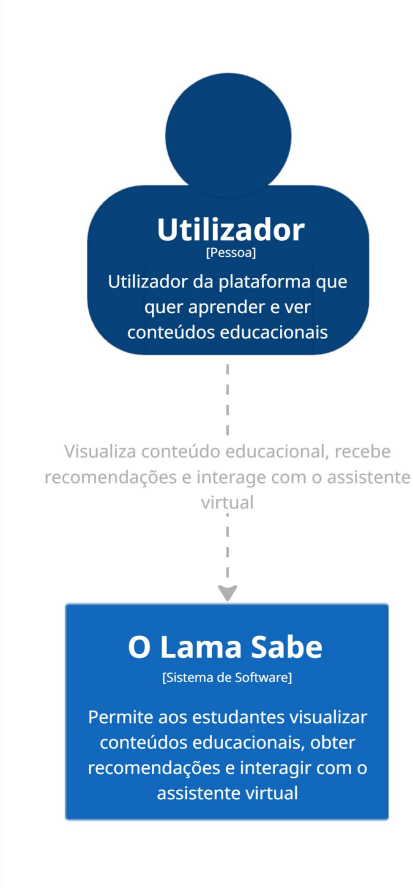

Figura 3.4: Arquitetura do Sistema

3.5.1 Arquitetura de Alto Nível do Sistema

A arquitetura de alto nível do sistema, representada na Figura 3.5 de acordo com o modelo C4 ao nível 2, ilustra os principais *containers* que compõem o sistema, bem como as interações entre eles. Esta representação permite compreender como a sistema está estruturada internamente.

O utilizador acede à plataforma via HTTPS interagindo com o *frontend*, a interface gráfica do sistema. Este *container* é responsável pela apresentação visual e interação com o utilizador, utilizando HTML e CSS para estruturar o conteúdo e definir o estilo visual das das páginas web e JavaScript para tornar as mesmas interativas. A comunicação entre o *frontend* e o *backend* é realizada através de chamadas API REST, com troca de dados em formato JSON. Além disso, é utilizada a tecnologia Jinja2 para renderização dinâmica de páginas HTML no lado do servidor.

O *backend* é responsável pela lógica, gestão de interações e dados. Este comunica com a base de dados através de SQL, onde são armazenadas informações essenciais, tais como dados de utilizadores, conteúdos educacionais, avaliações, recomendações, entre outros.

Adicionalmente, para funcionalidades relacionadas com o assistente virtual, o *backend* integra-se com serviços externos, que disponibiliza o LLM e mecanismos RAG. Estes serviços permitem melhorar a qualidade das respostas fornecidas pelo assistente, bem como personalizar comandos e interações, oferecendo mais funcionalidades na plataforma.

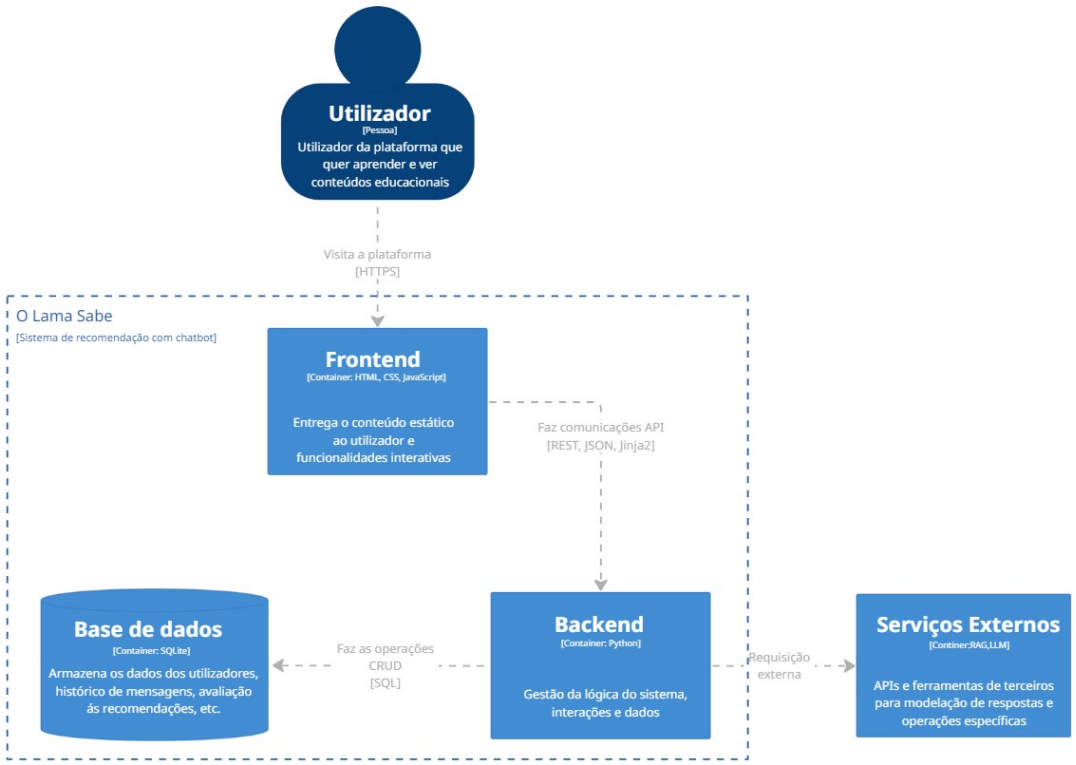

Figura 3.5: Arquitetura de Alto Nível do Sistema

Os diagramas da arquitetura do *Frontend* da Figura 3.6, *Backend* da Figura 3.7 e Serviços Externos da Figura 3.8, apresentados nos subcapítulos seguintes ilustram de acordo com o modelo C4 ao nível 3, os componentes de cada *container* da arquitetura de alto nível do sistema.

3.5.2 Arquitetura do Frontend

Cada componente do *container* do Frontend corresponde a uma página *web* da aplicação. Todas utilizam *HTML, CSS* e *JavaScript*, com exceção das páginas "Termos e Condições"e "Visão Geral", uma vez que são apenas informativas, não utilizam *JavaScript*.

Na página "Registo", o utilizador introduz os dados necessários para criar uma conta. Após o registo, utiliza a página "Login"para se autenticar na aplicação. Utilizadores não autenticados têm acesso apenas à página "Visão Geral".

As páginas "Vídeos", "Livros", "Documentos"e "Fichas de Exercícios"constituem os conteúdos educacionais disponibilizados na plataforma.

A página "Resolver Exercícios"permite ao utilizador resolver diretamente na própria plataforma, os exercícios disponibilizados nas Fichas de Exercícios.

A página "Recomendações"exibe as recomendações personalizadas de conteúdos ao utilizador.

O assistente virtual Tuga dispõe de uma página própria, "Assistente Virtual", bem como de uma página "Histórico", onde o utilizador pode consultar todas as mensagens trocadas com o mesmo.

Todas as páginas comunicam com a API do *backend*, exceto a página "Termos e Condições", que apresenta apenas informação estática.

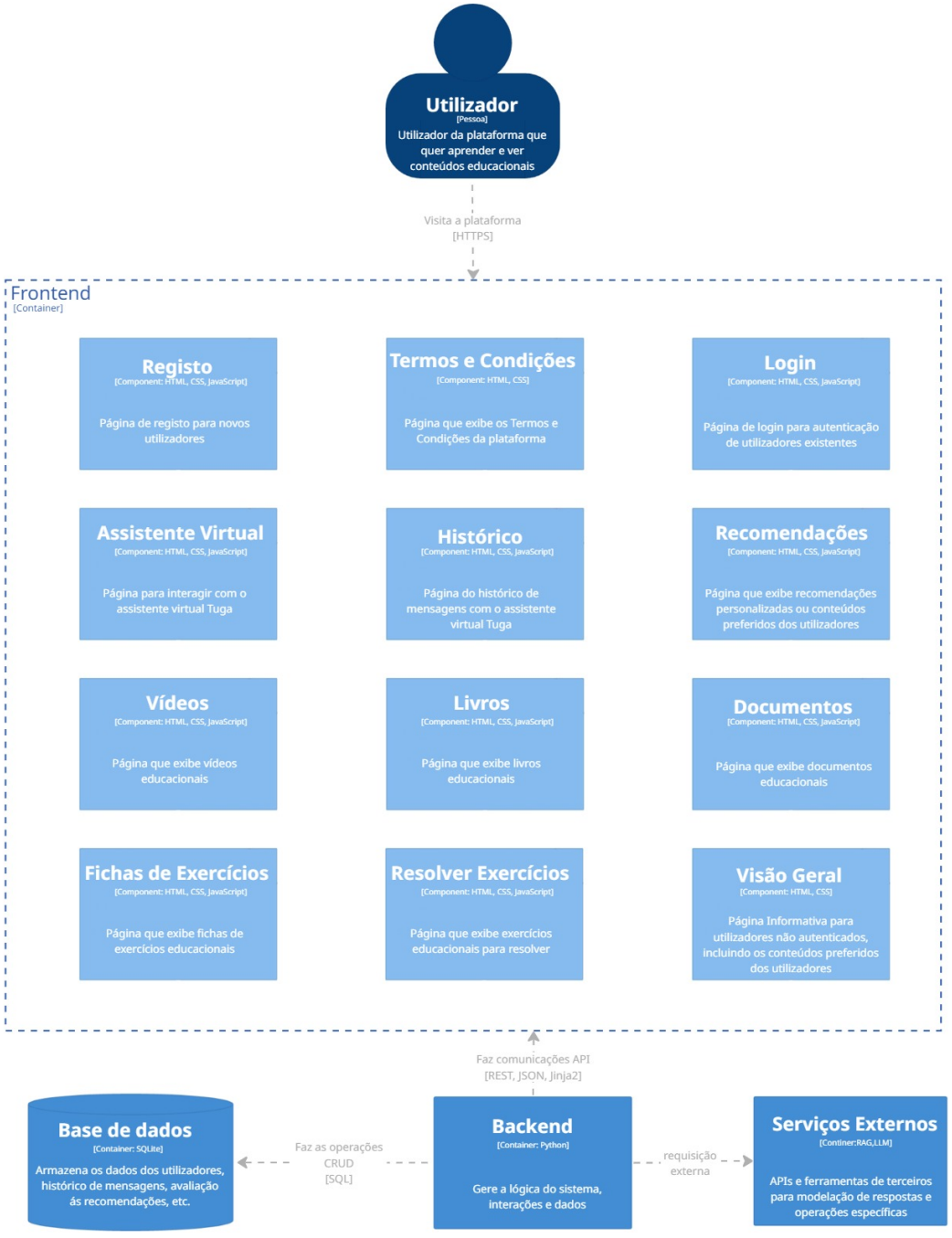

Figura 3.6: Arquitetura do *container Frontend*

3.5.3 Arquitetura do Backend

Os componentes do *Backend*, desenvolvidos em linguagem de programação *Python*, correspondem à lógica e gestão das funcionalidades e dados da plataforma.

O componente "Gestão de Autenticações"é responsável pela autenticação dos utilizadores, incluindo o registo, *login* e gestão de *tokens* de sessão. Este componente garante que apenas utilizadores autenticados têm acesso aos conteúdos e funcionalidades da plataforma.

O componente "Gestão de Conteúdos"é responsável pela exibição e interação com os diversos conteúdos disponíveis na plataforma. Permite que o utilizador visualize e interaja com os conteúdos educativos.

O componente "Gestão de Conteúdo Gostado"depende do funcionamento da "Gestão de Conteúdos"e tem como função gerir os conteúdos que o utilizador demonstrou preferência.

O componente "Recomendações"é responsável por criar recomendações de conteúdos personalizados com base nos conteúdos que o utilizador gostou.

O componente "Assistente Virtual"gere a interação entre o utilizador e o Tuga, permitindo a troca de mensagens e execução de comandos personalizados. Este componente utiliza o componente "Integração com Serviços Externos", responsável pela lógica de comunicação com serviços de externos para criar respostas ás perguntas do utilizador e carregamento de dados de fontes locais.

O componente "Validação de Exercícios", em conjunto com o componente "Integração com Serviços Externos", é responsável por analisar as respostas dos utilizadores aos exercícios e determinar se estão corretas ou incorretas.

O componente "Interação com a Base de Dados"centraliza a comunicação entre o *backend* e a base de dados, realizando operações CRUD, assegurando o acesso de dados como histórico de mensagens, autenticações, avaliações e preferências dos utilizadores.

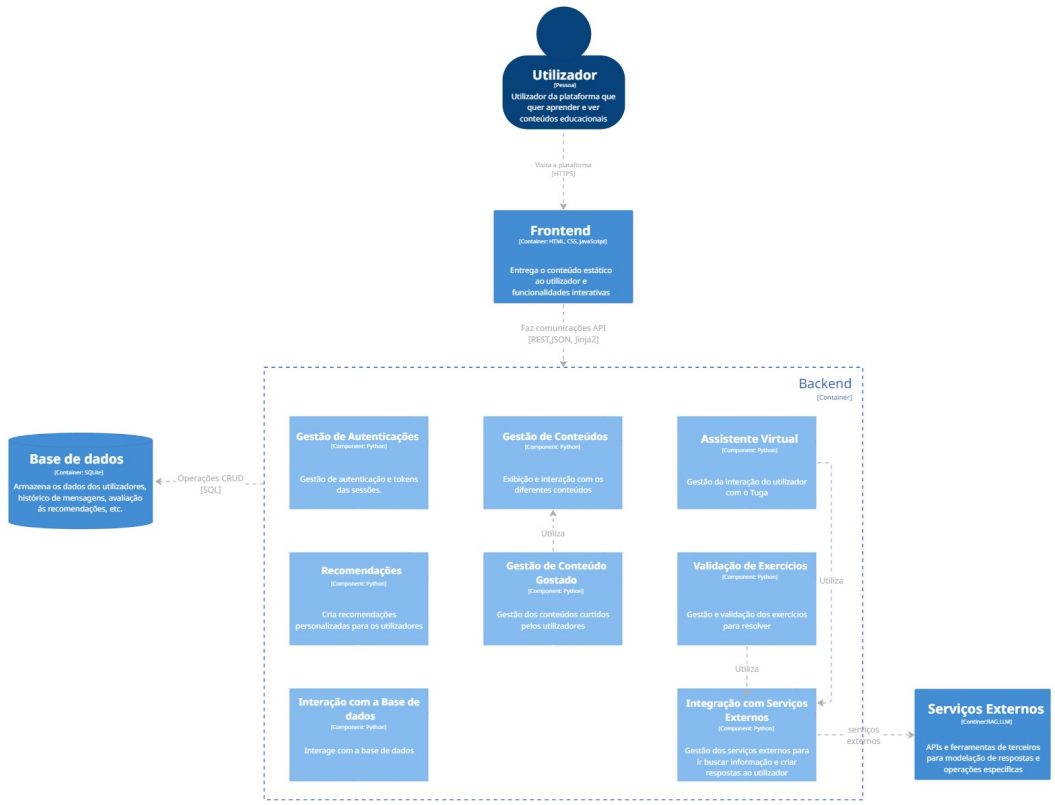

Figura 3.7: Arquitetura *container Backend*

3.5.4 Arquitetura dos Serviços Externos

Cada componente corresponde a um serviço externo a ser integrado no desenvolvimento do Assistente Virtual Tuga.

O *LightRAG* é um sistema RAG utilizado para melhorar as respostas de um modelo LLM através da recuperação semântica de informação relevante (Data Science Lab@HKU 2024). As representações vetoriais dos textos, fundamentais para este processo, são geradas por modelos especializados em *embeddings*, sendo utilizado o *Ollama* localmente para esse fim (Ollama 2023).

Além do *Ollama*, o *LightRAG* também integra o Groq, que disponibiliza modelos LLM de alto desempenho na *cloud* (Data Science Lab@HKU 2024; Groq 2024a; Ollama 2023). Estes modelos são utilizados para enriquecer as respostas do assistente virtual, podendo também ser configurados para identificar entidades e relações, que depois podem ser estruturadas num grafo de conhecimento. Para além disso, o LLM na *cloud* do *Groq* é utilizado no Assistente Virtual para responder aos utilizadores, comandos personalizados e na validação das respostas aos exercícios resolvidos pelos utilizadores(Groq 2024a).

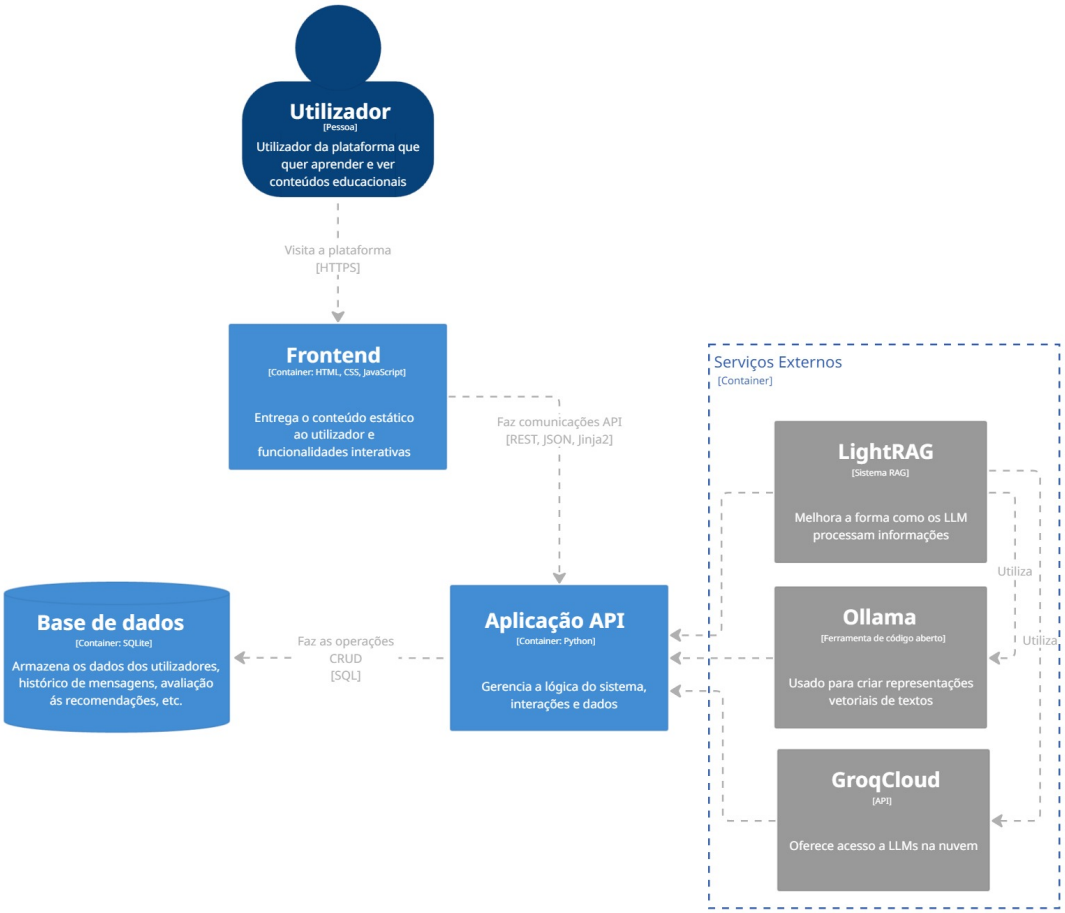

Figura 3.8: Vista Lógica do *container* Serviços Externos

3.5.5 Diagrama Entidade Relação (ER)

O diagrama Entidade-Relação (ER), ilustrado na Figura 3.9, descreve de forma clara e organizada os dados e as relações necessários para a implementação das funcionalidades da plataforma, assegurando o armazenamento do conteúdo essencial e da quantidade mínima necessária de dados sensíveis dos utilizadores.

A tabela principal *"users"* armazena os dados dos utilizadores, sendo o campo *"username"* a chave primária. Um utilizador pode ter zero ou múltiplos registos associados nas tabelas *"user_likes"*, *"chat_messages"* e *"recommendation_metrics"*, sendo o *"username"* utilizado como chave estrangeira nestas tabelas.

A tabela *"recommendation_metrics"* armazena os cálculos das métricas de avaliação que serão utilizados para avaliar o sistema de recomendação, onde a chave primária é o campo *"id"*.

A tabela *"resources"* armazena todas as informações relevantes sobre os conteúdos educacionais. Dependendo do tipo de conteúdo, poderá ser necessário armazenar se for um vídeo, a transcrição em texto no campo *"transcription"* para fornecer mais informações ao modelo LLM, ou se for um livro, o caminho para a imagem da capa no campo *"image_path"*. Um

conteúdo pode receber zero ou múltiplos gostos, sendo a sua chave primária referenciada como chave estrangeira na tabela *"user_likes"*.

A tabela *exercises* armazena as perguntas e respetivas respostas dos exercícios associados às fichas de exercícios. Cada exercício possui uma chave primária denominada *"id"*. Uma ficha de exercícios, contida na tabela *"resources"*, pode conter zero ou múltiplos exercícios. A relação entre ambas as tabelas é estabelecida através da chave estrangeira *"sheet_id"*, presente na tabela *"exercises"*.

A tabela *"user_likes"* regista os conteúdos, *"resources"*, apreciados pelos utilizadores, *"users"* sendo esta informação fundamental para a recomendação personalizada de conteúdos.

Por fim, a tabela *"chat_messages"* armazena o histórico de mensagens trocadas entre o utilizador e o Tuga, utilizando a data e hora atual para facilitar a consulta cronológica das interações, onde a chave primária é o campo *"id"*.

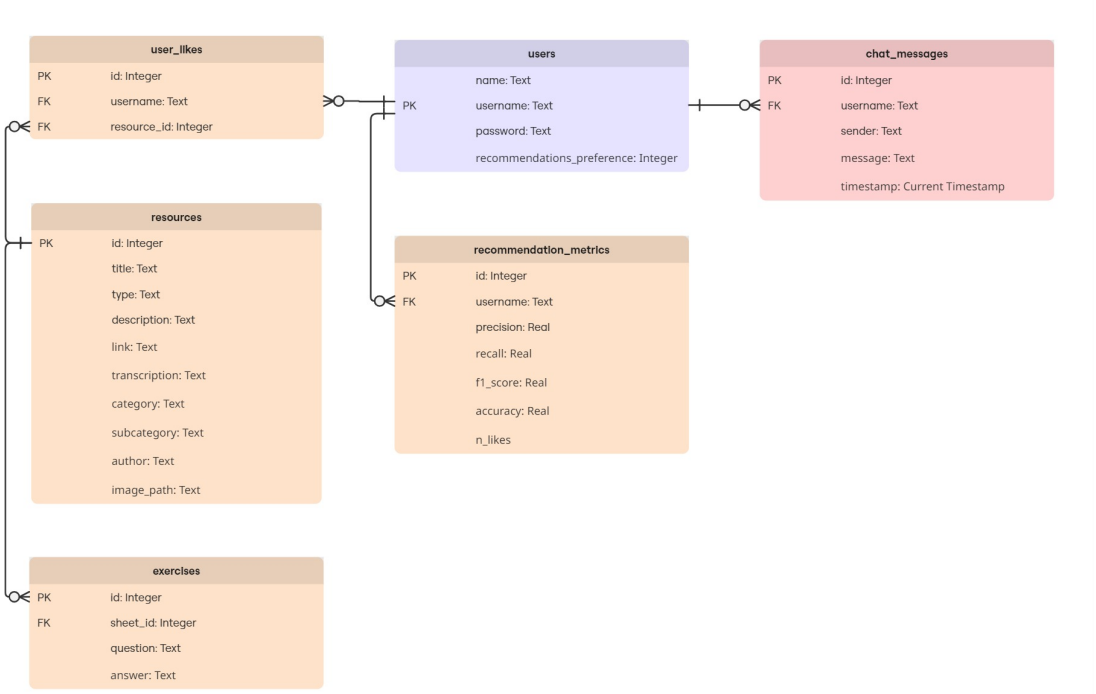

Figura 3.9: Diagrama Entidade Relação (ER)

3.5.6 Processo de Recomendação do Sistema

Os diagramas Business Process Model and Notation (BPMN) representam as etapas que o sistema de recomendação deve seguir na recomendação de conteúdos, tanto a partir da página de recomendações como através dos comandos personalizados do assistente virtual Tuga, incluindo as decisões associadas às diferentes possibilidades, de forma a mitigar problemas de *cold start*.

No fluxo do BPMN da recomendação de conteúdo na página de conteúdos, ilustrado na Figura 3.10, o utilizador inicia o processo ao solicitar recomendações, ou seja, ao aceder à página de recomendações. O sistema de recomendação verifica se no momento do registo na plataforma, o utilizador optou por receber recomendações variadas.

Se for o caso, o sistema de recomendação verifica se o utilizador já gostou de conteúdos. Caso existam gostos registados, o sistema gera 10 recomendações personalizadas com base nas preferências do utilizador, complementadas por 5 conteúdos selecionados aleatoriamente. Se, o utilizador ainda não tiver gostos registados, são apresentados 10 conteúdos aleatórios.

No momento do registo se o utilizador não tiver selecionado a opção de receber recomendações variadas, o sistema verifica também a existência de gostos a conteúdos. Se existirem gostos registados, são geradas 10 recomendações personalizadas. Caso não existam gostos, são apresentados os 10 conteúdos mais apreciados por outros utilizadores da plataforma.

Por fim, os resultados são retornados e apresentados ao utilizador.

