

**Aplicação de Modelos de Atenção em Campanhas de Marketing Digital e Realidade Virtual**

ISRC - Interdisciplinary Studies Research Center

2023 / 2024

1200690 Pedro Miguel Gaspar Alves Santos



**Aplicação de Modelos de Atenção em Campanhas de Marketing Digital e Realidade Virtual**

ISRC - Interdisciplinary Studies Research Center

2023 / 2024

1200690 Pedro Miguel Gaspar Alves Santos



Licenciatura em Engenharia Informática

Junho de 2024

Orientador ISEP: **Professora Ana Maria Madureira**

Supervisor: **Doutor** **Ivo Pereira**

*«Dedico este trabalho aos meus filhos pelo amor incondicional e à minha esposa pelo amor e apoio constantes ao longo destes anos.*»

Agradecimentos

É com grande satisfação que chego ao final de mais uma Licenciatura em Engenharia, culminando esta com a conclusão do projeto final do curso, a etapa final da Licenciatura em Engenharia Informática.

Estou imensamente grato à minha esposa, Sara, por todo o apoio que me deu, não me deixando nunca desistir, mesmo quando a conclusão do curso teve impacto na vida pessoal. Igualmente estou muito grato aos filhos, Rita e Vasco, por todo o apoio que me deram, às vezes mesmo sem darem conta disso. A esta família mais próxima, sem o vosso apoio e sacrifício, nunca teria conseguido concluir este curso.

É importante também reconhecer a instituição onde o projeto foi desenvolvido, juntamente com os professores e alunos das turmas a que pertenci, uma vez que fomentaram e facilitaram a aprendizagem, mantendo sempre um elevado nível de partilha de aprendizagens e conhecimentos.

Um dos maiores agradecimentos vai para o meu principal grupo de trabalho que desempenhou um papel crucial sem o qual nada disto teria sido possível. Enfrentámos juntos diversos desafios, mas conseguimos ultrapassar todos os obstáculos. Assim, Leandro Pinto, António Sousa e Jorge Santos, aqui fica o meu mais sincero agradecimento.

Finalmente, tenho também de agradecer à minha orientadora, Prof. Dr. Ana Madureira por me ter ajudado a concluir este projeto com sucesso e por me guiar no sentido de realizar um projeto que muito apreciei realizar, mas também que abre a porta futuramente a outros projetos na área de Inteligência Artificial.

Falta concluir

Resumo

O resumo do relatório (que só deve ser escrito após o texto principal do relatório estar completo) é uma apresentação abreviada e precisa do trabalho, sem acrescento de interpretação ou crítica, escrita de forma impessoal, podendo ter, por exemplo, as seguintes três partes:

1. Um parágrafo inicial de introdução do contexto e do problema/objetivo do trabalho.
2. Resumo dos aspetos mais importantes do trabalho descrito no presente relatório, que por sua vez documenta abordagem adotada e sistematiza os aspetos relevantes do trabalho realizado durante o estágio. Deve mencionar tudo o que foi feito, por isso deve concentrar-se no que é realmente importante e ajudar o leitor a decidir se quer ou não consultar o restante do relatório.
3. Um parágrafo final com as conclusões do trabalho realizado.

**Palavras-chave (Tema):** Incluir 3 a 6 palavras/expressões chave que caraterizem o projeto do ponto de vista de tema/área de intervenção.

**Palavras-chave (Tecnologias):** Incluir 3 a 6 palavras/expressões chave que caraterizem o projeto do ponto de vista de tecnologias utilizadas.

**(O Resumo só deve ocupar 1 página, cerca de 20 linhas**)

Índice

[1 Introdução 1](#_Toc168437717)

[1.1 Enquadramento/Contexto 1](#_Toc168437718)

[1.2 Descrição do Problema 2](#_Toc168437719)

[1.2.1 Objetivos 2](#_Toc168437720)

[1.2.2 Abordagem 3](#_Toc168437721)

[1.2.3 Contributos 3](#_Toc168437722)

[1.2.4 Planeamento do trabalho 4](#_Toc168437723)

[1.3 Estrutura do relatório 4](#_Toc168437724)

[2 Estado da arte 6](#_Toc168437725)

[2.1 Marketing Digital 6](#_Toc168437726)

[2.2 Machine Learning 7](#_Toc168437727)

[2.3 Modelos de Atenção 10](#_Toc168437728)

[2.4 Trabalhos relacionados 12](#_Toc168437729)

[2.5 Tecnologias existentes 14](#_Toc168437730)

[3 Análise e desenho da solução 17](#_Toc168437731)

[3.1 Domínio do problema 17](#_Toc168437732)

[3.2 Requisitos funcionais e não funcionais 19](#_Toc168437733)

[3.2.1 Requisitos Funcionais 19](#_Toc168437734)

[3.2.2 Requisitos Não Funcionais do sistema 20](#_Toc168437735)

[3.3 Desenho 24](#_Toc168437736)

[3.3.1 Arquitetura do sistema 24](#_Toc168437737)

[3.3.2 Solução proposta 26](#_Toc168437738)

[3.3.3 Solução alternativa 26](#_Toc168437739)

[3.3.4 Conclusões 27](#_Toc168437740)

[4 Implementação da Solução 29](#_Toc168437741)

[4.1 Descrição da implementação 29](#_Toc168437742)

[4.1.1. Web Service 29](#_Toc168437743)

[4.1.2. Modelos de Atenção 31](#_Toc168437744)

[4.1.3. User Interface 32](#_Toc168437745)

[4.1.4. Base de dados 34](#_Toc168437746)

[4.1.5. Controlo de versões 36](#_Toc168437747)

[4.1.6. CI/CD 37](#_Toc168437748)

[4.1.7. Controlo de tarefas 37](#_Toc168437749)

[4.2 Testes 38](#_Toc168437750)

[4.3 Avaliação da solução 42](#_Toc168437751)

[5 Conclusões 43](#_Toc168437752)

[5.1 Objetivos concretizados 43](#_Toc168437753)

[5.2 Limitações e trabalho futuro 43](#_Toc168437754)

[5.3 Apreciação final 43](#_Toc168437755)

[Referências 45](#_Toc168437756)

[Anexo A Conteúdo em anexos 47](#_Toc168437757)

[Referências 48](#_Toc168437758)

Índice de Figuras

Figura 1: Os 3 tipos de Aprendizagem Automática [9] 8

Figura 2: Os 3 tipos de Aprendizagem Automática e Aprendizagem Profunda [12] 10

Figura 3: Ranking do TIOBE [18] Figura 4: Ranking do PYPL [19] 14

Figura 5: Modelo FURPS+ [23] 17

Figura 6: Modelo de Domínio 18

Figura 7: Diagrama de Casos de Uso 19

Figura 8: Diagrama de componentes do sistema 25

Figura 9: Diagrama de implantação do sistema 25

Figura 10: Diagrama de componentes alternativo do sistema 26

Figura 11: Intercetador de pedidos web para inclusão de token de acesso 33

Figura 13: Classe User Serializer 34

Figura 14: Tabela da base de dados da Classe User Serializer 34

Figura 15: Classe Note Serializer 35

Figura 16: Tabela da Classe Note Serializer 35

Figura 17: Testes unitários de domínio 37

Figura 18: Teste unitário de criação de utilizador 37

Figura 19: Teste unitário de criação de nota 38

Figura 20: Teste de validação de criação do mapa de saliências 39

Figura 21: Resultado de Testes e2e 40

Figura 22: Testes e2e relativo à criação de um mapa de saliências de uma imagem sem criação de nota associada 40

**(O documento não pode exceder as 70 páginas, contando desde a primeira página da