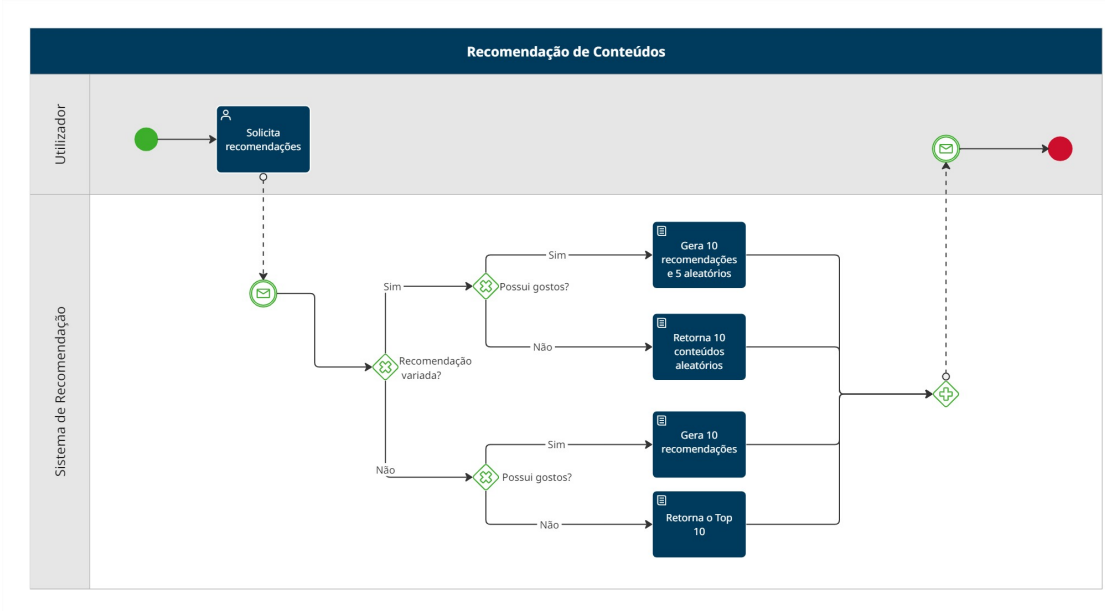

Figura 3.10: Fluxo BPMN de Recomendações Personalizadas na Página de Conteúdos

No fluxo do BPMN da recomendação de conteúdo através de comandos personalizados do assistente virtual Tuga, ilustrado na Figura 3.11, o utilizador inicia o processo enviando uma mensagem ao Tuga com um dos comandos personalizados de recomendação, podendo solicitar recomendações gerais ou recomendações específicas de um tipo de conteúdo, como por exemplo vídeos.

O assistente virtual Tuga encaminha o pedido ao sistema de recomendação, que verifica se o utilizador já gostou de conteúdos. Caso existam gostos, são geradas recomendações personalizadas, caso contrário são retornados os 10 conteúdos mais apreciados por outros utilizadores da plataforma.

No caso de o utilizador utilizar um comando para solicitar recomendações de um determinado tipo de conteúdo, o assistente virtual Tuga filtra as recomendações com base no tipo de conteúdo especificado, apresentando até 10 conteúdos correspondentes. Caso não existam recomendações personalizadas para o tipo solicitado, o sistema apresenta os 10 conteúdos mais apreciados por outros utilizadores desse mesmo tipo de conteúdo.

Por fim, os resultados são retornados e apresentados ao utilizador.

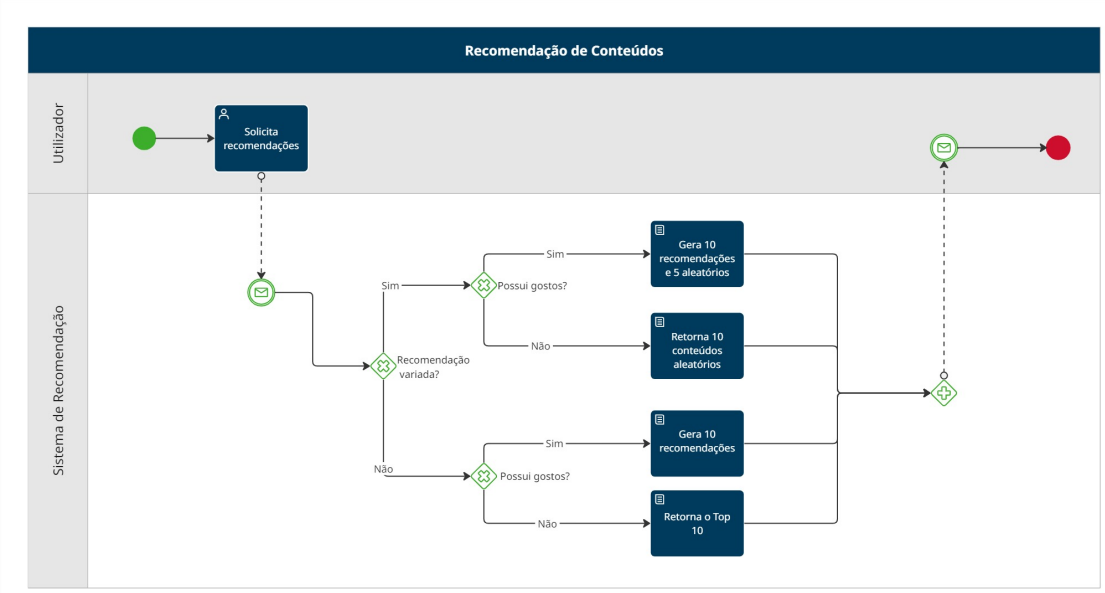

Figura 3.11: Fluxo BPMN de Recomendações Personalizadas Através de Comandos Personalizados do Assistente Virtual Tuga

3.6 Sumário

A plataforma O Lama Sabe é composta por quatro atores principais, o utilizador não autenticado, o utilizador autenticado, o sistema de recomendação e o assistente virtual Tuga, todos fundamentais na análise e desenho do sistema.

Os casos de uso permitiram descrever as principais funcionalidades da plataforma, desde o registo do utilizador até à recomendação personalizada de conteúdo. Os requisitos funcionais, não funcionais, de segurança, de interface e de integração definiram, respetivamente, o que o sistema deve fazer, como deve comportar-se, que medidas de segurança devem ser implementadas, de que forma o utilizador interage com a plataforma e como os serviços externos se integram com o sistema.

O modelo de domínio possibilitou a identificação dos principais conceitos da plataforma e das suas relações, destacando o utilizador como elemento central.

Com o objetivo de fornecer uma visão mais abrangente da lógica necessária à implementação, apresentou-se a visão lógica do modelo C4 ao nível 2. A análise detalhada ao nível 3, por *container*, permitiu compreender em profundidade outros requisitos técnicos de implementação, incluindo linguagens e serviços externos.

O diagrama ER descreveu a estrutura da base de dados, especificando como as entidades estão interligadas e como os dados devem ser organizados e estruturados.

Por fim, os modelos BPMN foram cruciais para definir o comportamento do sistema de recomendação nas diferentes condições, incluindo para mitigar o problema de *cold start* e a integração com o assistente virtual Tuga.

Capítulo 4

Implementação

Este capítulo descreve de forma detalhada a implementação do O Lama Sabe. São abordados todos os aspetos relevantes do seu desenvolvimento, incluindo as funcionalidades implementadas e a estrutura de cada página que compõe a aplicação.

4.1 Produto

A solução O Lama Sabe consiste numa plataforma educacional que disponibiliza conteúdos educativos, integrando um sistema de recomendação que sugere aos alunos conteúdos com base nos seus interesses, ou novos conteúdos para explorar. Adicionalmente, a plataforma O Lama Sabe inclui um assistente virtual, Tuga, que auxilia os alunos em caso de dúvidas ou quando não encontram as informações pretendidas, podendo ainda fornecer recomendações através de comandos personalizados. Estes comandos permitem solicitar recomendações gerais ou filtradas por tipo de conteúdo.

A plataforma O Lama Sabe disponibiliza as suas funcionalidades apenas a utilizadores autenticados, ou seja, que já efetuaram o *login* na plataforma. Utilizadores não autenticados têm acesso restrito à mesma.

4.2 Descrição e Organização

O foco inicial da implementação da plataforma O Lama Sabe centrou-se no desenvolvimento do assistente virtual Tuga, aplicando um conjunto de práticas para garantir respostas eficientes numa área de ensino específica. Posteriormente, foi dada prioridade à recolha de conteúdos educacionais, com o objetivo de suportar o desenvolvimento do sistema de recomendação da plataforma.

O desenvolvimento do *frontend* decorreu ao longo de todo o processo, tendo sido intensificado na fase final, com especial atenção à melhoria do *design*.

A fase de testes à plataforma permitiu validar as funcionalidades, identificando melhorias e à incorporação de novas funcionalidades, todas elas já descritas no presente capítulo.

4.3 Conjunto de Dados de Conteúdos Educacionais

Construiu-se um conjunto de dados educacionais em português, alinhado com os objetivos da solução proposta. Para tal, selecionou-se uma ontologia que teve em consideração o público-alvo dos testes da plataforma O Lama Sabe, focando-se em matérias lecionadas no Curso Profissional de Programação e Sistemas Informáticos.

O conjunto de dados é composto por diversas informações associadas a cada conteúdo educacional,permitindo a expansão futura das funcionalidades, caso necessário. Esses dados foram foram armazenadas e inseridas na base de dados com recurso a *Python* e SQL, sendo as colunas seguintes:

- **id:** Identificador único do conteúdo, garantindo a unicidade mesmo em casos de títulos repetidos com informações distintas;
- **title:** Título do conteúdo;
- **type:** Tipo de conteúdo educacional (ex.: vídeo);
- **description:** Breve descrição sobre o conteúdo;
- **link:** Hiperligação de acesso direto à fonte oficial do conteúdo;
- **transcription:** Transcrição textual do conteúdo, aplicável no caso de vídeos;
- **category:** Categoria do conteúdo;
- **subcategory:** Subcategoria do conteúdo;
- **author:** Autor(es) do conteúdo;
- **image_path:** Caminho da imagem associada ao conteúdo, usado na interface, essencial no caso serem livros.

Foram recolhidos conteúdos como vídeos, documentos, livros e fichas de exercícios provenientes de várias fontes de acesso público. Os livros de livrarias como a *Fnac* e a *Bertrand*, os vídeos do *YouTube*, e os documentos e exercícios maioritariamente obtidos de repositórios públicos de universidades e instituições de ensino superior.

O gráfico ilustrado na Figura 4.1 representa a distribuição dos conteúdos por tipo. No total o conjunto de dados é composto por 85 conteúdos, dos quais 40 são vídeos, 15 livros, 14 fichas de exercícios e 16 documentos. A proporção de vídeos é maior por uma razão específica. Livros e documentos contêm significativamente mais informação do que os vídeos, especialmente quando estes são segmentados em tópicos que permitem uma compreensão mais clara de cada conceito.

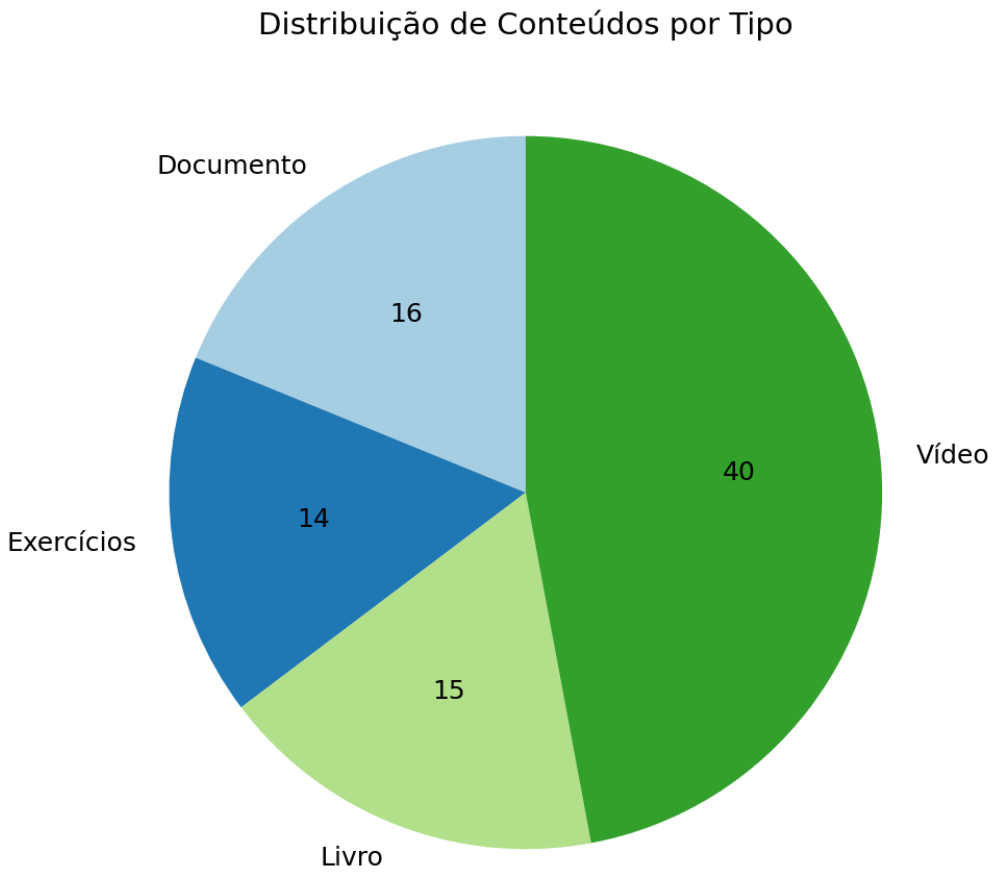

Figura 4.1: Distribuição de Conteúdos por Tipo

A utilização de categorias e subcategorias permite uma melhor organização do conjunto de dados, facilitando o acompanhamento das matérias de estudo já visualizadas durante a pesquisa de conteúdos para inserção no conjunto de dados. A Figura 4.2 ilustra as diferentes categorias e subcategorias associadas a cada matéria de estudo, destacando que em maioria está a Algoritmia e Linguagem de Programação C, as primeiras matérias a ser lecionada no Curso Profissional de Programação de Sistemas Informáticos.

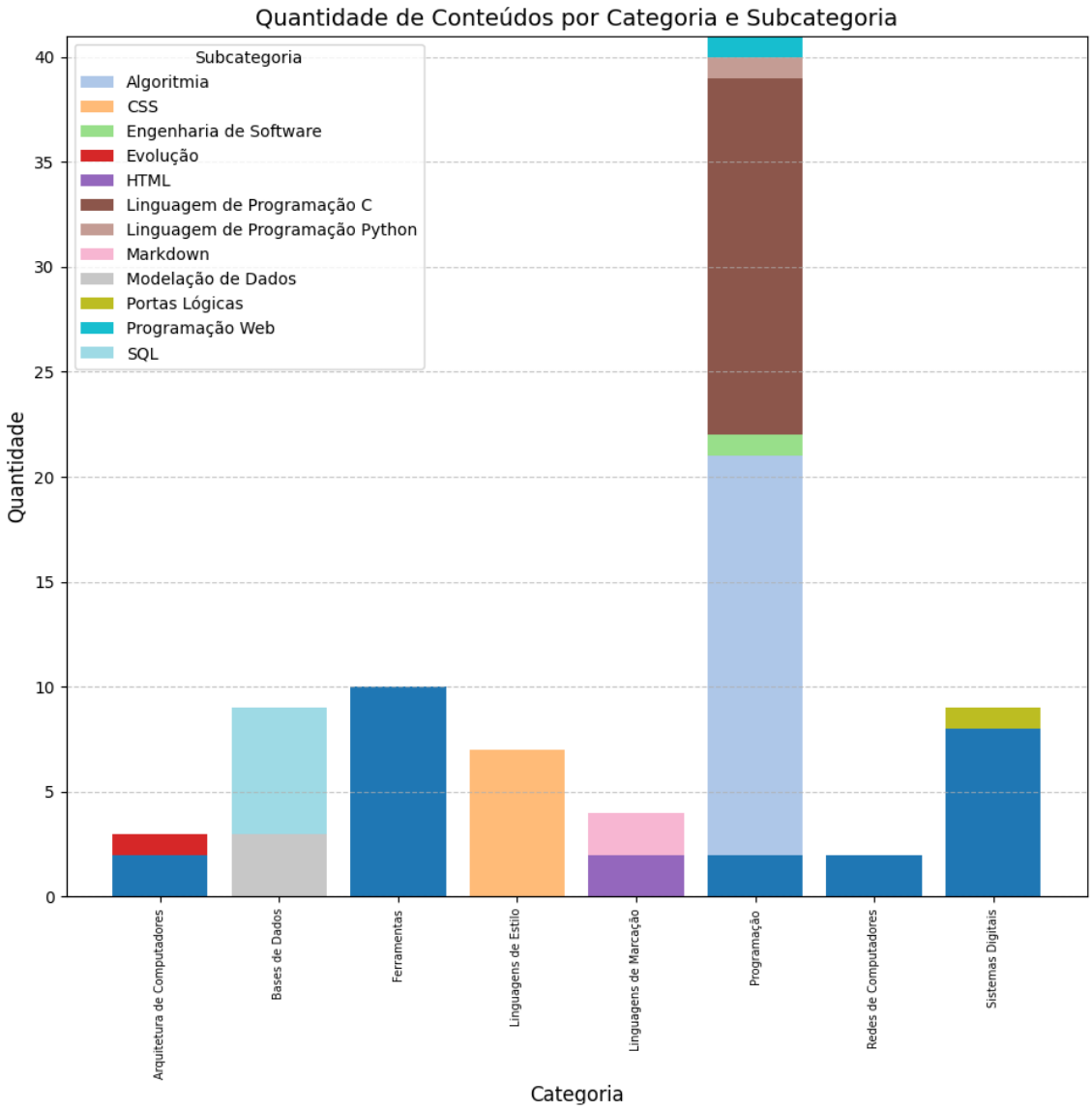

Figura 4.2: Quantidade de Conteúdos por Categoria e Subcategoria

4.4 Controlo de Acesso e Autenticação

A implementação dos mecanismos de controlo de acesso e autenticação revelou-se fundamental para assegurar que apenas utilizadores autenticados pudessem aceder às funcionalidades restritas da plataforma O Lama Sabe. Para tal, foram definidas páginas públicas e privadas, implementado um sistema de verificação do estado de sessão do utilizador, e estabelecidos redirecionamentos automáticos consoante o nível de acesso permitido.

4.4.1 Controlo de Acesso

A implementação de um gestor de controlo de acesso permitiu definir as áreas da plataforma O Lama Sabe às quais os utilizadores podem ou não aceder.

O fluxograma da Figura 4.3 representa o funcionamento do mecanismo de controlo de acesso da plataforma. Após a autenticação com sucesso, os utilizadores são automaticamente redirecionados para a página do Assistente Virtual Tuga, uma vez que este constitui um dos elementos centrais da experiência de utilização da plataforma. Caso o utilizador ainda não esteja autenticado, o sistema verifica se a página solicitada está incluída no conjunto de páginas públicas.

As páginas públicas englobam a "Visão Geral", o "Registo" e o *Login*, sendo estas acessíveis a qualquer utilizador não autenticado. Se for solicitada uma página restrita por um utilizador não autenticado, este é redirecionado para a página de *Login*, onde poderá proceder ao processo de autenticação. A decisão de restringir o acesso às restantes páginas da plataforma deve-se ao facto de a maioria das funcionalidades exigirem o conhecimento da identidade do utilizador autenticado, de forma a permitir o correto armazenamento e tratamento dos dados associados. A disponibilização da página "Visão Geral" permite, no entanto, que utilizadores não autenticados compreendam o propósito da plataforma e as funcionalidades que esta oferece.

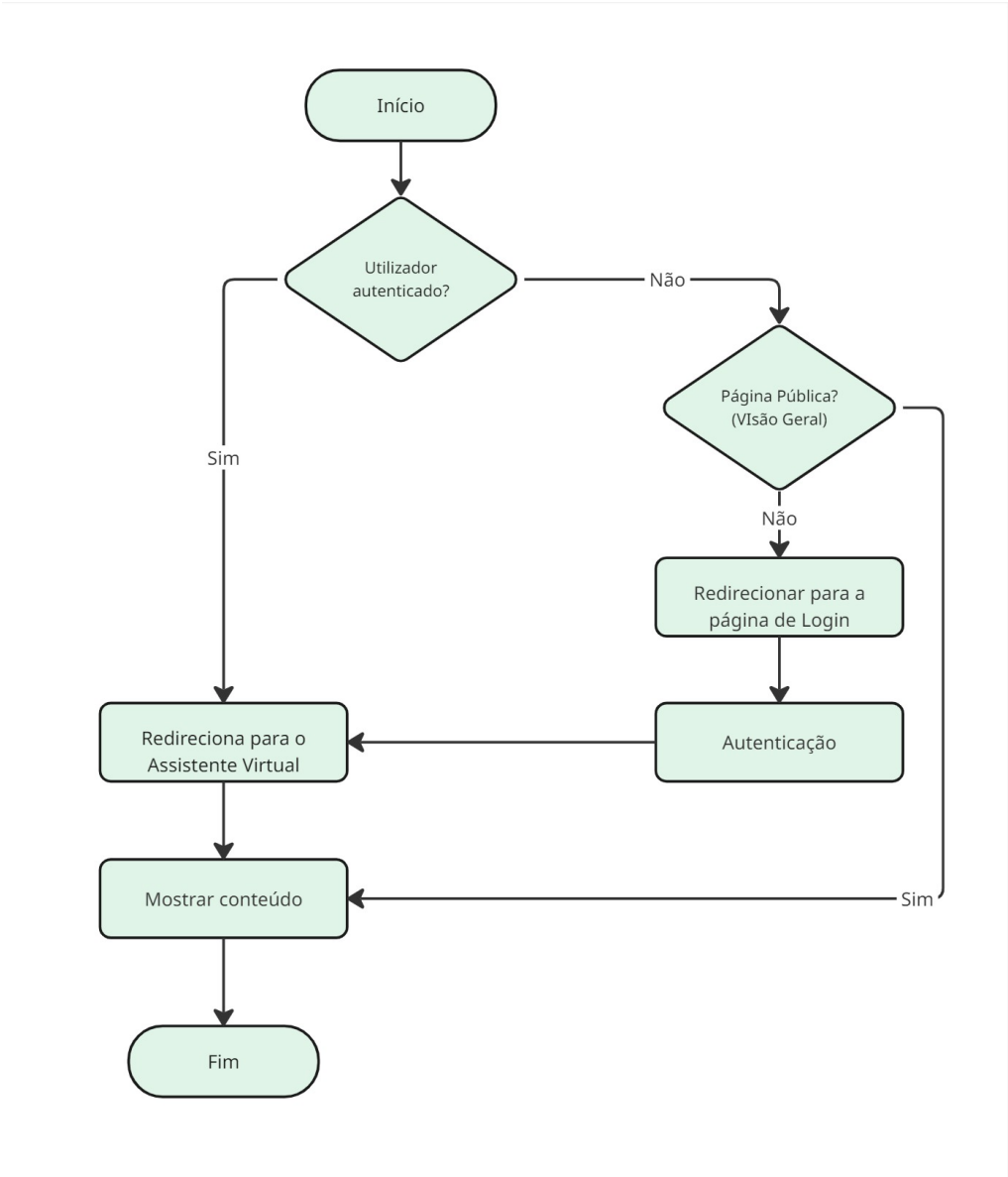

Figura 4.3: Fluxograma do Controlo de Acesso

4.4.2 Autenticação

O fluxograma ilustrado na Figura 4.4 representa o processo de autenticação, desde o registo inicial até ao término de sessão do utilizador.

Caso o utilizador já possua uma conta, deverá aceder à página de *login*, caso contrário deverá aceder à página de Registo. No processo de registo, o utilizador deve preencher os campos obrigatórios. Se os dados introduzidos cumprirem os requisitos definidos, a conta é criada com sucesso e o utilizador é redirecionado para a página de *login*. Em caso de erro, o sistema apresenta uma mensagem informativa com o motivo da falha.

Na página de *login*, o utilizador deve autenticar-se fornecendo as credenciais apropriadas. Se os dados estiverem corretos, o sistema valida a autenticação e o utilizador é encaminhado para a página do Assistente Virtual Tuga. Caso contrário, é apresentada uma mensagem de erro adequada, indicando a falha na autenticação.

Se o utilizador optar por terminar a sessão, o sistema procede à invalidação da de sessão e redireciona-o, de forma segura, para a página de *login*, garantindo a proteção da conta.

Durante o processo de registo, os dados dos utilizadores são armazenados de forma segura. São gerados *hashes* com o algoritmo SHA-256 para o nome e nome de utilizador, e utilizada a biblioteca *bcrypt* para a encriptação das palavras-passe, garantindo a integridade e confidencialidade da informação sensível.

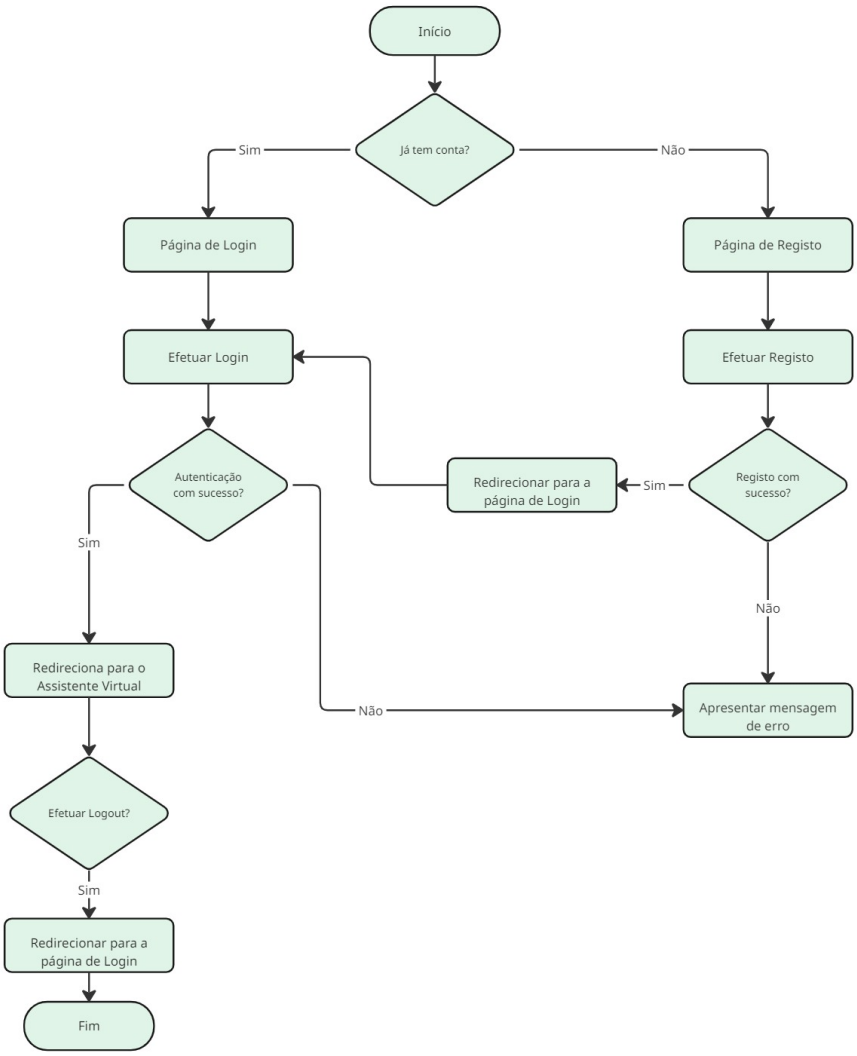

Figura 4.4: Fluxograma do Processo de Autenticação

4.5 Gestão de conteúdo curtido pelos utilizadores

A gestão dos conteúdos educacionais pelos quais o utilizador demonstrou interesse tem como principal objetivo suportar o sistema de recomendação, permitindo sugerir conteúdos relevantes e personalizados ao utilizador.

A utilização da funcionalidade de gostos permite uma interação mais simples e intuitiva para os utilizadores, em comparação com sistemas de avaliação mais complexos, como os de classificação por estrelas (*ratings*).

É possível demonstrar interesse pelo conteúdo educacional visualizado através da funcionalidade de 'gosto', disponível individualmente em cada conteúdo. Ao clicar no botão de gosto, é registada a ação, indicando que o utilizador demonstrou interesse no conteúdo. Caso o utilizador tenha interagido por engano ou tenha deixado de ter interesse no conteúdo em questão, poderá remover o gosto clicando novamente no botão, sendo assim eliminado o respetivo registo.

4.6 Recomendações Personalizadas

Para a recomendação de conteúdos educacionais ao utilizador, utilizou-se a técnica de filtragem baseada em conteúdo, tendo por base os conteúdos que o utilizador demonstrou interesse. A decisão deve-se à sua capacidade de identificar conteúdos com características semelhantes aos que o utilizador já demonstrou interesse. Desta forma, é possível proporcionar uma experiência de aprendizagem mais personalizada e relevante, assegurando recomendações mais precisas com base nos temas de estudo que refletem os interesses do utilizador.

É realizado um conjunto de etapas sequenciais até à apresentação dos conteúdos recomendados ao utilizador. As etapas encontram-se também ilustradas na figura 4.5

- **1. Seleção de dados:** Para garantir que são utilizadas as informações relevantes sobre os conteúdos, são juntados as colunas título (*title*), categoria (*category*) e subcategoria (*subcategory*) de cada conteúdo e concatenadas numa única coluna de texto. De forma a reforçar as palavras-chave presentes no título, que na representam com maior precisão o tema principal do conteúdo, é atribuído um peso cinco vezes superior a esta coluna, relativamente às restantes.
- **2. Limpeza dos títulos:** A remoção de caracteres especiais, *stopwords*, incluindo uma lista personalizada, conversão de todas as letras para minúsculas e eliminação de espaços redundantes nos títulos, permite uniformizar os dados e melhorar a eficácia na vetorização.
- **3. Identificação dos conteúdos apreciados:** Verifica-se quais os conteúdos que o utilizador gostou. Caso não tenha demonstrado interesse em nenhum conteúdo, é retornada uma lista de recomendações vazia.
- **4. Identificação dos títulos relevantes:** Com base nos conteúdos gostados pelo utilizador, verifica-se a correspondência dos respetivos títulos na coluna de títulos transformados, de forma a prevenir inconsistências causadas por erros ou diferenças após o pré-processamento.
- **5. Criação da matriz:** É construída uma matriz TF-IDF a partir dos títulos transformados, permitindo a representação vetorial dos termos presentes nos conteúdos.

- **6. Criação de recomendações:** Calcula-se a similaridade do cosseno entre os vetores dos conteúdos apreciados pelo utilizador e os restantes conteúdos disponíveis, ordenando os resultados por ordem decrescente de similaridade.
- **7. Retornar recomendações:** São selecionadas e retornadas até 20 recomendações, facilitando a gestão da quantidade de conteúdos a apresentar ao utilizador, especificados nos BPMN das figuras 3.10 e 3.11. Esta limitação ajuda também a mitigar o risco de não gerar recomendações relevantes dado a dimensão do conjunto de dados utilizado.

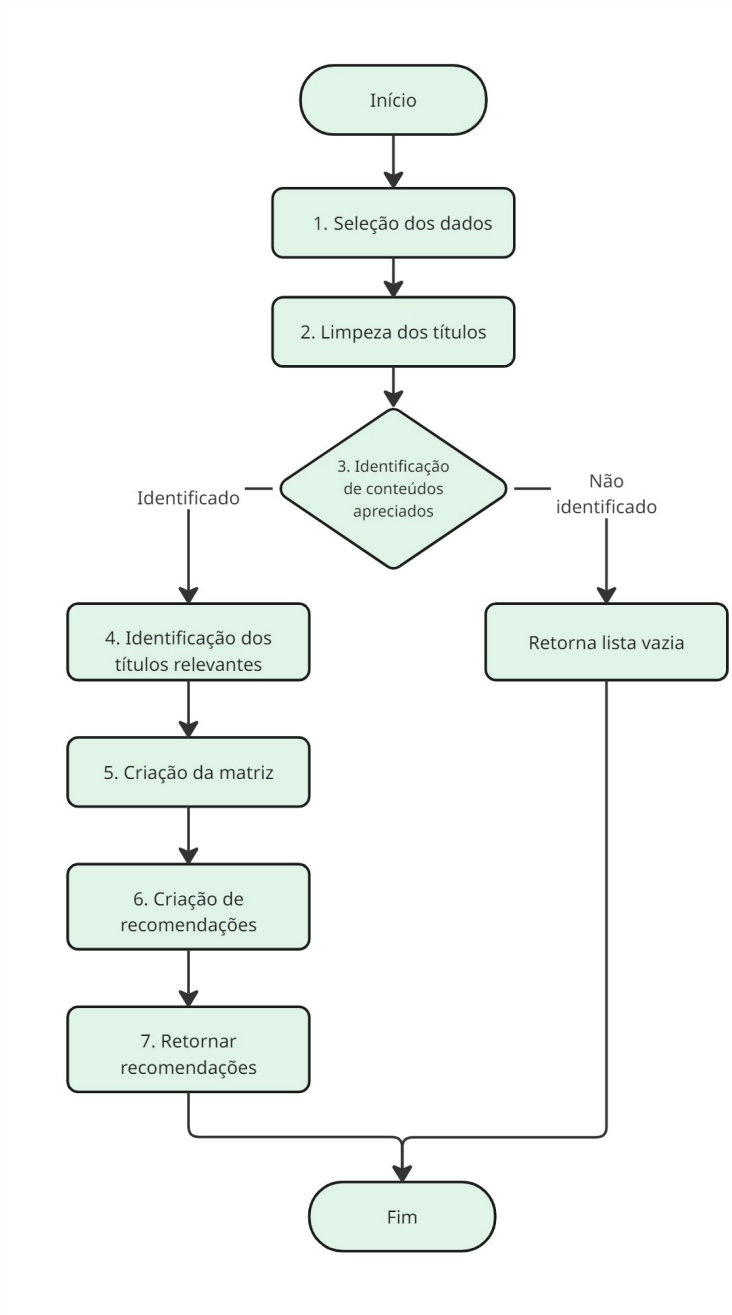

Figura 4.5: Fluxograma das Etapas do Sistema de Recomendação

4.7 Desenvolvimento do Assistente Virtual

O assistente virtual Tuga é um LLM alojado na nuvem, integrado com RAG, permitindo fornecer respostas mais precisas, completas e reduzir alucinações e outras incoerências nas respostas fornecidas ao utilizador.

4.7.1 Configuração do Sistema

Antes de selecionar o modelo ideal alojado na nuvem, foram inicialmente testados modelos locais através da ferramenta *Ollama*, em conjunto com o *LightRAG*, o RAG utilizado no desenvolvimento da plataforma O Lama Sabe. Para além de ser o principal meio de interação com o utilizador, o LLM é também responsável por processar a informação e os documentos fornecidos, extraindo entidades, relações e construir o grafo de conhecimento correspondente.

Análise de Modelos

Foram analisados diversos modelos locais com o objetivo de identificar aquele que melhor respondesse aos requisitos da solução proposta.

Os modelos locais *tinyllama*, *llama3.2:1b* e *qwen2.5-0.5B-Portuguese-v1:fp16* não foram capazes de extrair corretamente entidades e relações, ou de responder com precisão com base na informação fornecida (cnmoro 2024; Meta 2024a; P. Zhang et al. 2023). O modelo que apresentou resultados mais próximos foi o *tinyllama*, embora com erros, como a inserção de caracteres adicionais na resposta (P. Zhang et al. 2023). Estas limitações devem-se ao facto de os LLMs exigirem elevados recursos computacionais, quanto mais limitado for o hardware, mais restrito será os modelos que podem ser utilizados.

A adoção de LLMs na nuvem permitiu ultrapassar as limitações inerentes ao hardware local, tendo sido integrado a API do *Groq*, que possibilitou a utilização de um modelo suficientemente robusto para realizar corretamente a extração de entidades e relações, garantindo assim um desempenho adequado para os objetivos da plataforma O Lama Sabe (Groq 2024a).

Integração dos Serviços Externos

O código apresentado na Listagem 4.1, adaptado de (Data Science Lab@HKU 2024), demonstra como os diferentes serviços externos foram integrados. O modelo LLM é executado na nuvem através do *Groq*, utilizando o modelo *llama-3.3-70b-versatile*, que apresentou um desempenho superior tanto na interação com o utilizador como na extração de entidades e relações, dos disponibilizados na plataforma Groq (Groq 2024b; Meta AI] 2023). Para a representação vetorial, é utilizada uma instância local do *Ollama*, *nomic-embed-text*, responsável pelo processo de *embedding* dos textos (Nomic AI 2024).

```
rag = LightRAG(
    working_dir=WORKING_DIR,
    llm_model_func=groq_model_complete,
    llm_model_name="llama-3.3-70b-versatile",
    embedding_func=EmbeddingFunc(
        embedding_dim=768,
        max_token_size=8192,
        func=lambda texts: ollama_embed(
            texts, embed_model="nomic-embed-text", host="http://
    localhost:11434"
        ),
    ),
)
```

Listing 4.1: LightRAG com Groq e Ollama

As informações utilizadas no *LightRAG* consistem no conjunto de dados, documentos e fichas de exercícios que o integram.

O grafo de conhecimento ilustrado na figura 4.6 representa as entidades, as relações entre elas e os respetivos pesos. A intensidade da cor dos nós reflete a centralidade de proximidade de cada entidade, sendo que valores mais elevados indicam maior relevância na rede. É possível observar também uma secção ampliada do grafo, ilustrado na figura 4.7 que permite visualizar com mais detalhe a estrutura de uma parte das relações e entidades.

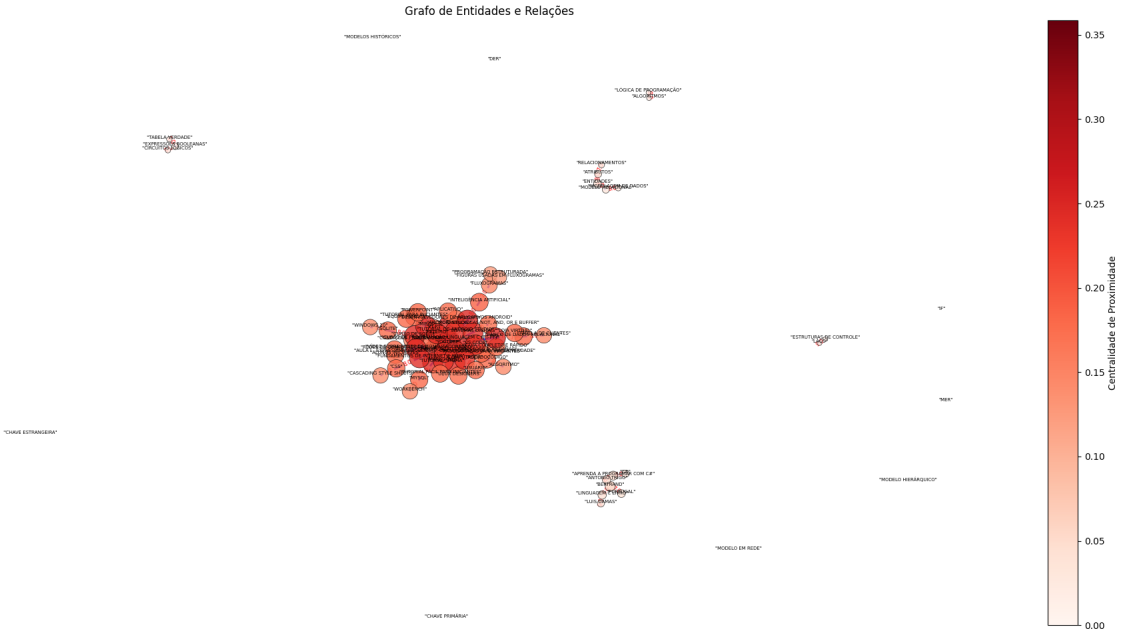

Figura 4.6: Grafo de Conhecimento

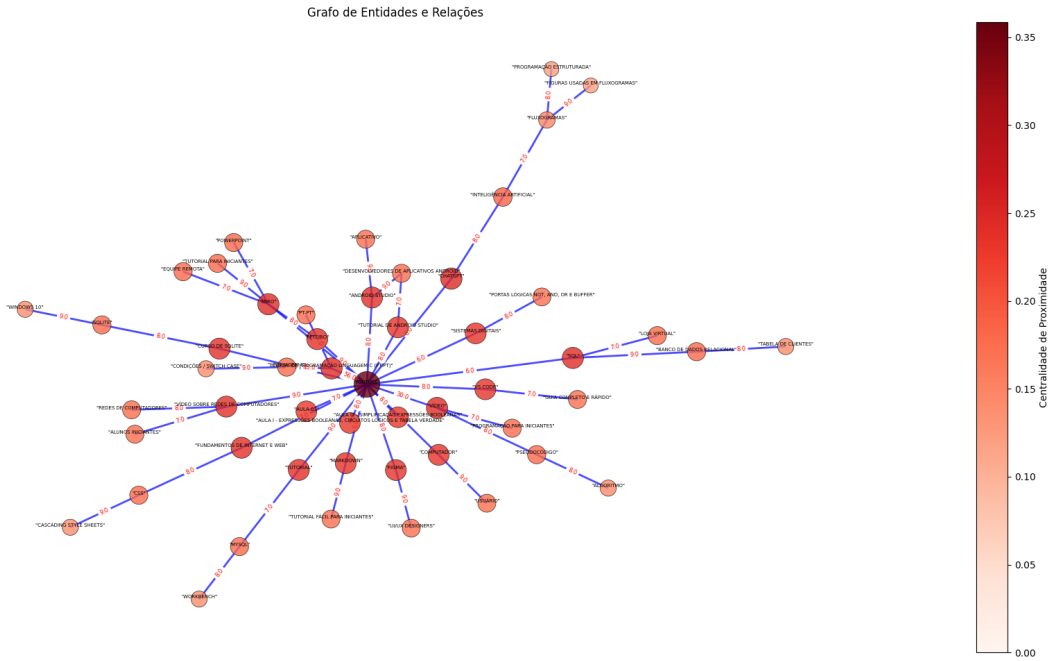

Figura 4.7: Grafo de Conhecimento Ampliado

4.7.2 Comandos de Resposta Personalizados

Os comandos personalizados implementados no assistente virtual Tuga foram concebidos com o objetivo de melhorar a experiência de utilização. Estes comandos permitem ao utilizador interagir de forma personalizada com o sistema de recomendação. Os comandos atualmente disponíveis são:

- **Ajuda-me Tuga:** Lista os comandos personalizados;
- **Recomenda:** Recomenda conteúdo educacional;
- **Recomenda Vídeos:** Recomenda apenas vídeos educativos;
- **Recomenda Livros:** Recomenda apenas livros educativos;
- **Recomenda Documentos:** Recomenda apenas documentos educativos;
- **Recomenda Exercícios:** Recomenda apenas exercícios educativos;

4.7.3 Interação com o Assistente Virtual Tuga

É utilizado um *prompt* configurado de forma a assegurar que o assistente virtual Tuga responda aos utilizadores em português de Portugal e inclua sempre a citação das fontes de onde obteve a informação.

O assistente virtual Tuga, realiza um conjunto de etapas até que o utilizador obtenha uma resposta. Estas etapas encontram-se também ilustradas na Figura 4.5:

- **1. Recebe a mensagem do utilizador:** O Tuga recebe a mensagem enviada pelo utilizador.
- **2. Verifica se é um comando:** O Tuga avalia se a mensagem corresponde a um comando personalizado ou a outro tipo de texto.
- **3. Executa o comando:** Caso a mensagem seja um comando personalizado, o Tuga procede à execução do respetivo comando programado.
- **4. Pesquisa localmente:** Se a mensagem não for um comando personalizado, é realizada uma pesquisa local através do *LightRAG*, para encontrar uma resposta adequada.
- **5. Verifica se encontrou uma resposta:** O Tuga verifica se conseguiu obter uma resposta satisfatória com base na informação local.
- **6. Efetua uma pesquisa híbrida:** Caso não seja encontrada uma resposta adequada localmente, o Tuga efetua uma pesquisa híbrida utilizando o LightRAG. A pesquisa híbrida consiste em combinar fontes locais e externas para gerar uma resposta mais completa e precisa.
- **7. Responde ao utilizador:** O Tuga fornece uma resposta ao utilizador, seja após a execução do comando, seja com base na informação obtida nas pesquisas.

Para determinar o método de pesquisa mais adequado a utilizar com o serviço *LightRAG*, foram realizados testes comparativos entre os diferentes modos disponíveis, pesquisa rápida (*naive*), pesquisa local, pesquisa global(fontes externas) e pesquisa híbrida (pesquisa local e global). A avaliação consistiu em solicitar informações sobre temas previamente incluídos nos dados fornecidos ao *LightRAG*, analisando a qualidade e completude das respostas obtidas. Os resultados demonstraram que a pesquisa local forneceu as respostas mais completas e relevantes, seguida da pesquisa híbrida. Com base nestes resultados, optou-se pela utilização preferencial destes dois métodos no funcionamento do assistente virtual Tuga.

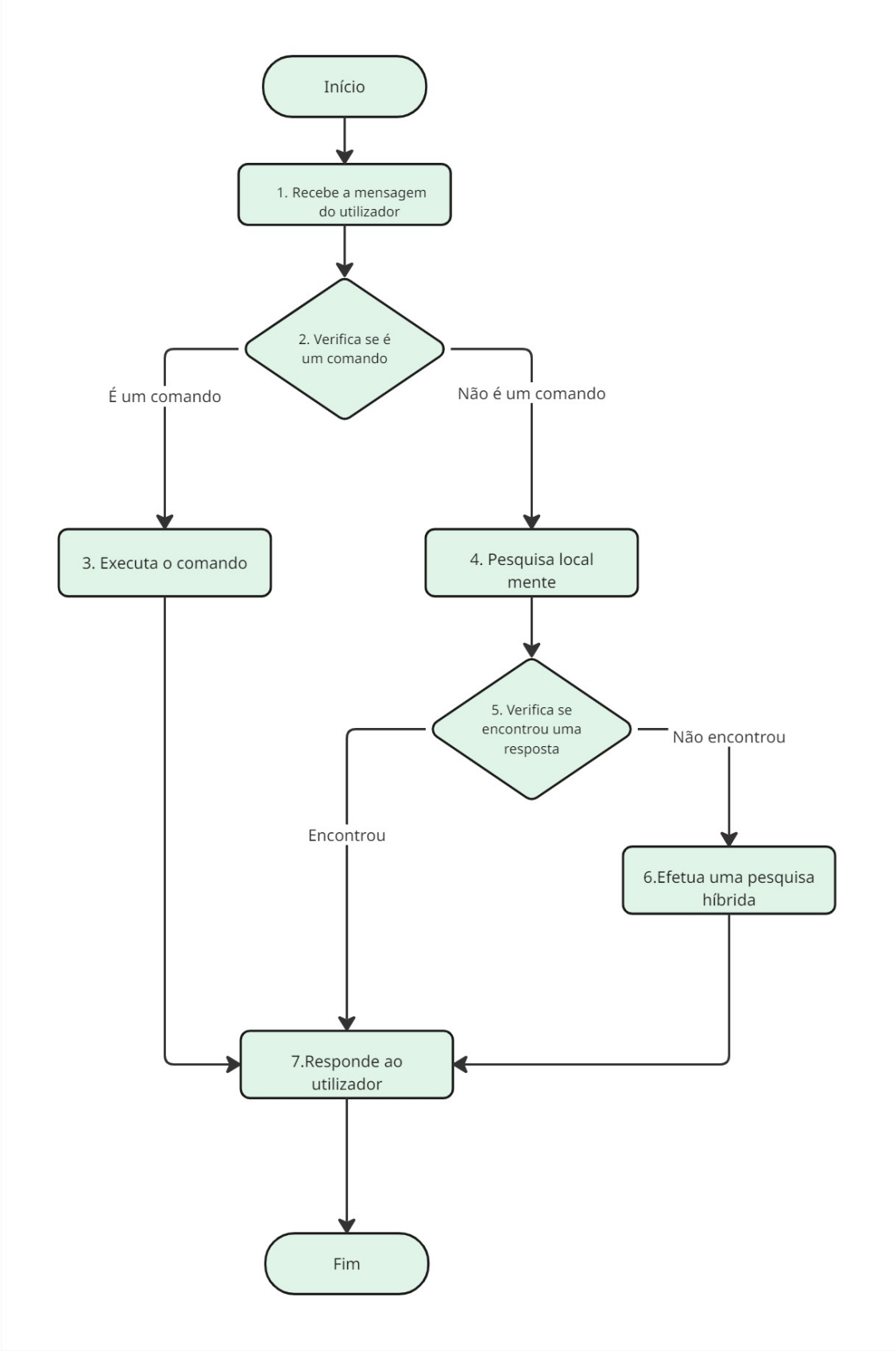

Figura 4.8: Fluxograma de Funcionamento do Tuga

O diagrama de sequência ilustrado na Figura 4.9, que aprofunda o fluxograma previamente apresentado, detalha as interações entre os diversos componentes do sistema durante o processamento de uma mensagem do utilizador até à obtenção da resposta por parte do Tuga. Os intervenientes neste processo são o Utilizador, o *Frontend*, a aplicação *FastAPI* (*backend*), adicionalmente o sistema de recomendação, a base de dados, e os serviços externos *LightRAG* e *Groq*.

O diagrama evidencia o fluxo completo. O utilizador inicia a interação ao escrever uma mensagem, que é apresentada no *frontend*. Esta mensagem é posteriormente enviada via método *POST* para a aplicação *FastAPI*, onde é armazenada na base de dados para possibilitar a consulta do histórico de mensagens. Em simultâneo, a aplicação *FastAPI* processa a mensagem e identifica o seu tipo. Caso se trate de um comando personalizado de recomendação, o sistema de recomendação é invocado para gerar sugestões personalizadas com base nos dados do utilizador. A resposta gerada é então devolvida ao *frontend* para visualização e adicionalmente, armazenada na base de dados.

Se a mensagem não corresponder a um comando, o serviço *LightRAG* é acionado para realizar uma pesquisa local ou híbrida, fornecendo a resposta à aplicação *FastAPI*, que a armazena na base de dados e a apresenta ao utilizador. No cenário em que não seja possível encontrar uma resposta adequada, o sistema recorre ao serviço externo *Groq*, utilizando o modelo *LLM* para gerar uma resposta alternativa, que é igualmente registada na base de dados e apresentada ao utilizador, garantindo assim que este obtém sempre um retorno do assistente virtual Tuga.

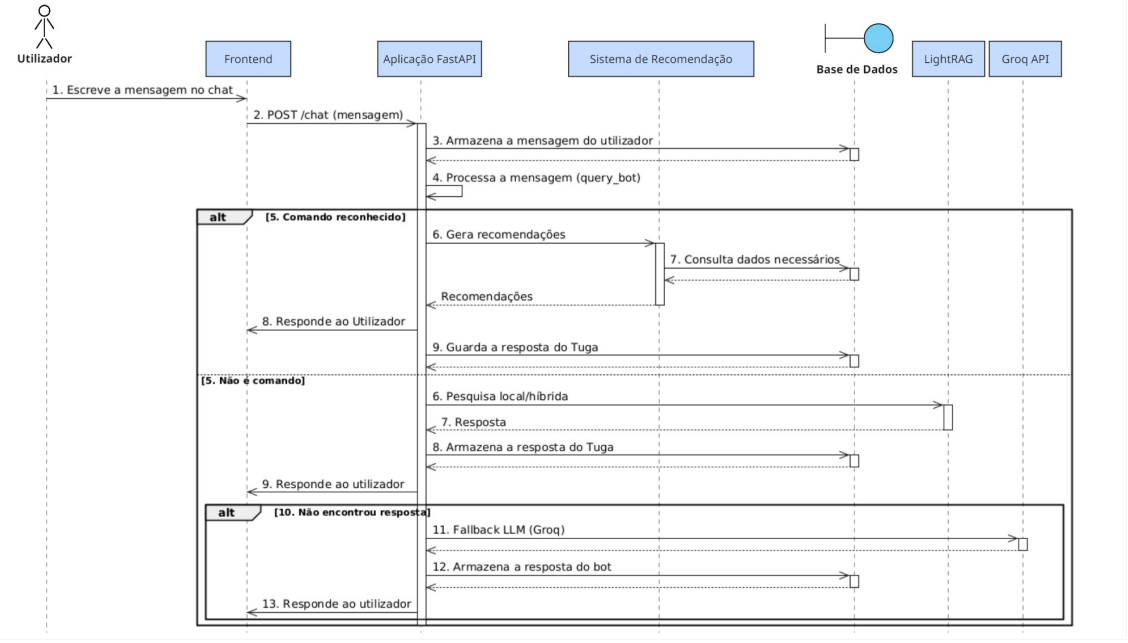

Figura 4.9: Diagrama de Sequência da Interação com o Tuga

4.8 Resolução de Exercícios

Com o intuito de desenvolver um mecanismo de resolução de exercícios simples, procedeu-se à conceção de uma abordagem que fornece respostas completas e úteis ao utilizador. Para tal, utilizou-se o modelo LLM *llama-3.3-70b-versatile* através da API da plataforma *Groq*, para a validação das respostas fornecidas pelo utilizador face aos exercícios propostos. Esta abordagem permite um processo de resolução mais robusto, com respostas mais claras e precisas, além de possibilitar a extensão do sistema para abranger exercícios de maior complexidade.

O modelo é utilizado para analisar a pergunta e a resposta correta armazenadas na base de dados, com base num *prompt* pré-definido, visualizado na lista 4.2. O sistema verifica a exatidão da resposta, indicando se está correta ou incorreta, e apresenta a resposta correta detalhando os erros cometidos. A Figura 4.22 exemplifica este processo de validação, facilitando a compreensão do mecanismo de aprendizagem.

```
prompt = f"""
            Pergunta: {question}
            Resposta correta: {correct_answer}
            Resposta: {user_answer}
            Indica se está correta ou incorreta.
            Responde em Português de Portugal.
            """
```

Listing 4.2: Prompt para Validar Resposta

4.9 Módulos

Cada módulo representado na arquitetura do *frontend*, ilustrado na Figura 3.6 corresponde a uma página da plataforma O Lama Sabe, através da qual o utilizador interage com as diferentes funcionalidades disponíveis. O design da interface foi concebido de forma a ser intuitivo e acessível, recorrendo a uma paleta de cores em tons de verde e azul-esverdeado, de modo a transmitir uma experiência visual agradável e coerente.

A navegação entre as diferentes páginas da plataforma é facilitada através da barra de navegação superior, cujo *layout* adaptou-se a partir de um exemplo disponibilizado pelo *W3Schools* (W3schools s.d.[a]). O mesmo para o *Login* e Registo, que foram adaptados a partir de um exemplo disponibilizado pelo *W3Schools* (W3schools s.d.[b]).

O logótipo da plataforma O Lama Sabe foi criado com recurso à ferramenta online disponível em Design.com s.d.

4.9.1 Visão Geral

A página "Visão Geral", ilustrada na Figura 4.10, disponibiliza informações introdutórias sobre a plataforma O Lama Sabe e o assistente virtual Tuga. Além disso, destaca os 10 conteúdos educacionais mais apreciados pelos utilizadores, com base no número de gostos atribuídos. Esta secção permite aos utilizadores conhecerem melhor a plataforma e as funcionalidades disponíveis.

Figura 4.10: Informação Sobre a Plataforma na Página Visão Geral

4.9.2 Registo

Para se registar, o utilizador deve fornecer o nome, *username*, palavra-passe e confirmação da palavra-passe. Caso deseje, o utilizador pode selecionar opcionalmente as preferências de recomendações disponíveis, como ilustrado na Figura 4.11. Além disso, deverá aceitar os Termos e Condições da plataforma, que podem ser consultados através da hiperligação disponibilizada.

Figura 4.11: Página de Registo

4.9.3 Login e Logout

O formulário de *Login*,ilustrado na figura 4.12 é semelhante ao de registo, sendo necessário apenas que o utilizador introduza o *username* e palavra-passe para se autenticar e obter acesso completo aos recursos e funcionalidades da plataforma.

Figura 4.12: Página Login

Para terminar a sessão, o utilizador deve efetua *Logout* da plataforma, clicando na opção correspondente na barra de navegação, conforme ilustrado na Figura 4.13.

Figura 4.13: Opção *Logout* na Barra de Navegação

4.9.4 Assistente Virtual

Ao aceder à página "Assistente Virtual", é apresentada ao utilizador uma mensagem introdutória, ilustrada na Figura 4.14, que o ajuda a compreender a principal função do Tuga. A Figura 4.15 exemplifica a execução de um comando personalizado, neste caso a resposta do sistema à instrução “recomenda livros”.

O *layout* do *chat* foi adaptado a partir de um exemplo existente no site do *Bootstrap*, e a imagem de perfil do Tuga utiliza como fundo uma fotografia da Ponte da Arrábida (Bootdey 2018; Gonçalves 2007).

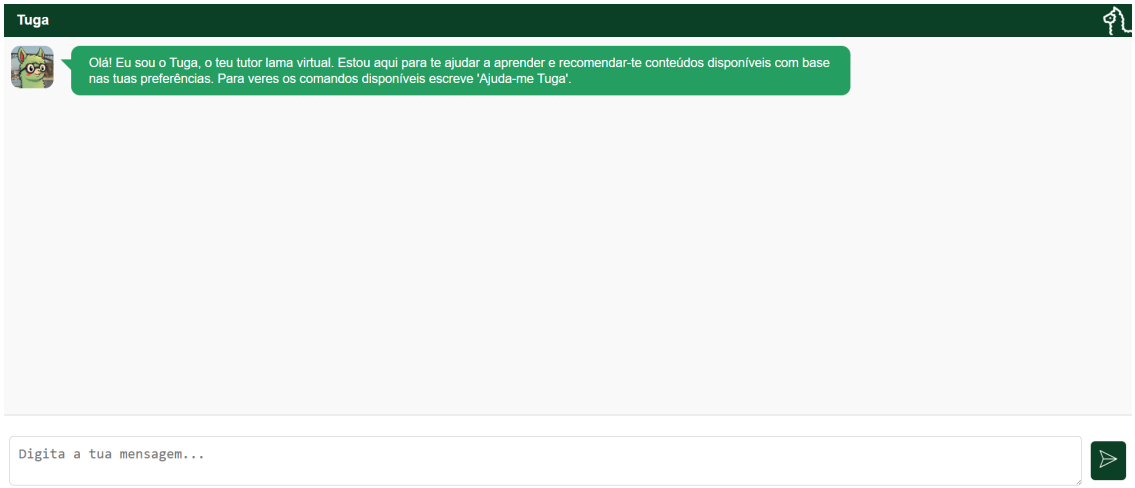

Figura 4.14: *Chat* do Assistente Virtual Tuga

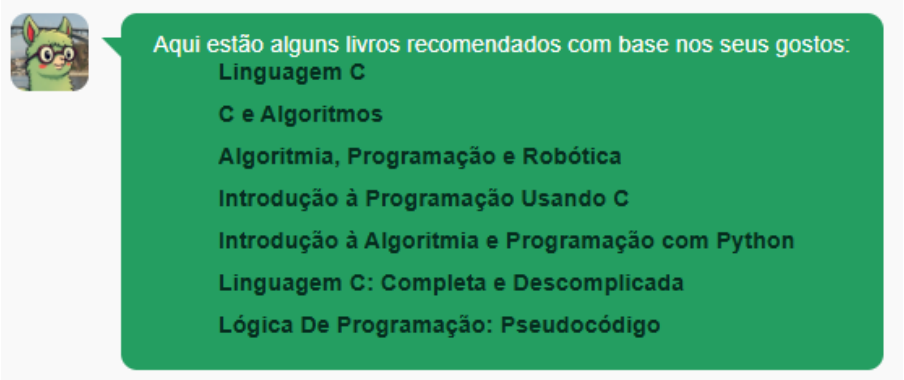

Figura 4.15: Resposta ao Comando Personalizado

4.9.5 Histórico

Na página de Histórico são apresentadas as mensagens trocadas entre o assistente virtual Tuga e o utilizador. Cada mensagem apresentada está acompanhada da respetiva data e hora, conforme ilustrado na Figura 4.16.

Figura 4.16: Histórico de Mensagens

4.9.6 Página de Recomendações

As recomendações personalizadas são apresentadas ao utilizador na página "Recomendações", para além dos comandos personalizados disponibilizados pelo assistente virtual. Para cada conteúdo recomendado, é possível o utilizar gostar do conteúdo, visualizar e ver o respetivo tipo de conteúdo, como ilustrado na Figura 4.17. A apresentação destas recomendações segue o fluxo representado no diagrama BPMN ilustrado na Figura 3.10.

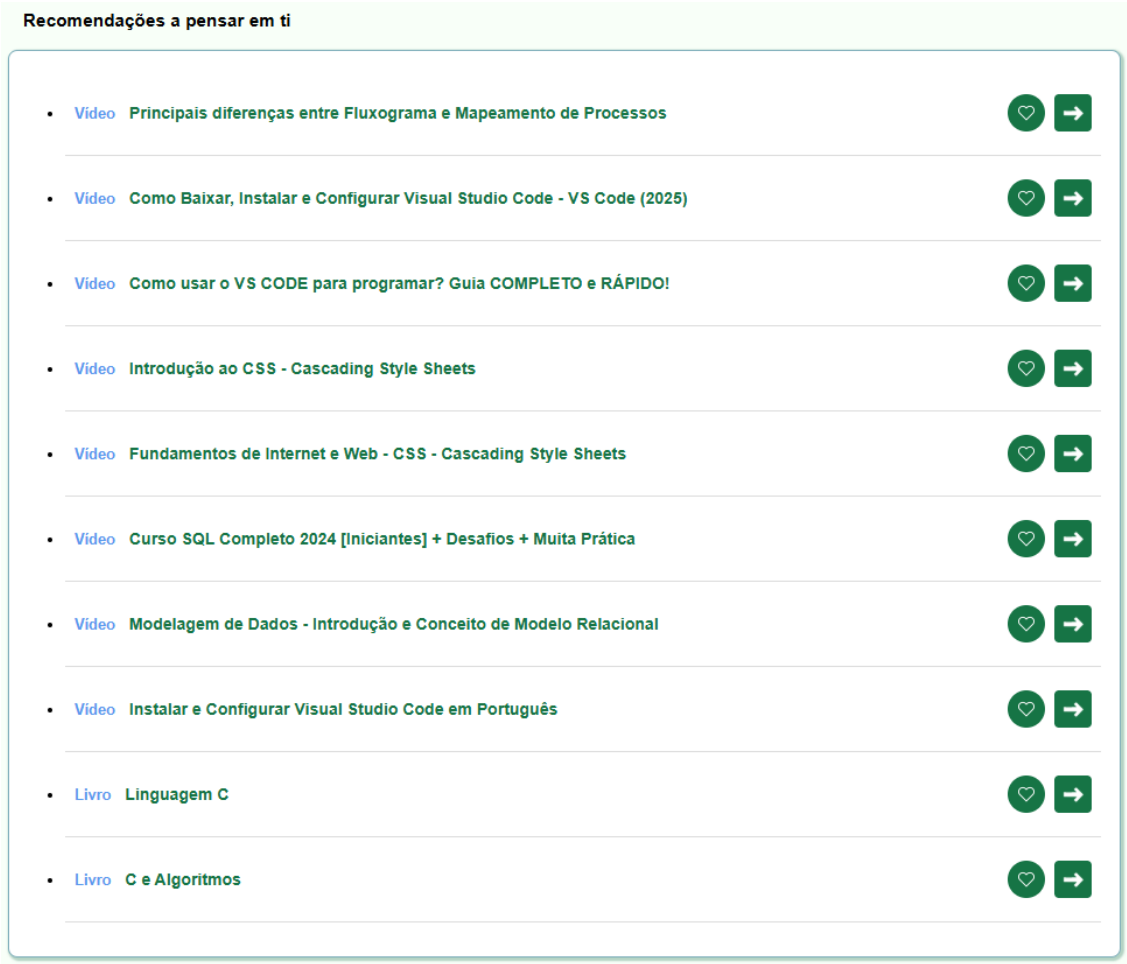

Figura 4.17: Página de Recomendações

4.9.7 Conteúdos Educacionais

Os conteúdos educacionais são apresentados ao utilizador separado por tipo de conteúdos, cada um tem a sua correspondente página "Vídeos", "Documentos", "Livros"e "Fichas de Exercícios"como ilustrado na Figura 4.18 da página "Livros". Para cada conteúdo o utilizador pode gostar , tirar o gosto, aceder ao conteúdo e filtrar os conteúdos por categoria.

Figura 4.18: Página Livros

Na página "Livros"é apresentada apenas a imagem representativa de cada livro. Nas páginas "Documentos"e "Exercícios", em vez da imagem, são exibidos o título e a descrição do conteúdo, conforme ilustrado nas figuras 4.19 4.21. Na página "Vídeos", ilustrada na figura 4.20, são apresentados o título e o próprio vídeo, incorporado através de um *player* do *YouTube*, permitindo a sua visualização diretamente na plataforma.

Figura 4.19: Página Documentos

Figura 4.20: Página Vídeos

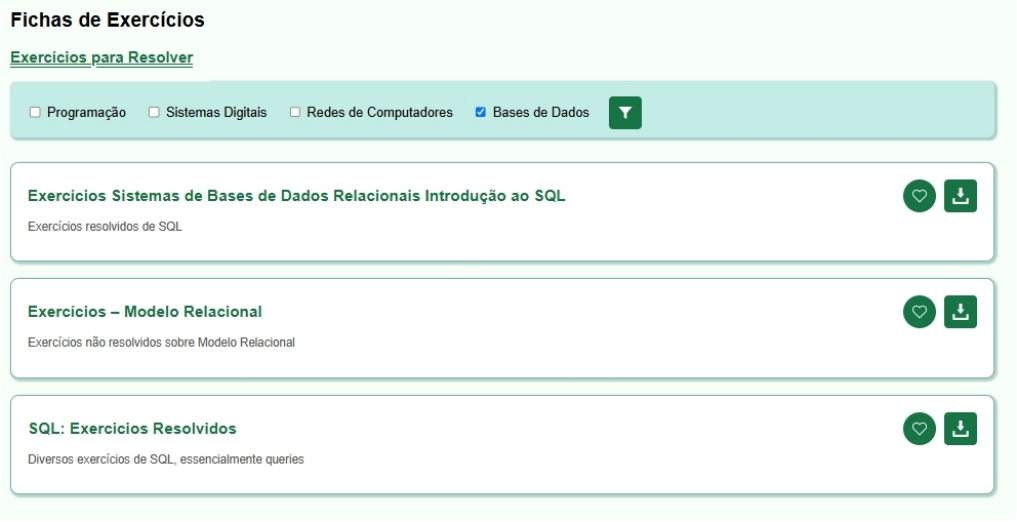

Figura 4.21: Página Exercícios

4.9.8 Resolver Exercícios

É possível resolver exercícios diretamente na plataforma, proporcionando uma experiência de aprendizagem mais interativa e centrada na prática.

Os exercícios disponíveis para resolução integram as fichas de exercícios previamente disponibilizadas na plataforma. A Figura 4.22 ilustra um dos exercícios resolvidos, cuja resposta foi validada pelo LLM. Como é possível observar, o LLM identificou que a resposta fornecida estava incorreta, apresentando a solução correta acompanhada de uma explicação que justifica o erro. Esta abordagem contribui para uma aprendizagem mais eficaz, ao permitir que o utilizador compreenda os erros cometidos e consolide o conhecimento de forma orientada.

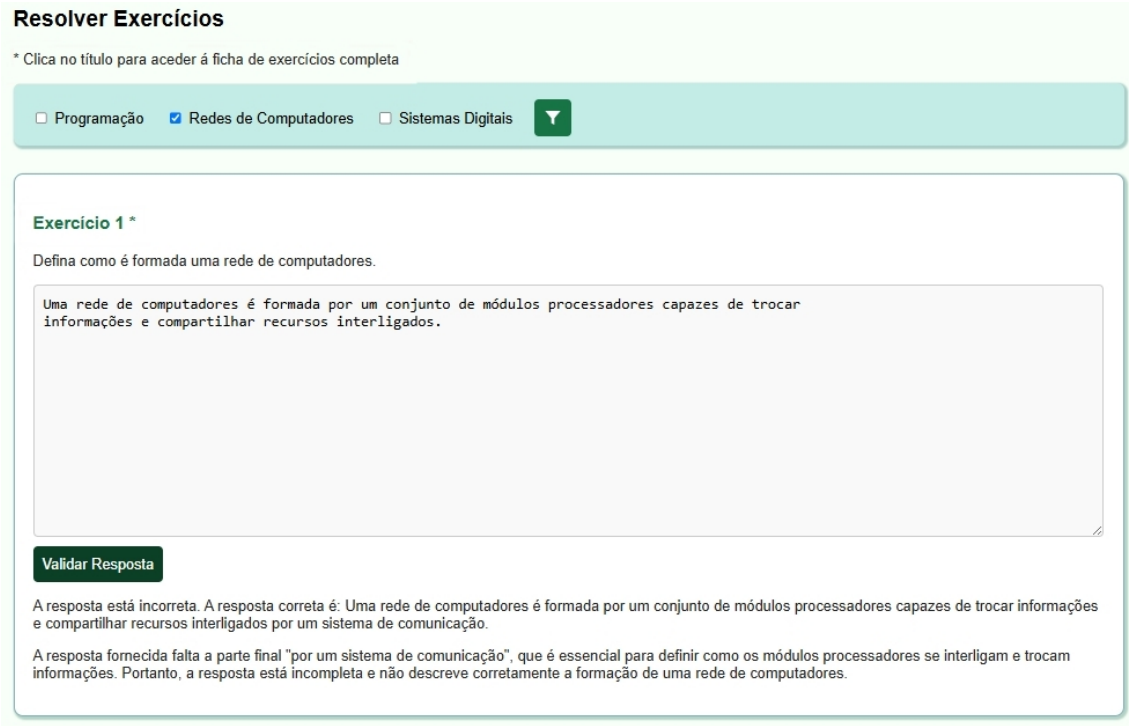

Figura 4.22: Resolução de Exercícios

4.10 Sumário

O presente capítulo descreveu a implementação da plataforma O Lama Sabe, abrangendo o desenvolvimento da própria plataforma, do assistente virtual Tuga e do sistema de recomendação. A explicação foi complementada com fluxogramas e um diagrama de sequência, que permitiram ilustrar de forma clara o fluxo de interação entre os componentes do sistema.

A plataforma foi concebida para ser intuitiva e funcional, oferecendo as ferramentas essenciais para apoiar a aprendizagem dos alunos. Foi construido um conjunto de dados orientado parte do público-alvo dos testes à plataforma, estudantes do Curso Profissional de Programação e Sistemas Informáticos, com base em conteúdos curriculares dessa área.

Os conteúdos educativos são organizados por tipo (livros, documentos, vídeos, fichas de exercícios) facilitando a consulta dos mesmos. Os utilizadores têm ainda a possibilidade de marcar conteúdos como favoritos, funcionalidade que suporta o sistema de recomendação baseado em filtragem por conteúdo.

Para além de permitir a visualização de conteúdos diretamente da fonte, a plataforma disponibiliza exercícios interativos que podem ser resolvidos na própria interface. Estes exercícios também se encontram organizados nas respetivas fichas, facilitando o seu enquadramento pedagógico.

O assistente virtual Tuga complementa a experiência da plataforma ao responder a dúvidas dos utilizadores e sugerir conteúdos através de comandos personalizados. A sua implementação integra diversos componentes, incluindo os serviços externos *LightRAG* e *Groq*, que permitem fornecer respostas mais precisas e completas. Em adição com os comandos personalizados, esta abordagem contribui para uma experiência mais personalizada e orientada às necessidades do utilizador.

Capítulo 5

Caso de Estudo

O presente capítulo descreve o caso de estudo utilizado para a avaliação da plataforma, do assistente virtual e do sistema de recomendação desenvolvidos. O principal objetivo é identificar oportunidades de melhoria da eficácia e usabilidade da plataforma, bem descobrir novas funcionalidades a serem implementadas. Adicionalmente, procura-se compreender se os utilizadores estariam dispostos a adotar a solução desenvolvida.

5.1 Propósito

O presente caso de estudo teve como objetivo principal avaliar a solução desenvolvida para a melhoria e expansão das funcionalidades da plataforma O Lama Sabe. Para além disso, pretendeu-se testar a eficácia da plataforma, do assistente virtual e do sistema de recomendação como ferramentas de apoio ao estudo. Os objetivos específicos do estudo são os seguintes:

- Verificar se as recomendações fornecidas aos utilizadores são adequadas, correspondendo às suas expectativas e demonstrando utilidade no processo de aprendizagem;
- Avaliar se o assistente virtual Tuga responde de forma pertinente, clara e satisfatória às dúvidas e necessidades apresentadas pelos utilizadores;
- Analisar se a plataforma é intuitiva, visualmente apelativa e se proporciona uma navegação fluida entre as diferentes secções do site;
- Avaliar a relevância dos conteúdos disponibilizados;
- Recolher *feedback* dos utilizadores, através de um questionário, para aferir a sua predisposição para utilizar a plataforma como ferramenta de apoio ao estudo no futuro, bem como identificar potenciais melhorias, funcionalidades adicionais a implementar.

5.2 Metodologia

No caso de estudo utilizou-se uma solução experimental da plataforma, que envolveu a participação de indivíduos selecionados para a realização de testes à plataforma. Paralelamente, realizou-se uma análise das respostas fornecidas pelo assistente virtual, com o objetivo de avaliar a sua eficácia na resolução das questões colocadas pelos utilizadores. Adicionalmente, procedeu-se paralelamente à avaliação do sistema de recomendação, recorrendo a métricas específicas de desempenho.

A seleção dos participantes baseou-se na acessibilidade e disponibilidade durante o período de realização do estudo, caracterizando-se como uma amostragem por conveniência. Como critério de seleção, os participantes deveriam possuir algum contacto prévio com linguagens de programação. Este critério garante um nível mínimo de literacia tecnológica necessário à utilização da plataforma. Adicionalmente, definiu-se como critério de inclusão o domínio básico da língua portuguesa, de forma a assegurar que todos os participantes conseguiam interagir com a plataforma e explorar eficazmente as suas funcionalidades.

5.3 Procedimento

O estudo foi desenvolvido em várias fases. A primeira consistiu em os participantes interagirem com a plataforma, iniciando pela leitura da página "Visão Geral", onde se encontra descrita a informação sobre as funcionalidades disponíveis. Em seguida, os participantes procederam ao registo na plataforma, preenchendo os dados solicitados.

Após a autenticação, foi dada liberdade aos participantes liberdade para explorar para explorar a plataforma autonomamente ou, em alternativa, seguir um conjunto de passos previamente definidos, que incluíam:

- Explorar as diferentes páginas da plataforma;
- Visualizar conteúdos educacionais e gostar de pelo menos três desses conteúdos;
- Interagir com o assistente virtual, colocando questões relacionadas com um dos temas dos conteúdos educacionais;
- Aceder à página de recomendações ou solicitar recomendações diretamente ao assistente virtual;
- Encerrar a sessão no site;

Durante o processo, cada um dos participantes foi acompanhado com o objetivo de esclarecer eventuais dúvidas relativas ao funcionamento da plataforma, bem como para observar e compreender possíveis dificuldades relacionadas com a sua interação com as diferentes funcionalidades da plataforma.

Adicionalmente, foram recolhidas opiniões e sugestões por parte dos participantes através de um questionário, com o intuito de identificar oportunidades de melhoria ao nível da usabilidade, do design e das funcionalidades existentes, assim como recolher propostas para a introdução de novas funcionalidades e possíveis melhorias.

A segunda fase consistiu na análise das respostas fornecidas pelo assistente virtual Tuga aos participantes, com o objetivo de melhorar a qualidade das interações e detetar eventuais erros ou limitações no seu desempenho.

A terceira fase focou-se na análise dos resultados obtidos a partir das métricas de avaliação do sistema de recomendações, que eram obtidos sempre que o utilizador acedesse à página de recomendações. O cruzamento desses dados com as opiniões recolhidas através do questionário, permitiu avaliar a relevância e utilidade das recomendações apresentadas aos utilizadores.

5.4 Recolha de dados

Os dados foram recolhidos através do preenchimento dos questionários e da inserção de dados na base de dados da plataforma. A Tabela 5.1 apresenta as questões incluídas no questionário. As perguntas identificadas como opcionais não eram de preenchimento obrigatório.

A estrutura das perguntas foi a seguinte:

- A pergunta 1 apresentou as opções de escolha múltipla "No ensino escolar", "No ensino superior"e "Já acabei os estudos ou estou a trabalhar mas quero aprender".
- As perguntas 2, 5, 6, 11 e 14 apresentou as opções de escolha múltipla "Sim"e "Não".
- A pergunta 3 apresentou as opções de escolha múltipla "Sim", "Sim, mas não utilizei"e "Não".
- As perguntas 8, 10 e 18 apresentaram as opções de escolha múltipla "Sim", "Talvez"e "Não".
- As perguntas 4 e 12 foram baseadas numa escala linear de 1 a 5.
- As perguntas 7, 9, 13, 15, 17 e 19 foram de resposta aberta (texto livre).

Tabela 5.1: Questionário de Opinião

Secção	**Pergunta**
Informação Geral	1. Estás atualmente a estudar?
	2. Usas ou usaste conteúdos educacionais além dos fornecidos pelos professores e outras entidades educadoras para te auxiliar nos estudos? (ex.: vídeos, outras plataformas, manuais)
Assistente Virtual Tuga	3. Tens conhecimento dos comandos personalizados de recomendação do Tuga?
	4. Quão relevante achaste a informação fornecida pelo Tuga?
	5. Gostaste da maneira que o Tuga escreve?
	6. Consideraste relevante a citação das fontes?
	7. Houve algo que pediste ou perguntaste ao Tuga em particular que não tenha correspondido às tuas expectativas? (opcional)
	8. Usarias o Tuga como apoio ao estudo quando não encontras o que precisas?
	9. Diz-nos algo que não tenhas gostado ou que gostarias de ver melhorado no Tuga. (opcional)
Recomendação de Conteúdo	10. Achaste que o conteúdo apresentado foi diversificado (ex.: vídeos)?
	11. Encontraste a página das recomendações personalizadas?
	12. Quão úteis foram as recomendações? (opcional)
	13. Diz-nos algo que não tenhas gostado ou que gostarias de ver melhorado em relação ao conteúdo e às recomendações. (opcional)
Plataforma Educacional O Lama Sabe	14. Achaste a plataforma intuitiva e fácil de usar?
	15. Caso tenhas respondido "Não", explica o motivo. (opcional)
	16. Quão visualmente apelativa achaste a plataforma?
	17. Diz-nos algo que não tenhas gostado ou que gostarias de ver melhorado em relação ao design e disposição das páginas da plataforma. (opcional)
Considerações Finais	18. Caso a plataforma O Lama Sabe fosse lançada com diversos temas de estudo acessíveis a todos, usarias?
	19. Se tiveres mais algum comentário a fazer, por favor partilha aqui. (opcional)

5.5 Sumário

No presente capítulo, definiu-se o caso de estudo aplicado à plataforma educacional O Lama Sabe, incluindo os seus componentes, o assistente virtual Tuga e o sistema de recomendação. Este estudo permite avaliar e melhorar a usabilidade da plataforma, otimizar funcionalidades existentes, introduzir novas funcionalidades e aperfeiçoar o design da interface. Para além disso, possibilitou compreender se os participantes considerariam utilizar a plataforma como ferramenta de apoio ao estudo. Importa referir que utilizou-se um protótipo experimental, não representando ainda o protótipo final da plataforma.

Capítulo 6

Resultados e Discussão

Neste capítulo são analisados e discutidos os resultados obtidos a partir do caso de estudo definido no Capítulo 5. Os dados analisados foram utilizados para melhoria da plataforma educacional O Lama Sabe, do assistente virtual Tuga e Sistema de Recomendação.

6.1 Caracterização da Amostra

Foi registada uma amostra total de 16 participantes no caso de estudo, dos quais 9 alunos do curso profissional de Gestão e Programação de Sistemas Informáticos da Escola Secundária de Valongo, 3 são estudantes do ensino superior e 4 são indivíduos já inseridos no mercado de trabalho.

Dos 16 participantes, apenas 1 referiu nunca ter utilizado conteúdos educacionais que não fossem fornecidos por professores ou instituições de ensino, o que evidencia uma elevada procura por conteúdos educacionais complementares aos disponibilizados pelas entidades educadoras.

6.2 Avaliação ao Assistente Virtual Tuga

Os resultados obtidos no âmbito do caso de estudo foram analisados com o intuito de aferir a eficácia do Tuga no apoio complementar aos processos de aprendizagem dos alunos.

6.2.1 Análise das Interações com o Tuga

As mensagens trocadas entre o assistente virtual Tuga e os participantes foram analisadas, incluindo a verificação da ocorrência de erros, a avaliação da escrita dos participantes e a análise da pertinência e completude das informações fornecidas pelo assistente.

Os temas abordados pelos participantes nas interações com o Tuga, ordenadas por frequência decrescente de ocorrência, foram as seguintes:

- Linguagem de programação C;
- Introdução à programação;
- Recomendação de conteúdos, para além dos comandos personalizados;
- Integração de Java com SQL;
- Linguagem de programação C#;
- *Machine Learning* em *Python*;

- Resolução de problemas relacionados com *bluescreen* em computadores;
- Interações sociais com o Tuga, incluindo perguntas sobre a identidade do Tuga e uso de *emojis*;
- Temas diversos dos quais, *Star Wars*, *anime*, entre outros;

De modo geral, os utilizadores interagiram com o Tuga de forma informal e direta. As respostas fornecidas revelaram-se adequadas e completas, sobretudo nas áreas relacionadas com programação, beneficiando da integração de informação adicional através do mecanismo RAG. A utilização da técnica RAG permitiu mitigar limitações associadas ao modelo LLM, assegurando respostas mais contextualizadas, coerentes e ajustadas ao domínio específico (Yunfan Gao et al. 2024).

Tanto os comandos personalizados como as respostas geradas pelo Tuga, funcionaram corretamente durante o período do caso de estudo, não se registando falhas de execução. Verificou-se, no entanto, um comportamento em que o assistente interpretava incorretamente frases que continham palavras-chave associadas a comandos, tratando-as como instruções diretas. Esta situação foi posteriormente corrigida, tendo sido também melhorado o formato de apresentação das fontes de informação, tornando-o mais claro e conciso.

A ocorrência evidenciou a importância da realização de testes abrangentes, que não tinham sido detetados em fases anteriores, sublinhando a relevância de efetuar avaliações com amostras de maior dimensão.

6.2.2 Análise do Feedback Sobre o Tuga

Nas respetivas perguntas do questionário sobre o Tuga, 9 participantes afirmaram ter utilizado os comandos personalizados, 1 demonstrou ter conhecimento dos mesmos mas não os utilizou, e os restantes 6 indicaram não ter conhecimento sobre esses comandos. No entanto, durante o acompanhamento realizado, verificou-se que todos os participantes utilizaram os comandos personalizados, ainda que de forma não intencional ou inconsciente.

O gráfico de barras ilustrado na Figura 6.1 ilustra as respostas relativas à relevância da informação fornecida pelo *Tuga*. Numa escala de 1 a 5, sete participantes atribuíram a cotação máxima (5), sete atribuíram a cotação 4 e dois participantes atribuíram a cotação 3, resultando numa média ponderada de 4,3.

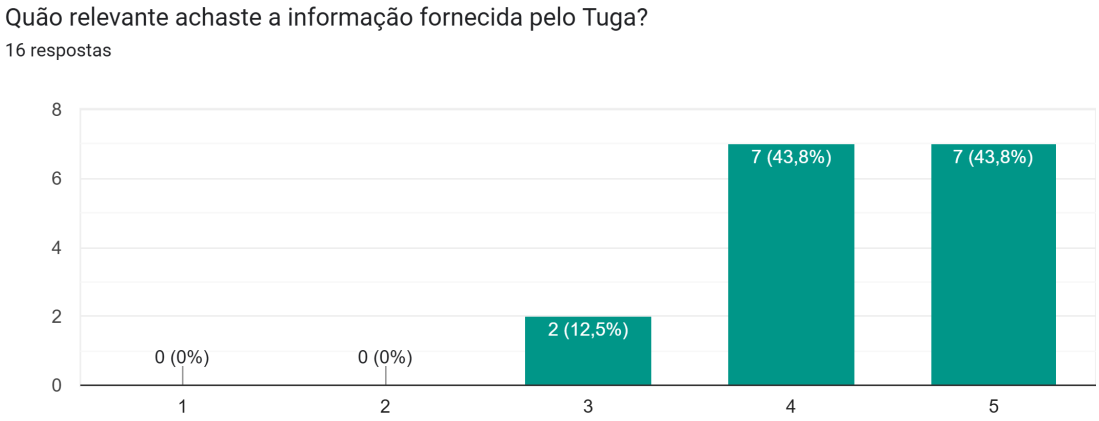

Figura 6.1: Análise à Relevância da Informação Fornecida pelo Tuga

Todos os participantes apreciaram a clareza da escrita e a presença de citações de fontes, com exceção de apenas um participante, que não considerou relevante a citação das fontes utilizadas.

Na pergunta 7 do questionário do caso de estudo, "Houve algo que pediste ou perguntaste ao Tuga em particular que não tenha correspondido às tuas expectativas?"apenas um participante relatou que solicitou sugestões de vídeos sobre a linguagem C#, tendo o Tuga sugerido canais do *YouTube* em vez de vídeos específicos, como desejado pelo participante . Adicionalmente, foi também referida, já anteriormente identificada e corrigida, de respostas com frases que ativavam comandos devido à presença de palavras-chave.

O gráfico de pizza apresentado na Figura 6.2 mostra as respostas relativas à intenção de utilização futura do *Tuga* como apoio complementar ao estudo. Metade dos participantes indicou que talvez viesse a utilizar o Tuga, enquanto 7 responderam que sim e apenas 1 indicou que não o utilizaria.

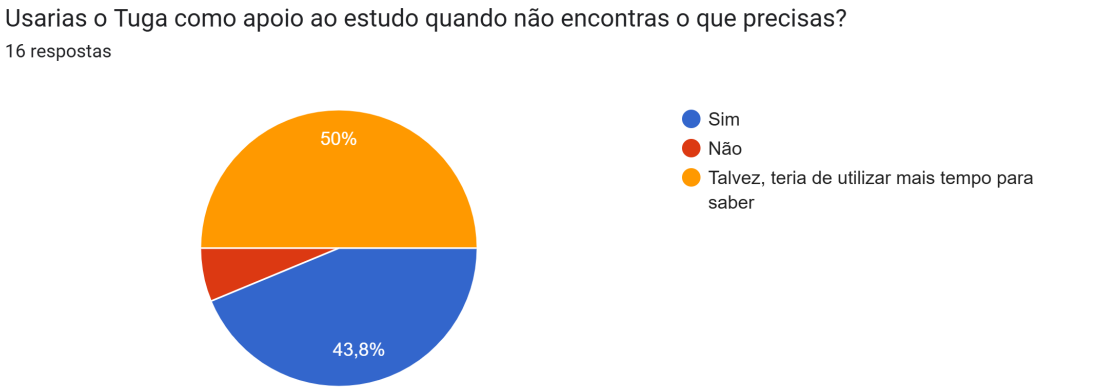

Figura 6.2: Análise ao Assistente Virtual Tuga como Apoio ao Estudo

Os resultados obtidos revelam uma perceção maioritariamente satisfatória relativamente ao desempenho do *Tuga*. No entanto, o potencial de utilização do assistente Virtual Tuga revelou-se em média neutra por parte dos participantes. Autores como (Chatwattana, Yangthisarn e Tabubpha 2024; Dam et al. 2024), embora ressaltem o elevado potencial dos LLMs na área educativa, ainda existe margem para aperfeiçoamento no contexto do ensino, de forma a aumentar o envolvimento e a adoção por parte dos utilizadores.

Adicionalmente, foram recolhidas diversas sugestões de novas funcionalidades e melhorias, nomeadamente:

- Envio de mensagens ao premir a tecla *Enter*;
- Suporte para escrita em múltiplas linhas utilizando as teclas *Shift + Enter*;
- Capacidade do Tuga processar e memorizar a conversa, permitindo fornecer exemplos com base no contexto se o utilizador pedir, após a resposta do Tuga;
- Bloqueio do envio de novas mensagens enquanto o Tuga não responde;
- Alteração do ícone do botão de envio enquanto o utilizador aguarda resposta, para indicar estado de "a carregar";
- Permitir editar e regenerar as respostas;
- Upload de Arquivos;
- Implementação de pesquisa com base em informações mais recentes;
- Abertura de *links* em novos separadores;
- Disponibilização de informações mais detalhadas e sugestões mais relevantes;
- Melhoria na apresentação textual para facilitar a leitura e compreensão;

Das sugestões recolhidas, foram implementadas aquelas consideradas mais relevantes, contidas nos requisitos funcionais do sistema. Estas melhorias permitiram aumentar a usabilidade e a eficiência na interação com o Assistente Virtual Tuga, proporcionando uma experiência mais intuitiva, eficaz e satisfatória para os utilizadores.

6.3 Análise do Conteúdo e Recomendações

As respostas às perguntas do questionário relativas aos conteúdos educacionais e às recomendações, analisadas em conjunto com as métricas obtidas, permitem uma avaliação mais detalhada da qualidade e relevância das sugestões fornecidas pelo sistema, bem como do seu impacto na experiência dos utilizadores.

6.3.1 Análise do Feedback Sobre o Conteúdo e Sistema de Recomendação

Dos participantes, 14 participantes consideraram os tipos de conteúdos apresentados como diversificados. Do qual 2 participantes selecionaram a opção "talvez", por não terem uma perceção clara de que outros tipos de conteúdos poderiam ser incluídos. Nenhum participante respondeu negativamente, o que indica que a plataforma disponibiliza, de forma satisfatória, uma diversidade de tipos de conteúdos adequada às expectativas da maioria dos utilizadores.

Relativamente à página de recomendações, alguns participantes reportaram dificuldades em localiza-la. Do total, 12 conseguiram encontrar, enquanto 4 não a localizaram, o que

motivou a alteração da sua posição na interface da plataforma, bem como a mudança do respetivo título da página.

Entre os 12 participantes que acederam à secção de recomendações, foi utilizada uma escala de 1 a 5 para avaliar a utilidade das recomendações apresentadas. Os resultados encontram-se ilustrados no gráfico de barras da Figura 6.3. Dos participantes, 7 atribuíram a pontuação máxima (5), 6 classificaram com 4, e dois com 2, resultando numa média ponderada de 4,3. No entanto, apenas 12 participantes indicaram ter encontrado a página de recomendações, e não 15, conforme registado, verificando-se um erro no preenchimento das respostas por parte de alguns participantes, o qual considerou-se devidamente na análise dos dados. As respostas dos participantes foram globalmente positivas, evidenciando uma perceção favorável quanto à utilidade das recomendações de conteúdos.

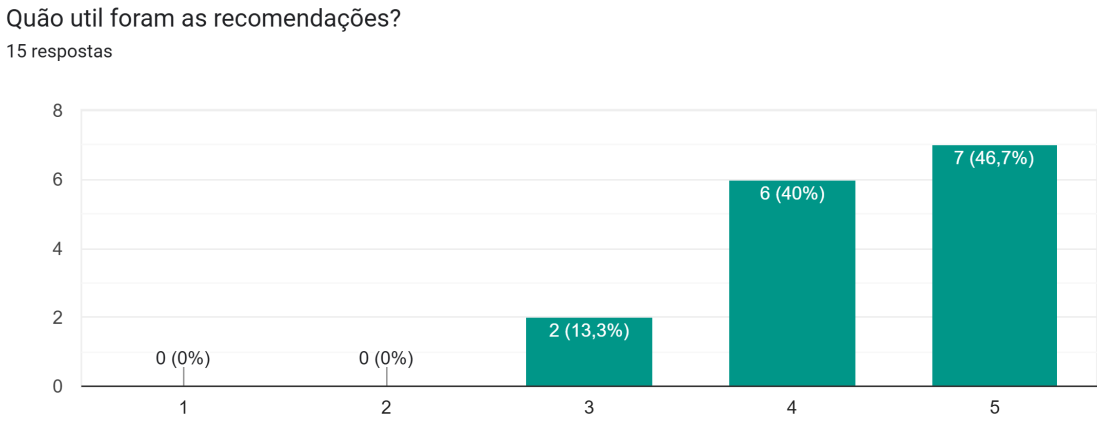

Figura 6.3: Análise à Utilidade das Recomendações

6.3.2 Métricas de Avaliação

Para avaliar a qualidade das recomendações fornecidas, foram utilizadas as métricas de precisão, *recall*, *F1-score* e acurácia, permitindo aferir a eficácia das sugestões apresentadas pelo sistema. Algumas destas métricas foram utilizadas pelos autores Bagunaid, Chilamkurti e Veeraraghavan 2022; Bhaskaran e Marappan 2023 na avaliação dos sistemas de recomendação por eles desenvolvidos, permitindo analisar o desempenho do sistema sem ser apenas através das opiniões dos utilizadores.

A análise foi realizada com recurso à média ponderada dos resultados, conforme ilustrado na Tabela 6.1, considerando a quantidade de gostos que o utilizador possuía no momento em que recebeu as recomendações. Esta abordagem possibilita uma avaliação mais aprofundada e contextualizada do desempenho do sistema de recomendação em diferentes cenários.

Tabela 6.1: Média Ponderada das Métricas de Avaliação por Número Gostos

Nº de Gostos	Precisão	Recall	F1-Score	Acurácia
1	0.002	0.045	0.005	0.763
2	0.061	0.750	0.111	0.644
3	0.150	1.000	0.261	0.800
4	0.079	0.667	0.140	0.572
5	0.074	0.800	0.136	0.400
6	0.208	0.833	0.333	0.765
9	0.124	0.889	0.217	0.324
10	0.118	0.800	0.205	0.271
11	0.136	0.818	0.234	0.306
17	0.200	0.882	0.326	0.271

Na análise dos valores obtidos das métricas apresentadas na Tabela 6.1, observa-se uma precisão consistentemente baixa em todos os casos, o que indica que apenas uma pequena parte das recomendações feitas foi, de facto, útil para os utilizadores.

Por outro lado, os valores de *recall* são elevados na maioria dos casos, excetuando-se a situação com apenas 1 gosto, o que demonstra que o sistema foi geralmente eficaz a identificar os conteúdos relevantes a recomendar.

Dado que o *F1-score* representa o equilíbrio entre precisão e *recall*, os seus valores baixos refletem a necessidade de melhorar esse equilíbrio, sobretudo aumentando a precisão sem comprometer a abrangência das recomendações.

Quanto à acurácia, observa-se uma tendência inicial para valores mais elevados, seguida de uma descida significativa à medida que aumenta o número de gostos. Isto sugere que o sistema de recomendação tende a funcionar melhor com utilizadores que possuem poucos gostos associados, possivelmente devido à quantidade de conteúdos disponíveis para recomendação.

Apesar dos resultados pouco satisfatórios obtidos na avaliação do sistema de recomendação com base em métricas, os dados recolhidos através do questionário revelaram uma perceção maioritariamente positiva. A maioria dos participantes achou útil as recomendações apresentadas, o que constitui um aspeto relevante na eficácia da recomendação de conteúdos, demonstrando que a avaliação ao sistema de recomendação pode não estar totalmente refletida apenas nas métricas.

6.3.3 Melhorias Finais ao Sistema de Recomendação

Com o objetivo de melhorar os resultados obtidos nas métricas de avaliação, foram introduzidas alterações no sistema de recomendação. Estas modificações incluíram a incorporação de uma lista personalizada de *stopwords*, o aumento da ponderação atribuída aos títulos dos conteúdos e a redução do número de recomendações apresentadas ao utilizador. Importa referir que a diminuição da quantidade de conteúdos recomendados foi aplicada apenas durante os testes, mantendo-se no funcionamento geral da plataforma a possibilidade de aceder a um número mais alargado de conteúdos, especialmente a pensar nos comandos personalizados do assistente virtual Tuga para filtragem. O processo de reformulação do sistema de recomendação encontra-se já descrito no Capítulo 4.

A Tabela 6.2 apresenta os novos valores das métricas obtidos após a implementação das melhorias no sistema de recomendação. Para a obtenção destes resultados, foram criados quatro utilizadores fictícios, cada um representando diferentes cenários de conteúdos apreciados. Este procedimento permitiu testar o desempenho do sistema em múltiplas situações, simulando diferentes perfis de utilização e preferências, de forma a avaliar o sistema de recomendação em contextos variados.

Tabela 6.2: Média das Métricas de Avaliação por Número de Gostos Após Melhoria ao Sistema de Recomendação

Nº de Gostos	**Precisão**	**Recall**	**F1-Score**	**Acurácia**
1	0.000	0.000	0.000	0.853
2	0.116	0.875	0.204	0.829
3	0.179	1.000	0.301	0.824
4	0.154	0.850	0.257	0.741
5	0.116	0.750	0.201	0.635
6	0.114	0.708	0.196	0.574
7	0.141	0.857	0.242	0.538
8	0.161	0.900	0.273	0.529
9	0.174	0.889	0.289	0.520
10	0.167	0.875	0.280	0.462
15	0.225	0.867	0.356	0.429

Os resultados obtidos evidenciam uma ligeira melhoria nos valores das métricas de avaliação, indicando um progresso no desenvolvimento do sistema de recomendação. No entanto, é importante considerar diversos fatores que podem influenciar estes resultados, sendo o principal o tamanho do conjunto de dados utilizado. Com apenas 85 conteúdos disponíveis, é provável que esta limitação tenha impacto significativo nos valores das métricas obtidas.

6.4 Interface da Plataforma

A interface da plataforma O Lama Sabe desempenha um papel crucial na captação da atenção dos utilizadores, sendo fundamental que proporcione uma experiência de utilização confortável, clara e intuitiva. As respostas obtidas através do questionário permitiram identificar oportunidades de melhoria.

Todos os participantes consideraram a interface intuitiva, com exceção de um utilizador, que referiu a ausência do botão de *Login/Logout* na página "Visão Geral"como uma limitação relevante. Esta observação foi tida em conta e ajustou-se a interface, passando a incluir esse botão também nessa página. Importa referir que, antes da realização dos testes, comunicou-se aos participantes que essa funcionalidade ainda se encontrava em fase de ajustes, motivo pelo qual essa funcionalidade não estava presente na barra de navegação da página mencionada.

De forma geral, conforme demonstrado no gráfico de barras da Figura 6.4, a plataforma é considerada intuitiva, mas com margem para melhorias. Numa escala de 1 a 5, 3 participantes atribuíram a pontuação máxima (5), 11 classificaram com 4 e 2 com 3, resultando numa média ponderada de 4,1.

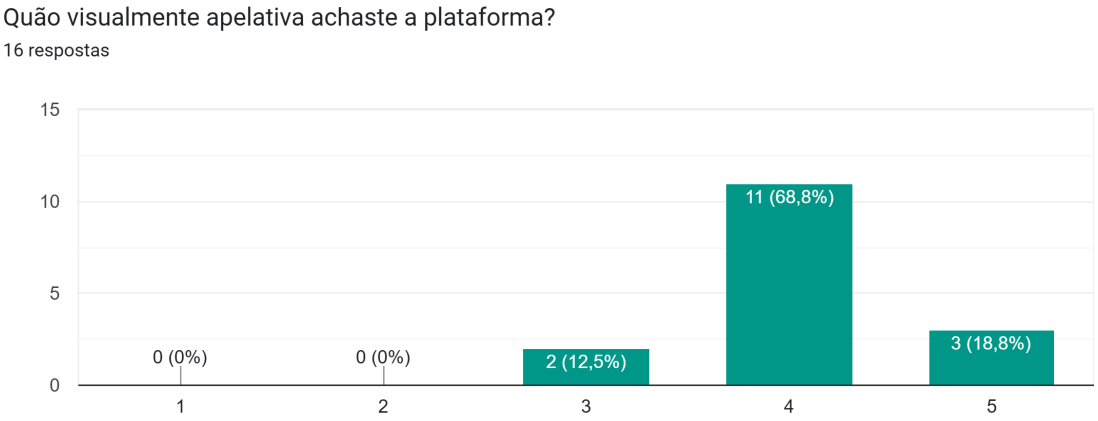

Figura 6.4: Análise à Interface da Plataforma

Adicionalmente, foram recolhidas diversas sugestões de melhoria da interface, sendo as seguintes:

- Alterar o ícone utilizado para abrir os conteúdos educacionais, por transmitir a ideia de partilha em vez de acesso;
- Uniformizar o *layout* das páginas de conteúdos, sugerindo que todas sigam o modelo da página de recomendações, com exceção da secção de livros, incluindo categorias e a possibilidade de expandir secções ao clicar no título dos documentos;
- Apresentar os conteúdos em formato de mosaico com imagens, semelhante à página "Livros", em vez de uma lista textual;
- Adicionar uma barra de pesquisa nas páginas dedicadas aos conteúdos educacionais;
- Agrupar as recomendações por tipo de conteúdo, organizando-as em filas distintas, e permitir que cada secção redirecione para a página correspondente à categoria;
- Inserir uma divisória entre os botões na barra de navegação, melhorando a separação visual;
- Melhorar a legibilidade, alterando a cor do fundo (verde) e do tipo de letra, considerados pouco apelativos;
- Tornar a barra de inserção de mensagens do assistente virtual Tuga dinâmica;
- Permitir o envio de mensagens através da tecla *Enter*, reservando *Ctrl + Enter* para inserção de quebras de linha;
- Nas recomendações de vídeos, apresentar os resultados em duas colunas verticais;
- Mudar a cor de fundo para preto;

As sugestões de melhoria relativas ao assistente virtual Tuga no que diz respeito aos botões de interação, já haviam sido apontadas por outros participantes na questão anterior relacionada com oportunidades de melhoria do Tuga. Parte das sugestões revelaram-se úteis para

a evolução da interface, tendo algumas sido implementadas, estas descritas nos requisitos funcionais e de interface no Capítulo 4.

Entre as melhorias aplicadas, se destaca em particular a substituição do ícone utilizado para aceder aos conteúdos educacionais, com o objetivo de tornar a função mais clara para o utilizador. Além disso, a barra de inserção de mensagens do assistente Tuga foi tornada dinâmica, permitindo ao utilizador ajustar a altura da mesma.

O estudo dos autores Abdi, Khosravi, Sadiq e Gasevic 2020 demonstrou que o excesso de informação pode dificultar a navegação para alguns utilizadores, evidenciando a importância de um design de interface simples e intuitivo, apresentando apenas a informação considerada necessária.Por essa mesma razão, as opiniões dos utilizadores foram fundamentais para orientar as melhorias implementadas na interface do assistente virtual Tuga.

6.5 Validação da Utilização da Plataforma

As perguntas do questionário destinadas a avaliar a plataforma e os seus componentes, nomeadamente o sistema de recomendação e o assistente virtual, como um todo, permitiram compreender a perceção global dos participantes da solução desenvolvida.

A questão que procurava perceber se os participantes utilizariam a plataforma caso esta fosse disponibilizada com uma maior variedade de temas de estudo revelou uma recetividade positiva. Conforme ilustrado no gráfico de pizza da Figura 6.5, nenhum dos participantes respondeu negativamente "Não", 5 participantes selecionaram a opção "Talvez"e os restantes 11 indicaram que utilizariam a plataforma "Sim". Estes resultados demonstram interesse por parte dos utilizadores, sugerindo que a plataforma poderá ser adotada como uma ferramenta útil no apoio ao estudo.

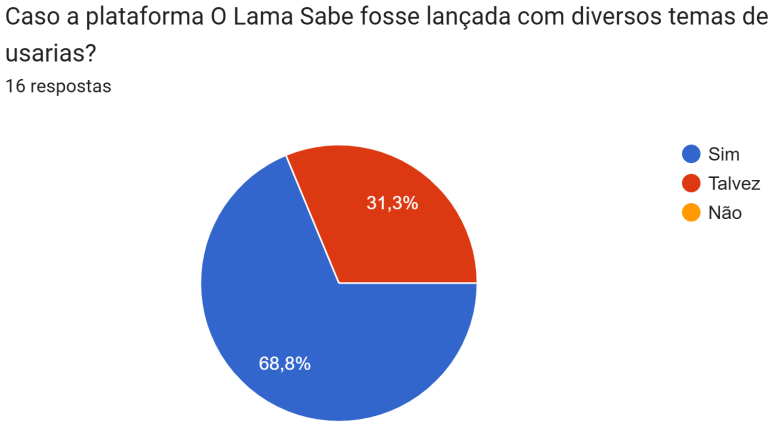

Figura 6.5: Análise à Utilização Futura da Plataforma

Adicionalmente, foram recolhidas várias sugestões de novas funcionalidades, bem como comentários finais dos participantes, sendo estes:

- Possibilidade de resolver exercícios diretamente na plataforma;
- Abrir os conteúdos em novos separadores, em vez de na mesma aba;

- Exibir imediatamente a mensagem do utilizador no *chat* do assistente virtual Tuga, indicando que este está a processar a resposta;
- Agrupar o histórico de conversas por sessão;
- Implementar deslocamento automático, *auto-scroll*, no *chat* do Tuga;
- Permitir a abertura do conteúdo ao clicar na imagem do livro;
- Introduzir um sistema de níveis com pontuação atribuída à visualização de conteúdos, que possa ser trocada por descontos em livros.

Para além das sugestões, dois dos participantes deixaram mensagens de incentivo, desejando sucesso no desenvolvimento do projeto. As respostas recolhidas revelaram-se uma combinação de comentários relacionados com funcionalidades específicas da plataforma, algumas das quais já haviam sido mencionadas pelos participantes em questões anteriores sobre o Tuga, o sistema de recomendação e a interface.

Das respostas propostas duas revelaram-se bastante interessantes, como a implementação de exercícios interativos e a introdução de mecanismos de gamificação através da atribuição de pontos. Destas duas sugestões, implementou-se na plataforma a funcionalidade de resolução de exercícios, perante o tempo disponível para desenvolvimento de novas funcionalidades e melhorias.

6.6 Sumário

Os resultados demonstraram que o assistente virtual Tuga funcionou corretamente, sem ocorrência de erros, apresentando respostas completas e bem estruturadas. Os participantes destacaram positivamente a clareza da linguagem utilizada, bem como a referência às fontes de informação incluídas nas respostas. Estes aspetos evidenciam que a utilização da abordagem RAG contribui para mitigar limitações comuns dos LLMs, permitindo fornecer respostas mais fiáveis e contextualizadas aos utilizadores (Yunfan Gao et al. 2024).

No entanto, apesar do bom desempenho técnico e da qualidade das interações, metade dos participantes manifestou incerteza quanto à utilização do assistente como apoio ao estudo. Este resultado sugere a necessidade de explorar o uso de LLMs de maior dimensão, bem como de ampliar a base de conteúdos disponibilizados ao modelo através do mecanismo de RAG. Importa referir, que modelos de maior escala implicam custos mais elevados, por exemplo monetários e de hardware, os quais podem constituir um entrave ao desenvolvimento e implementação da solução (Bang et al. 2024; Rodrigues, Vasconcelos e Maciel 2020). Adicionalmente, um dos participantes afirmou que não recorreria ao assistente virtual como ferramenta auxiliar, revelando assim incerteza quanto à utilidade prática do mesmo no contexto educativo.

A maioria dos participantes considerou os tipos de conteúdo disponibilizados na plataforma suficientemente diversificados, que destacou-se como um aspeto relevante para apoiar diferentes estilos de aprendizagem. Esta diversidade permite por exemplo, que alunos que preferem vídeos em vez de documentos, ou vice-versa, encontrem formatos que melhor se adequem às suas preferências.

Além disso, as recomendações fornecidas foram consideradas como úteis, tendo-se verificado um *feedback* positivo por parte dos participantes. No entanto, apesar das melhorias implementadas no sistema de recomendação, as métricas de avaliação obtidas continuaram

a apresentar valores pouco satisfatórios. Tal facto evidencia a necessidade de expandir o conjunto de conteúdos disponíveis e realizar mais testes, de forma a continuar o processo de aperfeiçoamento do sistema. Adicionalmente, poderá ser vantajoso considerar a integração de técnicas complementares, como a aplicação de regras e métodos de *clustering*, os quais têm demonstrado utilidade em diversos estudos (Bagunaid, Chilamkurti e Veeraraghavan 2022; Bhaskaran e Marappan 2023; Bouihi e Bahaj 2019).

Relativamente à interface da plataforma, todos os participantes a consideraram intuitiva com exceção de um, o que reforça a perceção geral de facilidade de utilização. A estética da plataforma foi igualmente bem recebida, embora ainda com oportunidades de melhoria a nível visual.

De forma global, cerca de três quartos dos participantes indicaram que utilizariam a plataforma O Lama Sabe caso esta fosse publicada com mais temas de estudo. O facto de nenhum participante ter respondido negativamente a esta questão é um sinal encorajador quanto à aceitação da solução, havendo possibilidade de aplicação da mesma num contexto mais alargado. Este resultado sugere que, com a devida expansão dos conteúdos e a continuidade na melhoria da experiência de utilização, a plataforma poderá tornar-se uma ferramenta eficaz de apoio ao estudo, com potencial para ser integrada em ambientes educativos.

O caso de estudo e as respostas recolhidas permitiram não só validar diversos aspetos da plataforma, do sistema de recomendação e do assistente virtual Tuga, como também identificar oportunidades concretas de melhoria. Com base neste *feedback*, foi possível implementar novas funcionalidades e ajustar a plataforma de forma a melhor corresponder às necessidades dos utilizadores.

Capítulo 7

Conclusões

O presente capítulo apresenta as conclusões decorrentes do desenvolvimento da solução proposta, englobando os resultados obtidos, os objetivos alcançados, as limitações identificadas durante o processo de desenvolvimento e as perspetivas para trabalho futuro.

7.1 Síntese

Com o aumento da utilização da tecnologia em Portugal, os estudantes recorrem cada vez mais a apoios digitais para complementar os seus estudos, utilizando recursos para além dos disponibilizados pelas instituições de ensino.

A revisão do estado da arte permitiu identificar diversos métodos de recomendação de conteúdo, para além dos métodos tradicionais, com especial foco na recomendação de materiais educacionais e cursos. Verificou-se ainda que a adoção de *chatbots* pode representar uma mais-valia no apoio aos estudantes. A análise de plataformas educacionais portuguesas revelou oportunidades de melhoria. Com exceção da *Khan Academy*, as plataformas que oferecem maior variedade de conteúdos e funcionalidades são, na sua maioria, de acesso pago através de subscrição (Khan Academy 2024).

A fase de análise e desenho do sistema foi desenvolvida em simultâneo com a implementação da plataforma O Lama Sabe, permitindo uma evolução contínua e integrada da solução. Esta abordagem proporcionou uma maior flexibilidade e capacidade de adaptação às necessidades que foram sendo identificadas ao longo do processo de desenvolvimento.

A plataforma integrou um conjunto de dados de diferentes tipos, vídeos, documentos, livros e fichas de exercícios, construído a pensar em parte da amostra dos participantes do caso de estudo. Além disso dispõe do assistente virtual Tuga para auxiliar os alunos e recomendações personalizadas, incluindo comandos personalizados disponíveis no assistente virtual.

O caso de estudo foi implementado em contexto real, sendo 9 dos participantes alunos do curso profissional de Gestão e Programação de Sistemas Informáticos da Escola Secundária de Valongo. Esta avaliação possibilitou não só a identificação de oportunidades de melhoria, como também a implementação de novas funcionalidades.

Os resultados do questionário foram maioritariamente positivos, à exceção prendeu-se com a perceção da utilidade do assistente virtual Tuga como apoio ao estudo, onde os participantes demonstraram alguma hesitação. Ainda assim, foram recolhidas diversas sugestões de melhoria por parte dos participantes, que contribuíram para o aperfeiçoamento da plataforma. As respostas obtidas evidenciam também o potencial da solução para ser aplicada em contextos educativos reais.

Em síntese, os objetivos propostos foram alcançados, tendo a plataforma desenvolvida e os seus componentes ido ao encontro das expectativas. Reconhece-se, no entanto, que existe espaço para evolução, sendo esse um ponto a explorar em trabalhos futuros.

7.2 Objetivos Cumpridos

A dissertação cumpriu com os objetivos inicialmente propostos, tendo todos sido alcançados. O estado da arte permitiu realizar uma análise aprofundada do funcionamento do ensino em Portugal, identificar as entidades educativas, os recursos de apoio ao estudo disponíveis e os métodos de recomendação aplicados ao conteúdo educacional. Além disso, foi investigada a integração de *chatbots* no contexto educativo.

Foi desenvolvida uma plataforma que superou as expectativas iniciais, funcionando como um recurso complementar de apoio ao estudo. O processo de desenvolvimento contemplou todos os requisitos definidos inicialmente pensados, tendo ainda sido integrados requisitos adicionais que surgiram ao longo do projeto.

A recolha de dados junto dos participantes revelou-se essencial, proporcionando contributos relevantes e sugestões de melhoria que possibilitaram o aperfeiçoamento contínuo da plataforma.

7.3 Resultados Alcançados

A plataforma O Lama Sabe disponibiliza um vasto conjunto de funcionalidades, permitindo o acesso gratuito a conteúdos educacionais diversificados, como vídeos, documentos, livros e exercícios. Os utilizadores podem classificar os conteúdos através de gostos ou removê-los das suas preferências, contribuindo assim para que o sistema de recomendação gere recomendações personalizadas, adaptadas ao perfil e estilo de aprendizagem de cada aluno. Adicionalmente, a plataforma possibilita a resolução de exercícios diretamente no ambiente digital, promovendo o envolvimento ativo dos estudantes.

Em comparação com a maioria das plataformas portuguesas analisadas, que exigem o pagamento de uma subscrição para aceder à totalidade das funcionalidades ou apresentam apenas um acesso parcial e limitado, a solução desenvolvida distingue-se de forma significativa (Escola Mágica 2024; Escola Virtual 2024; Leya Aula Digital 2024; Matematica.pt 2024). É totalmente gratuita e oferece uma maior diversidade de tipos de conteúdos, ajustando-se assim às necessidades individuais dos utilizadores .

Para complementar a experiência de aprendizagem, a plataforma integra um assistente virtual baseado num LLM alojado na nuvem, suportado por RAG. Este assistente responde às dúvidas dos utilizadores e executa comandos personalizados, como a solicitação de recomendações de conteúdos específicos. Apesar de não ter sido possível analisar todas as funcionalidades das plataformas pagas devido à necessidade de subscrição, nenhuma das plataformas portuguesas analisadas, das que permitiam acesso às funcionalidades, disponibilizava um assistente virtual (Khan Academy 2024; Mentes Brilhantes 2024; O Bichinho do Saber 2024).

No estado da arte realizado, a utilização de sistemas de recomendação demonstrou-se relevante no apoio à aprendizagem dos estudantes, permitindo uma orientação mais eficaz na seleção de conteúdos (Abdi, Khosravi, Sadiq e Gasevic 2020; Bhaskaran e Marappan 2023; Chatwattana, Yangthisarn e Tabubpha 2024). Além disso, ficou evidenciado que a

integração de LLMs com apoio de RAG pode ser uma mais-valia no contexto educativo, proporcionando respostas mais completas, fundamentadas e personalizadas (Dai et al. 2023; Yunfan Gao et al. 2024; Y. Zhang et al. 2023; W. X. Zhao et al. 2024).

O assistente virtual correspondeu às expectativas dos utilizadores no que respeita à qualidade da linguagem utilizada, à citação de fontes de informação e à completude das respostas. No entanto, não conseguiu captar plenamente o interesse dos participantes enquanto ferramenta de apoio ao estudo, o que indica margem para melhorias na sua integração e usabilidade.

O sistema de recomendação permitiu orientar os alunos na seleção dos conteúdos mais adequados às suas preferências e estilos de aprendizagem, como a escolha por vídeos em vez de documentos. Contudo, os resultados das métricas de avaliação evidenciaram a necessidade de aperfeiçoamento do modelo e de um conjunto de dados mais robusto e diversificado, de forma a aumentar a precisão e a relevância das recomendações.

De forma geral, a plataforma demonstrou potencial para ser utilizada como um recurso complementar no apoio ao estudo, mas revelou também oportunidades de melhoria, nomeadamente ao nível da interface e da experiência do utilizador. No que diz respeito à segurança, foram implementadas apenas medidas básicas, pelo que o reforço da segurança constitui uma prioridade para trabalhos futuros.

7.4 Limitações

Durante o desenvolvimento do projeto, surgiram diversas limitações que impactaram diretamente o progresso e os resultados obtidos.

Uma das principais limitações prendeu-se com a ausência de um conjunto de dados adequado. Não foi encontrado nenhum *dataset* que incluísse conteúdos alinhados com parte público-alvo do caso de estudo,os estudantes do curso profissional de Gestão e Programação de Sistemas Informáticos, e que estivesse também disponível em português. Por essa razão, foi necessário construir um conjunto de dados de raiz, de forma a garantir que estes correspondiam aos temas lecionados. Apesar de conter apenas 85 conteúdos, a recolha e organização dos dados revelou-se bastante morosa, cujo tamanho do conjunto de dados poderá ter influenciado negativamente os resultados obtidos nas métricas de avaliação do sistema de recomendação.

Outra limitação significativa relacionou-se com os recursos computacionais disponíveis. Devido à exigência de elevado poder de processamento por parte dos modelos de linguagem, foi necessário recorrer a um LLM baseada na nuvem.

Durante o desenvolvimento, verificaram-se dificuldades na implementação da funcionalidade de *login* e *logout* na página "Visão Geral", uma vez para essa funcionalidade funcionar corretamente necessitaria de utilizar o *manager* de sessão do *FastAPI*, que identifica o utilizador autenticado e restringe automaticamente o acesso a utilizadores não autenticados. Este comportamento contrariava o objetivo principal da página, que era estar acessível a todos os utilizadores, independentemente do seu estado de autenticação. Para contornar esta limitação, foi criada uma cópia da página "Visão Geral"exclusivamente para utilizadores autenticados, permitindo assim preservar a lógica de controlo de acesso sem comprometer a experiência de navegação entre os estados de *login* e *logout*.

Adicionalmente, a integração do*Groq* revelou-se particularmente desafiante, devido a incompatibilidades entre versões de bibliotecas, incluindo a integração do *LightRAG* inicialmente. Foi despendido um tempo considerável na tentativa de implementar esta funcionalidade

através da função original *groq_model_complet*, fornecida pela biblioteca `LightRAG`, sem sucesso. Para além disso, não foi encontrada qualquer documentação ou informação online sobre a criação manual da função, para além da sua invocação direta. Estas limitações conduziram à necessidade de recorrer à IA. Este processo contou com o apoio da ferramenta *GitHub Copilot*, que auxiliou na construção da função. Além disso, o LLM utilizado (llama-3.3-70b-versatile), embora gratuito, apresentava limitações de uso com um limite máximo de 131.072 *tokens* por sessão.

O grafo de conhecimento gerado pelo *LightRAG* apresentou nós isolados, indicando limitações na capacidade de estabelecer conexões entre entidades. Esta limitação deve-se às características do LLM de dimensão média utilizado, que necessitaria de um modelo de maior capacidade para superar as restrições impostas pelas limitações de hardware disponível.

A dimensão da amostra no caso de estudo foi também limitada. Apesar do envolvimento de 16 participantes, o objetivo consistia em abranger um número mais alargado de participantes para uma análise mais representativa. Diversas pessoas enquadradas no perfil pretendido foram contactadas, mas poucas se mostraram disponíveis para colaborar. Inicialmente, previa-se a publicação da plataforma online para facilitar o processo de testes, contudo, tal revelou-se inviável devido aos custos associados à integração do LLM, uma vez que as ferramentas disponíveis eram exclusivamente pagas e com preços elevados. Por este motivo, os testes foram realizados localmente, à exceção do preenchimento do questionário, o que permitiu garantir o anonimato dos participantes.

7.5 Trabalho Futuro

De modo a potenciar os resultados e a evolução das técnicas e ferramentas utilizadas no sistema de recomendação e no assistente virtual, recomenda-se, para trabalhos futuros, a expansão do estado da arte. É aconselhável analisar e comparar métodos de recomendação presentes noutros estudos a plataformas educacionais, incluindo internacionais, de forma a identificar abordagens inovadoras e boas práticas. No âmbito do assistente virtual, sugere-se a investigação de novos LLMs e estratégias para proporcionar uma experiência de utilização mais personalizada e envolvente. Relativamente às plataformas educacionais, revela-se pertinente analisar um maior número de soluções, tanto portuguesas como internacionais, para melhor posicionar a plataforma desenvolvida no contexto global.

No que respeita à evolução da plataforma, é recomendada a implementação de novas funcionalidades, com especial destaque para a introdução de elementos de gamificação, como níveis e conquistas (*achievements*), conforme sugerido por um dos participantes do caso de estudo. Esta abordagem poderá aumentar o envolvimento dos utilizadores e promover uma utilização mais frequente da plataforma.

Recomenda-se também o reforço do assistente virtual Tuga, nomeadamente através da expansão da sua base de conhecimento, permitindo fornecer respostas mais completas, precisas e contextualizadas.

No que concerne à validação de exercícios, sugere-se o desenvolvimento de métodos mais sofisticados para a validação das respostas dos utilizadores. Este aperfeiçoamento permitirá superar as limitações da atual metodologia, que combina comparação direta de dados armazenados com processamento via API do *Groq*, implementando a solução que demonstrar maior eficácia.

Outro aspeto fundamental consiste no alargamento e diversificação do conjunto de dados, de modo a permitir uma avaliação mais rigorosa do sistema de recomendação e a melhorar os resultados das métricas de desempenho.

Adicionalmente, sugere-se o envolvimento de entidades de ensino, como professores e educadores, desenvolvendo funcionalidades específicas para estes utilizadores, nomeadamente a possibilidade de sugerirem conteúdos educacionais aos seus alunos através da plataforma.

A segurança da plataforma deverá ser reforçada, nomeadamente com a introdução de mecanismos como a autenticação de dois fatores (*2FA*) e outras medidas que garantam a privacidade e confiança dos utilizadores.

Recomenda-se ainda a melhoria do design da interface, tornando-a mais atrativa e acessível, bem como a inclusão de um tema escuro para proporcionar uma melhor experiência de utilização.

Por fim, considera-se importante a realização de novos testes, preferencialmente de maior duração e com uma amostra mais alargada, para avaliar o impacto da plataforma no percurso escolar dos alunos e na sua motivação para a aprendizagem.

Bibliografia

Abdi, Solmaz, Hassan Khosravi, Shazia Sadiq e Dragan Gasevic (2020). «Complementing educational recommender systems with open learner models». Em: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. LAK '20. Frankfurt, Germany: Association for Computing Machinery, pp. 360–365. isbn: 9781450377126. doi: 10.1145/3375462.3375520. url: https://doi.org/10.1145/3375462.3375520.

Abdi, Solmaz, Hassan Khosravi, Shazia Sadiq e Dragan Gašević (2019). «A Multivariate ELO-based Learner Model for Adaptive Educational Systems». Em: *Proceedings of the Educational Data Mining Conference*, pp. 462–467.

Aementes Brilhantes (2024). Acesso em: 11 de novembro de 2024. url: https://www.aementesbrilhantes.com.pt/servicos-de-apoio.

Alotaibi, Naif (2024). *Image Generator [GPT personalizado]*. https://chat.openai.com/gpts. Acedido em: 22 de maio de 2025.

Amazon Web Services (2024a). *AWS SageMaker: Treine e Implante Modelos de Machine Learning*. https://aws.amazon.com/pt/sagemaker/. Acesso em: 28 dez. 2024.

– (2024b). *What is Retrieval-Augmented Generation?* Acesso em: 28 dez. 2024. url: https://aws.amazon.com/what-is/retrieval-augmented-generation/.

Areal Editores (1984). *Areal Editores - Manuais e Recursos Educativos*. https://www.arealeditores.pt/. Acesso em: 15 de janeiro de 2025.

Baek, Seungyun, Seunghwan Lee e Junhee Seok (2024). «Strategic Insights in Korean-English Translation: Cost, Latency, and Quality Assessed through Large Language Model». Em: *2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 551–553. doi: 10.1109/ICUFN61752.2024.10625181.

Baevski, A. e M. Auli (2018). «Adaptive input representations for neural language modeling». Em: *arXiv preprint arXiv:1809.10853*.

Bagunaid, Wala, Naveen Chilamkurti e Prakash Veeraraghavan (2022). «AISAR: Artificial Intelligence-Based Student Assessment and Recommendation System for E-Learning in Big Data». Em: *Sustainability* 14.17. issn: 2071-1050. doi: 10.3390/su141710551. url: https://www.mdpi.com/2071-1050/14/17/10551.

Bang, Jehyeon et al. (2024). *vTrain: A Simulation Framework for Evaluating Cost-effective and Compute-optimal Large Language Model Training*. arXiv: 2312.12391 [cs.LG]. url: https://arxiv.org/abs/2312.12391.

Barkan, Oren et al. (2020). «Explainable Recommendations via Attentive Multi-Persona Collaborative Filtering». Em: RecSys '20. Virtual Event, Brazil: Association for Computing Machinery, pp. 468–473. isbn: 9781450375832. doi: 10.1145/3383313.3412226. url: https://doi.org/10.1145/3383313.3412226.

Bender, Emily M. et al. (2021). «On the dangers of stochastic parrots: Can language models be too big?» Em: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM.

Bhaskaran, S. e Raja Marappan (2023). «Design and analysis of an efficient machine learning based hybrid recommendation system with enhanced density-based spatial clustering for

digital e-learning applications». Em: *Complex Intell. Syst.* 9, pp. 3517–3533. doi: 10.1007/s40747-021-00509-4.

Bootdey (2018). *Card chat list - Bootstrap snippet*. Acesso: março de 2025. url: https://www.bootdey.com/snippets/view/card-chat-list#html.

Borgeaud, S. et al. (2022). «Improving language models by retrieving from trillions of tokens». Em: *International Conference on Machine Learning*. PMLR, pp. 2206–2240.

Bouihi, Bouchra e Mohamed Bahaj (ago. de 2019). «Ontology and Rule-Based Recommender System for E-learning Applications». Em: *International Journal of Emerging Technologies in Learning (iJET)* 14.15, pp. 4–13. doi: 10.3991/ijet.v14i15.10566. url: https://online-journals.org/index.php/i-jet/article/view/10566.

Bouraga, Sarah et al. (abr. de 2014). «Knowledge-Based Recommendation Systems: A Survey». Em: *International Journal of Intelligent Information Technologies* 10, pp. 1–19. doi: 10.4018/ijiit.2014040101.

Breese, John, David Heckerman e Carl Kadie (jan. de 2013). «Empirical Analysis of Predictive Algorithm for Collaborative Filtering». Em: *UAI*.

Brown, T. et al. (2020). «Language models are few-shot learners». Em: *Advances in Neural Information Processing Systems* 33, pp. 1877–1901.

Bui, Tuan et al. (2024). «Cross-Data Knowledge Graph Construction for LLM-enabled Educational Question-Answering System: A Case Study at HCMUT». Em: *Proceedings of the 1st ACM Workshop on AI-Powered Q&A Systems for Multimedia*. AIQAM '24. Phuket, Thailand: Association for Computing Machinery, pp. 36–43. isbn: 9798400705472. doi: 10.1145/3643479.3662055. url: https://doi.org/10.1145/3643479.3662055.

Burke, Robin (mai. de 2000). «Knowledge-Based Recommender Systems». Em: *Encyclopedia of library and information systems* 69.

Çano, Erion e Maurizio Morisio (2017). «Hybrid Recommender Systems: A Systematic Literature Review». Em: *Intelligent Data Analysis* 21.6, pp. 1487–1524. issn: 1088-467X.

Cao, Yang et al. (2025). «Explainable Session-Based Recommendation via Path Reasoning». Em: *IEEE Transactions on Knowledge and Data Engineering* 37.1, pp. 278–290. doi: 10.1109/TKDE.2024.3486326.

Carta, Salvatore et al. (2023). *Iterative Zero-Shot LLM Prompting for Knowledge Graph Construction*. arXiv: 2307.01128 [cs.CL]. url: https://arxiv.org/abs/2307.01128.

Chaccour, Christina et al. (2024). «Telecom's Artificial General Intelligence (AGI) Vision: Beyond the GenAI Frontier». Em: *IEEE Network* 38.5, pp. 21–28. doi: 10.1109/MNET.2024.3425594.

Chatwattana, Pinanta, Piyada Yangthisarn e Areeya Tabubpha (jun. de 2024). «The Educational Recommendation System with Artificial Intelligence Chatbot: A Case Study in Thailand». Em: *International Journal of Engineering Pedagogy (iJEP)* 14.5, pp. 51–64. doi: 10.3991/ijep.v14i5.48491. url: https://online-journals.org/index.php/i-jep/article/view/48491.

Chen, Xiaojun, Shengbin Jia e Yang Xiang (2020). «A review: Knowledge reasoning over knowledge graph». Em: *Expert Systems with Applications* 141, p. 112948. issn: 0957-4174. doi: https://doi.org/10.1016/j.eswa.2019.112948. url: https://www.sciencedirect.com/science/article/pii/S0957417419306669.

Chiang, Jeffery (2021). *7 Types of Hybrid Recommendation System*. Medium. Acesso em dezembro de 2024. url: https://medium.com/analytics-vidhya/7-types-of-hybrid-recommendation-system-3e4f78266ad8.

Chockalingam, Annamalai et al. (2023). *Large Language Model eBooks*. Acesso em: 21 jan. 2025. url: https://resources.nvidia.com/en-us-large-language-model-ebooks.

cnmoro (2024). *Qwen2.5-0.5B-Portuguese-v1 - Modelo personalizado em português*. Acesso: março de 2025. url: https://ollama.com/cnmoro/Qwen2.5-0.5B-Portuguese-v1.

Dai, W. et al. (2023). «Can large language models provide feedback to students? A case study on ChatGPT». Em: *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*. IEEE, pp. 323–325.

Dam, Sumit Kumar et al. (2024). «A Complete Survey on LLM-based AI Chatbots». Em: arXiv: 2406.16937 [cs.CL]. url: https://arxiv.org/abs/2406.16937.

Das, B. C., M. H. Amini e Y. Wu (2024). «Security and privacy challenges of large language models: A survey». Em: *arXiv preprint arXiv:2402.00888*.

Data Science Lab@HKU (2024). *LightRAG: Lightweight Retrieval-Augmented Generation*. Acesso em: 10 de março de 2025. url: https://github.com/HKUDS/LightRAG.

Design.com (s.d.). *Design.com*. Acesso em: abril de 2025. url: https://www.design.com.

Diário da Repúbica (1986). *Lei n.º 46/86 - Lei de Bases do Sistema Educativo*. Acesso em: 18 de janeiro de 2025. url: https://diariodarepublica.pt/dr/legislacao-consolidada/lei/1986-34444975.

Diário de Notícias (2024). *Educação: 80 milhões para garantir acesso à Internet de qualidade nas escolas*. Acessado em: 3 de janeiro de 2025. url: https://www.dn.pt/8400543320/educacao-80-milhoes-para-garantir-acesso-a-internet-de-qualidade-nas-escolas/.

Direção-Geral da Educação (2011). *Missão e Atribuições*. https://www.dge.mec.pt/missao-e-atribuicoes. Acesso em: 15 de janeiro de 2025.

Direção-Geral da Educação (DGE) (2019). *Atividades de Apoio às Escolas*. Acesso em: 13 de janeiro de 2025. url: https://apoioescolas.dge.mec.pt/Atividades.

Direção-Geral do Ensino Superior (2024). *Informação Geral para Instituições de Ensino Superior*. Acesso em: 11 de novembro de 2024. url: https://www.dges.gov.pt/pt/pagina/informacao-geral-institucionais?plid=593.

Do, Minh-Phung, Dung Nguyen e Loc Nguyen (ago. de 2010). «Model-based approach for Collaborative Filtering». Em: *The 6th International Conference on Information Technology for Education (IT@EDU2010)*.

Escola Mágica (2024). Acesso em: 10 de outubro de 2024. url: https://www.escolamagica.pt/.

Escola Virtual (2024). Acesso em: 10 de outubro de 2024. url: https://www.escolavirtual.pt/.

Eurydice – Comissão Europeia (2024). *Portugal - Sistema Nacional de Educação*. Acesso em: 11 de novembro de 2024. url: https://eurydice.eacea.ec.europa.eu/pt-pt/national-education-systems/portugal/portugal.

Falkner, A., A. Felfernig e A. Haag (2011). «Recommendation technologies for configurable products». Em: *AI Magazine* 32, pp. 99–108. doi: 10.1609/aimag.v32i3.2369.

Felfernig, A., G. Friedrich et al. (2006). «An integrated environment for the development of knowledge-based recommender applications». Em: *International Journal of Electronic Commerce* 11, pp. 11–34. doi: 10.2753/JEC1086-4415110201.

Felfernig, A., K. Isak et al. (2007). «The VITA financial services sales support environment». Em: *Proceedings of the 19th National Conference on Innovative Applications of Artificial Intelligence - Volume 2, IAAI'07*. AAAI Press, pp. 1692–1699.

Felfernig, A., V. Le e S. Lubos (2023). «Conjunctive query based constraint solving for feature model configuration». Em: *Proceedings of the 12th Conference on Information Technology and Its Applications (CITA 2023)*. Vol. 734. Lecture Notes in Networks and Systems. Springer, pp. 357–367.

Gao, Y. et al. (2023). «Chat-Rec: Towards Interactive and Explainable LLMs-Augmented Recommender System». Em: *arXiv preprint arXiv:2303.14524*.

Gao, Yunfan et al. (2024). *Retrieval-Augmented Generation for Large Language Models: A Survey*. arXiv: 2312.10997 [cs.CL]. url: https://arxiv.org/abs/2312.10997.

Geng, S. et al. (2022). «Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5)». Em: *Proceedings of the 16th ACM Conference on Recommender Systems*, pp. 299–315.

GitHub, OpenAI (2021). *GitHub Copilot*. Acedido em 24 de maio de 2025. url: https://github.com/features/copilot.

Gonçalves, José (2007). *Ponte da Arrábida - Porto*. Acesso: abril de 2025. url: %5Curl%7Bhttps://commons.wikimedia.org/wiki/File:Ponte_da_Arrabida_-_Porto.JPG%7D.

Groq (2024a). *Groq Cloud*. Acesso em: março de 2025. url: https://console.groq.com/home.

– (2024b). *Groq Platform — Model Documentation*. Acesso: março de 2025. url: https://console.groq.com/docs/models.

Hadi, Muhammad Usman et al. (jul. de 2023). *Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects*. doi: 10.36227/techrxiv.23589741.

Hornik, K., M. Stinchcombe e H. White (1989). «Multilayer feedforward networks are universal approximators». Em: *Neural Networks* 2.5, pp. 359–366.

Hosseinzadeh Aghdam, Mehdi et al. (2015). «Adapting Recommendations to Contextual Changes Using Hierarchical Hidden Markov Models». Em: *Proceedings of the 9th ACM Conference on Recommender Systems*. RecSys '15. Vienna, Austria: Association for Computing Machinery, pp. 241–244. isbn: 9781450336925. doi: 10.1145/2792838.2799684. url: https://doi.org/10.1145/2792838.2799684.

Instituto de Avaliação Educativa (IAVE) (2024a). *Estudo Internacional PISA – IAVE*. Acesso em: 22 de novembro de 2024. url: https://iave.pt/estudo-internacional/pisa/.

– (2024b). *Provas de Aferição – Ensino Básico*. Acesso em: 11 de novembro de 2024. url: https://iave.pt/provas-e-exames/provas-e-exames/provas-de-afericao-eb/.

Instituto Politécnico do Porto (2020). *Regulamento do Código de Boas Práticas e de Conduta do Instituto Politécnico do Porto*. Acesso em: 18 de dezembro de 2024. url: https://www.iscap.ipp.pt/regulamentos/CodigoboaspraticasedecondutaIPP.pdf.

Intayoad, Wacharawan, Till Becker e Punnarumol Temdee (mai. de 2017). «Social Context-Aware Recommendation for Personalized Online Learning». Em: *Wireless Personal Communications* 97.1, pp. 163–179. doi: 10.1007/s11277-017-4499-2.

Ishak, Iskandar et al. (jan. de 2023). «Deep Learning-Based Recommendation System: Systematic Review and Classification». Em: *IEEE Access* PP, pp. 1–1. doi: 10.1109/ACCESS.2023.3323353.

Jannach, Dietmar, Massimo Quadrana e Paolo Cremonesi (2022). «Session-based Recommender Systems». Em: *Recommender Systems Handbook*. Ed. por Francesco Ricci et al. 3rd. Springer, pp. 301–335.

Javed, U. et al. (mar. de 2021). «A Review of Content-Based and Context-Based Recommendation Systems». Em: *International Journal of Emerging Technologies in Learning (iJET)* 16.03, pp. 274–306. doi: 10.3991/ijet.v16i03.18851. url: https://doi.org/10.3991/ijet.v16i03.18851.

Javed, Umair et al. (fev. de 2021). «A Review of Content-Based and Context-Based Recommendation Systems». Em: *International Journal of Emerging Technologies in Learning (iJET)* 16. doi: 10.3991/ijet.v16i03.18851.

Júnior, R. D. Torres (2004). «Combining Collaborative and Content-based Filtering to Recommend Research Paper». Tese de mestrado. Porto Alegre: Universidade Federal do Rio Grande do Sul.

Kang, Wang-Cheng et al. (2023). «Do LLMs Understand User Preferences? Evaluating LLMs On User Rating Prediction». Em: arXiv: 2305.06474 [cs.IR]. url: https://arxiv.org/abs/2305.06474.

Khan, A. e A. Hoffmann (2003). «Building a case-based diet recommendation system without a knowledge engineer». Em: *Artificial Intelligence in Medicine* 27, pp. 155–179. doi: 10.1016/S0933-3657(02)00113-6.

Khan Academy (2024). Acesso em: 2 de outubro de 2024. url: https://pt-pt.khanacademy.org/.

Khosravi, Hassan, Kirsty Kitto e Williams Joseph (2019). «RiPPLE: A Crowdsourced Adaptive Platform for Recommendation of Learning Activities». Em: *Journal of Learning Analytics* 6.3, pp. 91–105.

Kim, Sein et al. (2024). «Large Language Models meet Collaborative Filtering: An Efficient All-round LLM-based Recommender System». Em: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. KDD '24. Barcelona, Spain: Association for Computing Machinery, pp. 1395–1406. isbn: 9798400704901. doi: 10.1145/3637528.3671931. url: https://doi.org/10.1145/3637528.3671931.

Kudo, T. (2018). «Subword regularization: Improving neural network translation models with multiple subword candidates». Em: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 66–75.

Lewis, Patrick et al. (2020). «Retrieval-augmented generation for knowledge-intensive nlp tasks». Em: *Advances in neural information processing systems*. Vol. 33, pp. 9459–9474.

LeYa (2008). *LeYa - Grupo Editorial*. https://www.leya.com/.

Leya Aula Digital (2024). Acesso em: 10 de outubro de 2024. url: https://auladigital.leya.com/.

LinkedIn Learning (2018). *Como se Comunicar com Confiança*. https://www.linkedin.com/learning/como-se-comunicar-com-confianca.

Lorenzi, F. e F. Ricci (2005). «Case-based recommender systems: a unifying view». Em: *Intelligent Techniques for Web Personalization*. Ed. por B. Mobasher e S. Anand. Vol. 3169. Lecture Notes in Computer Science. Berlin; Heidelberg: Springer, pp. 89–113.

Lozic, E. e B. Stular (2023). «ChatGPT v Bard v Bing v Claude 2 v Aria v Human-Expert. How good are AI chatbots at scientific writing? (ver. 23Q3)». Em: *arXiv preprint arXiv:2309.08636*.

Lucas, Joel P. et al. (2013). «A hybrid recommendation approach for a tourism system». Em: *Expert Systems with Applications* 40.9, pp. 3532–3550. issn: 0957-4174. doi: https://doi.org/10.1016/j.eswa.2012.12.061. url: https://www.sciencedirect.com/science/article/pii/S0957417412013024.

Martins, Constantino e Catarina Figueiredo (2024). *Aulas de Sistemas de Recomendação*. Mestrado de Engenharia Informática. Instituto Superior de Engenharia do Porto.

Matematica.pt (2024). Acesso em: 10 de outubro de 2024. url: https://www.matematica.pt/index.php.

Mbaye, Baba (2018). «Recommender System: Collaborative Filtering of E-Learning Resources». Em: *Proceedings of the International Conference e-Learning 2018*. Laboratory of ELLIADD, University of Franche-Comté. França, pp. 213–217. isbn: 978-989-8533-78-4.

Mentes Brilhantes (2024). Acesso em: 10 de outubro de 2024. url: https://mentesbrilhantespt.weebly.com/.

Meta (2024a). *LLaMA 3.2*. Acesso: março de 2025. url: https://ollama.com/library/llama3.2.

– (2024b). *Llama3.2*. Acesso em: 05 de março de 2025. url: https://ollama.com/library/llama3.2.

Meta AI] (2023). *LLaMA 3.3*. Acesso: março de 2025. url: https://console.groq.com/docs/model/llama-3.3-70b-versatile.

Microsoft (2024). *Azure Content Moderator Overview*. https://learn.microsoft.com/en-us/azure/ai-services/content-moderator/overview. Acesso em: 28 dez. 2024.

Narayanan, Deepak et al. (2021). *Scaling Language Model Training to a Trillion Parameters Using Megatron*. Accesso em: 24 de fevereiro de 2025. url: https://developer.nvidia.com/blog/scaling-language-model-training-to-a-trillion-parameters-using-megatron/.

Naveed, Humza et al. (2024). *A Comprehensive Overview of Large Language Models*. arXiv: 2307.06435 [cs.CL]. url: https://arxiv.org/abs/2307.06435.

Netflix (2015). *Educational Screenings of Netflix Documentaries*. Acesso em: 13 de janeiro de 2025. url: https://help.netflix.com/pt-pt/node/57695.

Nomic AI (2024). *nomic-embed-text*. Acesso: março de 2025. url: https://ollama.com/library/nomic-embed-text.

O Bichinho do Saber (2024). Acesso em: 10 de outubro de 2024. url: https://www.obichinhodosaber.com/.

Obeid, Charbel et al. (2018). «Ontology-based Recommender System in Higher Education». Em: *Companion Proceedings of the The Web Conference 2018*. WWW '18. Lyon, France: International World Wide Web Conferences Steering Committee, pp. 1031–1034. isbn: 9781450356404. doi: 10.1145/3184558.3191533. url: https://doi.org/10.1145/3184558.3191533.

Ollama (2023). *Ollama*. https://ollama.com/. Acesso em: março de 2025.

OpenArt (2022). *OpenArt: AI Art Generator*. https://openart.ai/home. Acesso: abril de 2025.

Ortiz, S. (2023). *What is Bing Chat? Here's everything you need to know*. [Accessed: Sept. 24, 2024]. url: https://www.zdnet.com/article/what-is-thenew-bing-heres-everything-you-need-to-know/.

Panagiotakis, Costas e Harris Papadakis (2022). «Session-Based Recommendation by combining Probabilistic Models and LSTM». Em: RecSysChallenge '22. Seattle, WA, USA: Association for Computing Machinery, pp. 39–44. isbn: 9781450398565. doi: 10.1145/3556702.3556846. url: https://doi.org/10.1145/3556702.3556846.

Pennock, David M. et al. (2013). *Collaborative Filtering by Personality Diagnosis: A Hybrid Memory- and Model-Based Approach*. arXiv: 1301.3885 [cs.IR]. url: https://arxiv.org/abs/1301.3885.

Phalle, Ms e Prof Bhushan (mar. de 2024). «Content Based Filtering And Collaborative Filtering: A Comparative Study». Em: *Journal of Advanced Zoology* 45, pp. 96–100. doi: 10.53555/jaz.v45iS4.4158.

Porto Editora (1994). *Porto Editora - Manuais Escolares e Recursos Educativos*. https://www.portoeditora.pt/. Acesso em: 15 de janeiro de 2025.

Rabahallah, Kahina, Latifa Mahdaoui e Faical Azouaou (mar. de 2018). «MOOCs Recommender System using Ontology and Memory-based Collaborative Filtering». Em: pp. 635–641. doi: 10.5220/0006786006350641.

Rahutomo, Faisal, Tohru Kitasuka e Masayu Aritsugi (jan. de 2014). «Semantic Cosine Similarity». Em: url: https://www.researchgate.net/publication/262525676_Semantic_Cosine_Similarity.

Ram, O. et al. (2023). «In-context retrieval-augmented language models». Em: *arXiv preprint arXiv:2302.00083*.

Rodrigues, Jairson B., Germano C. Vasconcelos e Paulo R. M. Maciel (2020). «Time and Cost Prediction Models for Language Classification Over a Large Corpus on Spark». Em: *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1702–1709. doi: 10.1109/SSCI47803.2020.9308299.

Sala do Futuro (2024). Acesso em: 11 de dezembro de 2024. url: https://saladofuturo.pt/.

Saravia, E. (2022). *Prompt Engineering Guide*. https://github.com/dairai/Prompt-Engineering-Guide. Acessado pelo autor do artigo "A Comprehensive Overview of Large Language Models"em: Dezembro 2022.

Schober, P., C. Boer e L. A. Schwarte (mai. de 2018). «Correlation Coefficients: Appropriate Use and Interpretation». Em: *Anesthesia & Analgesia* 126.5, pp. 1763–1768. doi: 10.1213/ANE.0000000000002864.

Sennrich, R., B. Haddow e A. Birch (2016). «Neural machine translation of rare words with subword units». Em: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1715–1725.

Shinde, Subhash K. e Uday V. Kulkarni (ago. de 2011). «Hybrid Personalized Recommender System Using Fast K-medoids Clustering Algorithm». Em: *Journal of Advances in Information Technology* 2.3, pp. 152–158. url: https://www.jait.us/uploadfile/2014/1222/20141222035318591.pdf.

Shishehchi, Saman et al. (fev. de 2012). «Ontological Approach in Knowledge Based Recommender System to Develop the Quality of E-learning System». Em: *Australian Journal of Basic and Applied Sciences* 6, pp. 115–123.

Smith, Shaden et al. (2022). *Using DeepSpeed and Megatron to Train Megatron-Turing NLG 530B, A Large-Scale Generative Language Model*. arXiv: 2201.11990 [cs.CL]. url: https://arxiv.org/abs/2201.11990.

Soared, Maria (2021). *Escola de Português iei, Jogos Educativos*. Acessado em: 3 de janeiro de 2025. url: https://iei.pt/jogos-iei/.

Soh, Harold et al. (2017). «Deep Sequential Recommendation for Personalized Adaptive User Interfaces». Em: *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. IUI '17. Limassol, Cyprus: Association for Computing Machinery, pp. 589–593. isbn: 9781450343480. doi: 10.1145/3025171.3025207. url: https://doi.org/10.1145/3025171.3025207.

Superprof (2024). Acesso em: 11 de novembro de 2024. url: https://www.superprof.pt/.

SuperTmatik (s.d.). *SuperTmatik - Jogo de Cartas Educativo*. https://www.supertmatik.net/app/cartas/. Acesso em: 15 de janeiro de 2025.

Tänzer, M., S. Ruder e M. Rei (2021). «Memorisation versus generalisation in pretrained language models». Em: *arXiv preprint arXiv:2105.00828*.

Tavakol, Maryam e Ulf Brefeld (2014). «Factored MDPs for detecting topics of user sessions». Em: *Proceedings of the 8th ACM Conference on Recommender Systems*. RecSys '14. Foster City, Silicon Valley, California, USA: Association for Computing Machinery, pp. 33–40. isbn: 9781450326681. doi: 10.1145/2645710.2645739. url: https://doi.org/10.1145/2645710.2645739.

Touvron, H. et al. (2023). «Llama 2: Open foundation and fine-tuned chat models». Em: *arXiv preprint arXiv:2307.09288*.

Towle, B. e C. Quinn (2000). «Knowledge based recommender systems using explicit user models». Em: *AAAI Technical Report WS-00-04*. AAAI Press, pp. 74–77.

Tutor.pt (2024). Acesso em: 11 de novembro de 2024. url: https://www.tutor.pt/servicos/.

Uta, Mathias et al. (2024). «Knowledge-based recommender systems: overview and research directions». Em: *Frontiers in Big Data* 7. issn: 2624-909X. doi: 10.3389/fdata.2024.1304439. url: https://www.frontiersin.org/journals/big-data/articles/10.3389/fdata.2024.1304439.

Valmeekam, K. et al. (2022). «Large language models still can't plan (a benchmark for LLMs on planning and reasoning about change)». Em: *arXiv preprint arXiv:2206.10498*.

Vaswani, A. et al. (2017). «Attention is all you need». Em: *Advances in Neural Information Processing Systems 30*.

W3schools (s.d.[a]). *CSS Horizontal Navigation Bar*. Acesso: março de 2025. url: https://www.w3schools.com/css/css_navbar_horizontal.asp.

– (s.d.[b]). *How TO - Login Form*. Acesso: março de 2025. url: https://www.w3schools.com/howto/howto_css_login_form.asp.

Wang, S. et al. (2023). «Can ChatGPT write a good boolean query for systematic review literature search?» Em: *arXiv preprint arXiv:2302.03495*.

Webson, A. e E. Pavlick (2021). «Do prompt-based models really understand the meaning of their prompts?» Em: *arXiv preprint arXiv:2109.01247*.

Webster, J. J. e C. Kit (1992). «Tokenization as the initial phase in NLP». Em: *COLING 1992 Volume 4: The 14th International Conference on Computational Linguistics*.

Wikipedia (2023). *Bard (chatbot)*. [Accessed: Sept. 25, 2024]. url: https://en.wikipedia.org/wiki/Bard_(chatbot).

Wikipedia contributors (1987). *Era Uma Vez... a Vida — Wikipédia, a enciclopédia livre*. Acesso em: 13 de janeiro de 2025. url: https://pt.wikipedia.org/wiki/Era_Uma_Vez..._a_Vida.

Yang, Z. (2023). *Chinese tech giant Baidu just released its answer to ChatGPT*. [Accessed: Sept. 25, 2024]. url: https://www.technologyreview.com/2023/03/16/1069919/baidu-erniebot-chatgpt-launch/.

Yuan, Zhihang et al. (2024). *LLM Inference Unveiled: Survey and Roofline Model Insights*. arXiv: 2402.16363 [cs.CL]. url: https://arxiv.org/abs/2402.16363.

Zhang, C. et al. (2021). «Understanding deep learning (still) requires rethinking generalization». Em: *Communications of the ACM* 64.3.

Zhang, Peiyuan et al. (2023). *TinyLlama: An Open-Source Small Language Model*. Acedido em abril de 2025. url: https://ollama.com/library/tinyllama.

– (2024). *TinyLlama: An Open-Source Small Language Model*. arXiv: 2401.02385 [cs.CL].

Zhang, Y. et al. (2023). «Siren's song in the ai ocean: A survey on hallucination in large language models». Em: *arXiv preprint arXiv:2309.01219*.

Zhao, Qian et al. (2019). «From preference into decision making: modeling user interactions in recommender systems». Em: *Proceedings of the 13th ACM Conference on Recommender Systems*. RecSys '19. Copenhagen, Denmark: Association for Computing Machinery, pp. 29–33. isbn: 9781450362436. doi: 10.1145/3298689.3347065. url: https://doi.org/10.1145/3298689.3347065.

Zhao, Wayne Xin et al. (2024). «A Survey of Large Language Models». Em: *arXiv preprint*. url: https://github.com/RUCAIBox/LLMSurvey.

Zhao, Zihuai et al. (2024). *Recommender Systems in the Era of Large Language Models (LLMs)*. arXiv: 2307.02046 [cs.IR]. url: https://arxiv.org/abs/2307.02046.

Zhiyuli et al. (2023). «BookGPT: A General Framework for Book Recommendation Empowered by Large Language Model». Em: *arXiv preprint arXiv:2305.15673*.

Zhu, Yue et al. (2024). «Optimizing GPU Multiplexing for Efficient and Cost-Effective Access to Diverse Large Language Models in GPU Clusters». Em: *2024 32nd International Conference on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS)*, pp. 1–8. doi: 10.1109/MASCOTS64422.2024.10786517.

Apêndice A

Apêndice A - Planeamento do Projeto

No planeamento do projeto, são definidos conteúdos essenciais antes do seu desenvolvimento. Destaca-se o escopo do projeto as competências necessárias para determinar as áreas a serem trabalhadas, e o cronograma, que estabelece as diferentes etapas e estimativas de datas de execução.

A.1 Riscos do Projeto

É essencial identificar, analisar e avaliar os riscos que podem surgir durante o desenvolvimento do projeto, de forma a planear as ações necessárias para a sua mitigação ou resolução, caso estes se concretizem.

ID	Descrição	Causa	Efeito	Risk Owner	Probabilidade (1-5)	Impacto (1-5)	PI Score	Resultado Esperado, Nenhuma Ação	Tipo de Resposta ao Risco	Descrição da Resposta
	Descrição do Risco	Causa do Risco	Efeito no Projeto	Nome da pessoa que monitoriza o risco	Estimativa aproximada obtida pelo grupo da probabilidade de tal ocorrer	Rough estimate of how significant the impact of this risk	Estimativa aproximada de quão significativo é o impacto deste risco	O que acontecerá se o risco se tornar um problema e não forem tomadas medidas	Decisão tomada pelo grupo sobre como responder a este risco	Como sabe que está na hora de colocar a resposta em ação
1	Dados insuficientes	Não ser possível encontrar quantidade necessária	Os testes apresentarem resultados pouco satisfatórios	Constantino Martins	1	3	3	Estes problemas podem influenciar negativamente a avaliação do sistema	Mitigar	Quando afetar a qualidade dos testes e/ou reparar-se que há necessidade de obter mais dados
2	Escolha incorreta dos métodos de recomendação	Estudos insuficientes	Más recomendações	Constantino Martins	1	5	5	Resultados insatisfatórios no desenvolvimento do projeto	Mitigar	Quando as previsões resultarem em resultados sem sentido
3	Atrasos no desenvolvimento dos entregáveis	Contratempos no desenvolvimento e complexidade	Menos tempo para executar testes e documentação	Constantino Martins	1	4	4	O atraso pode impedir a conclusão de algumas tarefas.	Aceitar	Quando uma tarefa está a demorar mais que o previsto

Figura A.1: Riscos do Projeto

A.2 Competências

Procedeu-se a uma autoavaliação e heteroavaliação das competências, o que permitiu identificar os pontos fortes, bem como as áreas de competência que necessitam de desenvolvimento.

A.2.1 Competências mais Desenvolvidas

- **Colaboração e Trabalho em Equipa**: Manifesta-se uma boa capacidade de trabalho em equipa, ouvindo as opiniões dos demais e pela disposição para auxiliar os colegas em áreas onde se possui maior conhecimento. Esta competência é fundamental para promover um ambiente colaborativo e produtivo entre os membros da equipa.

- **Ética e Responsabilidade Social**: Atua-se de forma ética e socialmente responsável, respeitando os princípios éticos de cada entidade. Assume-se o compromisso a contribuir para um ambiente profissional positivo e a tomar decisões que não prejudiquem os demais.
- **Gestão do Tempo**: Consegue-se gerir de forma eficiente o tempo disponível para a realização de uma tarefa, cumprindo os prazos estabelecidos, mesmo quando surgem contratempos. Priorizam-se as tarefas com base nos prazos e na sua importância, assegurando a realização eficiente das atividades mais urgentes e relevantes.

A.2.2 Competências a desenvolver

Durante a fase de desenvolvimento do projeto, foquei em melhorar as competências nas quais tenho apresentado maiores dificuldades, com o objetivo de as fortalecer. O plano de ação apresentado descreve as medidas a serem adotadas para o desenvolvimento de cada uma dessas competências.

- **Gestão de Stress e Ansiedade**: Com o objetivo de melhorar a gestão do stress e da ansiedade, implementei uma organização mais eficiente das tarefas diárias. Esta abordagem promoveu um equilíbrio saudável entre as responsabilidades e períodos de descanso.
- **Melhorar a Comunicação**: Com o intuito de melhorar a comunicação e a confiança ao comunicar, realizei o curso *Communicating with Confidence* no *LinkedIn Learning*, com o objetivo de comunicar com clareza e assertividade em diferentes contextos profissionais (LinkedIn Learning 2018). O certificado de conclusão encontra-se no *link*.

A.3 Âmbito do Projeto

O projeto é composto por diversos elementos essenciais, sendo crucial identificar os principais *stakeholders* envolvidos. Além disso, é fundamental evidenciar os benefícios que o projeto trará ao ser desenvolvido.

A.3.1 Principais Stakeholders

O projeto envolve diferentes *stakeholders* essenciais para o seu desenvolvimento, sendo os principais:

- **Estudante**: Principal responsável desde o planeamento até ao desenvolvimento e conclusão da dissertação.
- **Orientador**: Responsável por orientar o estudante, acompanhando o progresso do projeto, assegurando que as tarefas são cumpridas de acordo com o cronograma e que o trabalho está em conformidade com as exigências da instituição ISEP.
- **Co-orientador**: Apoia o autor na ausência do orientador, ajudando também no desenvolvimento da pesquisa e resolução de questões técnicas e científicas ao longo do processo de elaboração da dissertação.

A.3.2 Benefícios do Projeto

O projeto é desenvolvido tendo em consideração os diferentes estilos de aprendizagem dos alunos, disponibilizando conteúdos educacionais diversificados e adaptados às suas necessidades individuais.

A integração de um sistema de recomendações personalizadas visa reduzir o tempo despendido pelos alunos na procura de materiais que se adequem às suas preferências.

Adicionalmente, a implementação de um *chatbot* terá como objetivo apoiar os alunos no processo de aprendizagem, fornecendo respostas imediatas a dúvidas e complementando os conteúdos disponibilizados na plataforma.

Adicionalmente, a plataforma é de uso gratuito, considerando que muitas das principais plataformas educativas portuguesas analisadas no estudo do estado da arte são parcialmente gratuitas ou pagas.

A.4 Work Breakdown Structure(WBS)

O diagrama *Work Breakdown Structure (WBS)* A.2 ilustra as fases do desenvolvimento da dissertação, bem como os respetivos entregáveis e tarefas. Este instrumento foi fundamental para estruturar o conteúdo, permitindo uma visão clara das diferentes fases de entrega. Além disso, facilitou a preparação de tópicos para discussão durante as reuniões com os orientadores, assegurando um alinhamento eficaz do projeto.

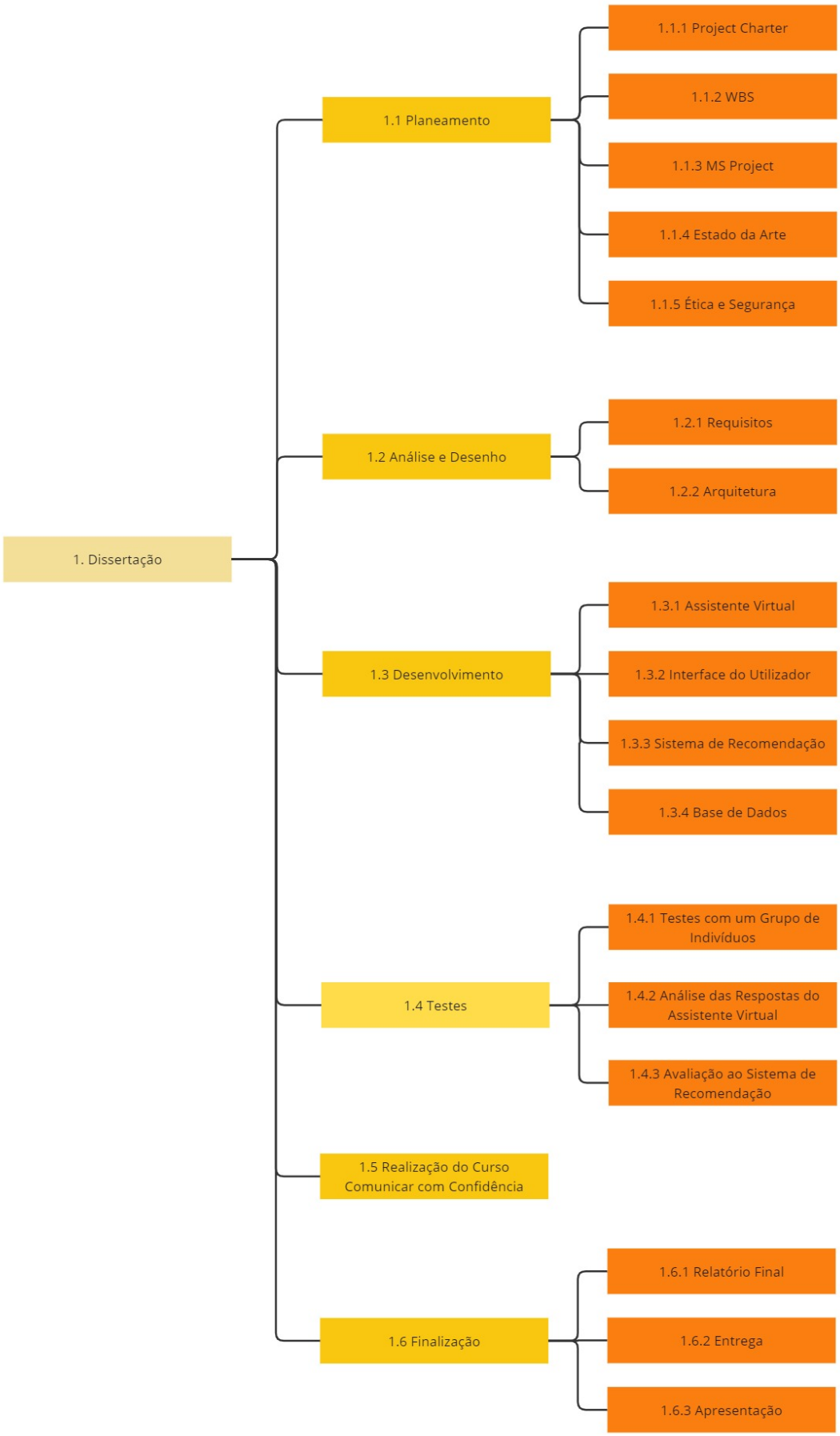

Figura A.2: WBS

A.5 Cronograma

O gráfico de *Gantt* descreve o cronograma do projeto, incluindo os níveis de WBS, os recursos necessários para cada tarefa e o progresso de execução das tarefas. A partir da fase de análise, são atribuídas datas estimadas, uma vez que até à conclusão do planeamento do projeto, não é possível determinar com precisão as datas reais de execução.

	Task Mode	WBS	Task Name	Nível WBS	Duration	Start	Finish	Predece	Resource Names	% Complete
1		**1**	**Dissertação**	**Projeto**	**206 days?**	**Tue 01/10/24**	**Wed 16/07/25**			**99%**
2		**1.1**	**Planeamento**	**Fase**	**70 days**	**Tue 01/10/24**	**Sat 04/01/25**			**100%**
3		1.1.1	Project Charter	Entregável	29 days	Mon 28/10/24	Thu 05/12/24		Estudante	100%
4		1.1.2	WBS	Entregável	3 days	Wed 06/11/24	Fri 08/11/24		Estudante	100%
5		1.1.3	MS Project	Entregável	30 days	Mon 11/11/24	Fri 20/12/24	4	Estudante	100%
6		1.1.4	Estado da Arte	Entregável	65 days	Tue 01/10/24	Mon 30/12/24		Estudante	100%
7		1.1.5	Ética e Segurança	Entregável	1 day	Tue 01/10/24	Tue 01/10/24		Estudante	100%
8		**1.2**	**Análise e Desenho**	**Fase**	**56 days?**	**Mon 20/01/25**	**Mon 07/04/25**	**6**		**100%**
9		1.2.1	Requisitos	Entregável	56 days?	Wed 15/01/25	Wed 02/04/25	6	Estudante;Orienta	100%
10		1.2.2	Arquitetura	Entregável	56 days?	Wed 15/01/25	Wed 02/04/25	6	Estudante;Orienta	100%
11		**1.3**	**Desenvolvimento**	**Fase**	**86 days?**	**Mon 03/02/25**	**Mon 02/06/25**	**2**		**100%**
12		1.3.1	Assistente Virtual	Entregável	86 days?	Tue 21/01/25	Tue 20/05/25			100%
13		1.3.2	Interface do Utilizador	Entregável	80 days?	Mon 03/02/25	Fri 23/05/25			100%
14		1.3.3	Sistema de Recomendação	Entregável	86 days?	Mon 03/02/25	Mon 02/06/25			100%
15		1.3.4	Base de Dados	Entregável	80 days?	Mon 03/02/25	Fri 23/05/25			100%
16		**1.4**	**Testes**	**Entregável**	**15 days?**	**Mon 05/05/25**	**Fri 23/05/25**			**100%**
17		1.4.1	Testes com um grupo de índividuos	Entregável	15 days?	Mon 05/05/25	Fri 23/05/25			100%
18		1.4.2	Análise das Respostas do Assistente Virtual	Entregável	1 day?	Wed 21/05/25	Wed 21/05/25			100%
19		1.4.3	Avaliação ao Sistema de Recomendação	Entregável	10 days?	Fri 09/05/25	Thu 22/05/25			100%
20		1.5	Realização do Curso Comunicar com Confidência	Tarefa	1 day	Mon 03/03/25	Mon 03/03/25			100%
21		**1.6**	**Finalização**	**Fase**	**91 days?**	**Mon 10/03/25**	**Mon 14/07/25**			**99%**
22		1.6.1	Relatório Final	Entregável	80 days?	Mon 10/03/25	Sat 28/06/25		Estudante	100%
23		1.6.2	Entrega	Entregável	1 day	Mon 02/06/25	Sun 29/06/25	22	Estudante	100%
24		1.6.3	Apresentação	Entregável	1 day?	Mon 14/07/25	Mon 14/07/25		Estudante	0%
25		**1.7**	**Milestones**	**Milestones**	**166 days?**	**Tue 26/11/24**	**Wed 16/07/25**			**0%**
26		1.7.1	Plano Aprovado	Milestones	0 days	Tue 26/11/24	Tue 26/11/24			100%
27		1.7.2	Validação do Relatório	Milestones	0 days?	Mon 16/06/25	Mon 16/06/25			0%
28		1.7.3	Projeto Concluído	Milestones	0 days?	Wed 16/07/25	Wed 16/07/25			0%

Figura A.3: Gráfico de *Gantt*, Tarefas

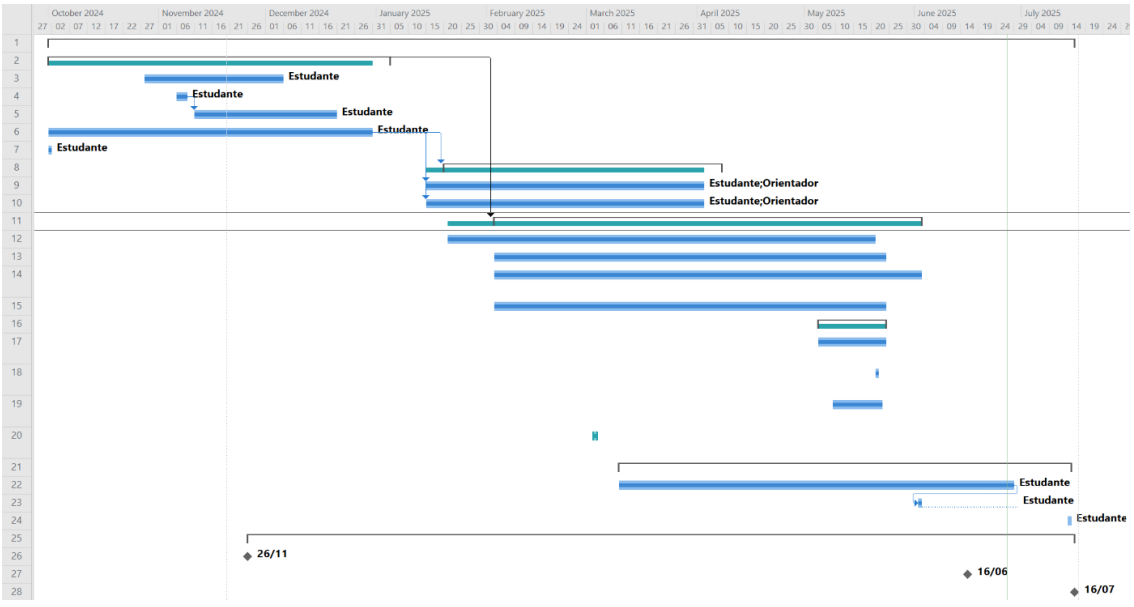

Figura A.4: Gráfico de *Gantt*, Cronograma

A.6 Monitorização e Controlo

Durante o planeamento do projeto, foram realizadas reuniões online recorrentes com o orientador, utilizando a plataforma *Microsoft Teams*. Numa fase inicial, o principal objetivo consistiu em compreender mais em detalhe o projeto a ser desenvolvido e identificar potenciais áreas para pesquisa no estado da arte.

A partir de novembro, as reuniões passaram a ter uma periodicidade quinzenal, algumas delas realizadas com a presença do coorientador. Estas sessões permitiram acompanhar o progresso do trabalho, garantir o cumprimento do planeamento estabelecido, esclarecer eventuais dúvidas e sugestões por parte dos orientadores.

Ambos os orientadores têm acesso ao documento da dissertação, que está a ser desenvolvido em *LaTeX* através do editor online *Overleaf*.

Apêndice B

Apêndice B - Requisitos Funcionais Complementares

Neste apêndice encontram-se especificados os requisitos funcionais complementares do sistema, conforme referenciado no Capítulo 3. Estes requisitos representam uma decomposição detalhada dos requisitos principais.

RF1.1	**Inserir dados pessoais no registo**
Descrição	O sistema deve solicitar no momento do registo, os dados do utilizador, nome, nome de utilizador (*username*), palavra-passe e confirmação da palavra-passe.
Prioridade	Alta

Tabela B.1: RF1.1 – Inserir dados pessoais no registo

RF1.2	**Validação do Nome**
Descrição	Durante o registo, o sistema deve validar se o nome do utilizador não contém números e possui entre 2 e 50 caracteres.
Prioridade	Alta

Tabela B.2: RF1.2 – Validação do Nome

RF1.3	**Validação do Nome de Utilizador**
Descrição	Durante o processo de registo, o sistema deve verificar se o nome de utilizador (*username*) introduzido já se encontra registado. Caso não esteja, deve validar se o mesmo contém entre 3 e 20 caracteres e se é composto apenas por letras, números e o *underscore* (_).
Prioridade	Alta

Tabela B.3: RF1.3 – Validação do Nome de Utilizador

RF1.4	**Validação da Palavra-passe**
Descrição	O sistema deve validar se a palavra-passe introduzida durante o registo possui entre 6 e 20 caracteres, apresentando uma mensagem de erro caso este critério não seja cumprido.
Prioridade	Alta

Tabela B.4: RF1.4 – Validação da Palavra-passe

RF1.5	**Verificação da Confirmação da Palavra-passe**
Descrição	O sistema deve assegurar que a confirmação da palavra-passe coincide com a palavra-passe introduzida, apresentando uma mensagem de erro em caso de divergência.
Prioridade	Alta

Tabela B.5: RF1.5 – Verificação da Confirmação da Palavra-passe

RF1.6	**Consulta e Aceitação dos Termos e Condições**
Descrição	O sistema deve disponibilizar ao utilizador os Termos e Condições para consulta e exigir a sua aceitação para que o registo possa ser concluído com sucesso.
Prioridade	Alta

Tabela B.6: RF1.6 – Consulta e Aceitação dos Termos e Condições

RF1.7	**Preferência de Recomendação de Conteúdo Variado**
Descrição	O sistema deve permitir que o utilizador, opcionalmente, indique se pretende receber recomendações adicionais e diversificadas, incluindo conteúdo ainda não explorado, para além das recomendações personalizadas habituais.
Prioridade	Baixa

Tabela B.7: RF1.7 – Preferência de Recomendação de Conteúdo Variado

RF2.1	**Verificar se o Utilizador Existe**
Descrição	Durante o processo de autenticação, o sistema deve validar se *username* do utilizador existe, apresentando mensagem de erro em caso não encontrar.
Prioridade	Alta

Tabela B.8: RF2.1 – Verificar se o Utilizador Existe

RF2.2	**Verificar a Palavra-passe do Utilizador**
Descrição	Caso o *username* exista, o sistema deve validar a palavra-passe introduzida, comparando-a com o valor armazenado de forma segura através de *hashing*.
Prioridade	Alta

Tabela B.9: RF2.2 – Verificar a Palavra-passe do Utilizador

RF5.1	**Filtragem de Conteúdo**
Descrição	O sistema deve permitir ao utilizador filtrar os conteúdos por categoria, já separados por tipo, apresentando uma lista com os respetivos nomes das categorias, da qual poderá selecionar uma ou mais categorias para aplicar o filtro.
Prioridade	Baixa

Tabela B.10: RF5.1 – Filtragem de Conteúdo

RF5.2	**Fluxo de Navegação dos Conteúdos**
Descrição	O sistema deve abrir o conteúdo educativo solicitado pelo utilizador num novo separador, garantindo um fluxo de utilização contínuo e fluido na plataforma.
Prioridade	Baixa

Tabela B.11: RF5.2 – Fluxo de Navegação dos Conteúdos

RF5.3	**Gostar de Conteúdos Educacionais**
Descrição	O sistema deve permitir ao utilizador gostar dos conteúdos educacionais.
Prioridade	Alta

Tabela B.12: RF5.3 – Gostar de Conteúdos Educacionais

RF5.4	**Visualizar Conteúdos Gostados**
Descrição	O sistema deve permitir ao utilizador ver se já gostou ou não do conteúdo.
Prioridade	Alta

Tabela B.13: RF5.4 – Visualizar Conteúdos Gostados

RF5.5	**Retirar Gosto de Conteúdos**
Descrição	O sistema deve permitir ao utilizador remover o gosto de um conteúdo previamente assinalado, caso se tenha enganado a clicar ou tenha deixado de ter interesse no mesmo.
Prioridade	Média

Tabela B.14: RF5.5 – Retirar Gosto de Conteúdos

RF5.6	**Ordenar Conteúdos Gostados**
Descrição	O sistema deve apresentar em primeiro na lista de conteúdos, aqueles que o utilizador já deu "gosto", de forma a facilitar o seu acesso aos conteúdos gostados.
Prioridade	Baixa

Tabela B.15: RF5.6 – Visualizar Conteúdos Curtidos

RF6.1	**Validação da Resposta**
Descrição	O sistema deve validar a resposta submetida pelo utilizador, indicando se está correta ou incorreta, e apresentar sempre a resposta correta, independentemente do resultado da validação.
Prioridade	Média

Tabela B.16: RF6.1 – Validação da Resposta

RF6.2	**Resposta Incorreta**
Descrição	O sistema deve apresentar sempre que a resposta estiver incorreta, uma explicação que auxilie o utilizador a compreender o erro cometido, promovendo a melhoria da aprendizagem.
Prioridade	Média

Tabela B.17: RF6.2 – Resposta Incorreta

RF6.3	**Controlo da Validação**
Descrição	O sistema deve desativar temporariamente o botão de validação da resposta, mantendo-o desativado até que a resposta seja apresentada ao utilizador. Esta medida previne erros no processamento e apresentação das respostas ao utilizador.
Prioridade	Média

Tabela B.18: RF6.3 – Controlo da Validação

RF7.1	**Recomendação Baseada em Conteúdo**
Descrição	O sistema deve utilizar recomendação baseada em conteúdo para identificar conteúdo semelhante aos que o utilizador já gostou.
Prioridade	Alta

Tabela B.19: RF7.1 – Recomendação Baseada em Conteúdo

RF7.2	**Página de Recomendações**
Descrição	O sistema deve permitir a visualização de uma página com as recomendações, de até 10 conteúdos recomendados que o utilizador ainda não tenha gostado, aumentando para 15 caso o utilizador tenha selecionado a opção de recomendação variada durante o registo, desde que esses conteúdos ainda não tenham sido curtidos. Além disso, o sistema deve indicar o tipo de conteúdo (vídeo, documento, livro, exercícios) agrupado por tipo e possibilitar ao utilizador curtir o conteúdo apresentado.
Prioridade	Alta

Tabela B.20: RF7.2 – Visualizar Recomendações

RF7.3	**Exclusão de Conteúdos Previamente Gostados**
Descrição	O sistema deve garantir que conteúdos previamente gostados pelo utilizador não sejam novamente recomendados, com o objetivo de apresentar sempre ao utilizador conteúdos ainda não descobertos
Prioridade	Alta

Tabela B.21: RF7.3 – Exclusão de Conteúdos Previamente Gostados

RF7.4	**Recomendação Variada de Conteúdo**
Descrição	Caso o utilizador tenha optado por visualizar recomendações variadas, o sistema deve apresentar 10 conteúdos recomendados com base nos conteúdos que o utilizador gosta e 5 conteúdos aleatórios.
Prioridade	Alta

Tabela B.22: RF7.4 – Recomendação Variada de Conteúdo

RF7.5	**Recomendações para Utilizadores sem Gostos**
Descrição	Caso o utilizador ainda não tenha gostado de nenhum conteúdo, o sistema deve apresentar os 10 conteúdos preferidos dos utilizadores, no caso de ter seleciono visualizar recomendações variadas exibe aleatoriamente conteúdo.
Prioridade	Alta

Tabela B.23: RF7.5 – Recomendações para Utilizadores sem Gostos

RF7.6	**Avaliações das Recomendações**
Descrição	O sistema deve avaliar e armazenar os resultados das recomendações com base em métricas específicas, para efeitos de testes e análise de desempenho.
Prioridade	Média

Tabela B.24: RF7.6 – Avaliações das Recomendações

RF8.1	**Resposta ao Utilizador**
Descrição	O assistente virtual Tuga deve primeiramente, procurar a informação localmente para responder ao utilizador. Apenas no caso de não encontrar dados relevantes, deverá recorrer a fontes externas para responder ao utilizador.
Prioridade	Alta

Tabela B.25: RF8.1 – Resposta ao Utilizador

RF8.2	**Texto Multi Linha**
Descrição	O sistema deve suportar a inserção de texto em múltiplas linhas por parte do utilizador no *chat* do assistente virtual Tuga.
Prioridade	Média

Tabela B.26: RF8.2 – Texto Multi Linha

RF8.3	**Mensagem Introdutória**
Descrição	O assistente virtual Tuga deve apresentar, no início da interação, uma mensagem introdutória com o objetivo de sugerir ao utilizador tipos de perguntas que pode realizar e indicar o comando de ajuda.
Prioridade	Alta

Tabela B.27: RF8.3 – Mensagem Introdutória

RF8.4	**Comando de Ajuda**
Descrição	O assistente virtual Tuga deve disponibilizar um comando de ajuda que apresente ao utilizador a lista dos comandos personalizados disponíveis, facilitando a sua utilização.
Prioridade	Alta

Tabela B.28: RF8.4 – Comando de Ajuda

RF8.5	**Reconhecimento Textual**
Descrição	O assistente virtual Tuga deve impedir a associação de palavras reservadas a comandos quando estas forem utilizadas em frases, assegurando a distinção clara entre a invocação de comandos e o uso dessas palavras em frases.
Prioridade	Alta

Tabela B.29: RF8.5 – Reconhecimento Textual

RF8.6	**Citação de Fontes**
Descrição	O assistente virtual Tuga deve apresentar as fontes de informação utilizadas na geração das respostas, garantindo a transparência dos conteúdos fornecidos.
Prioridade	Baixa

Tabela B.30: RF8.6 – Citação de Fontes

RF8.7	**Conformidade Linguística do Tuga**
Descrição	O assistente virtual Tuga deve interagir com os utilizadores em Português de Portugal, respeitando as normas linguísticas e culturais.
Prioridade	Média

Tabela B.31: RF8.7 – Conformidade Linguística do Tuga

RF8.8	**Controlo de Envio de Mensagens**
Descrição	O sistema deve desativar temporariamente o botão de envio de mensagens no *chat* do assistente virtual Tuga imediatamente após o utilizador submeter uma mensagem, mantendo-o desativado até que a resposta seja apresentada. Esta medida previne erros no processamento e apresentação das respostas ao utilizador.
Prioridade	Alta

Tabela B.32: RF8.8 – Controlo de Envio de Mensagens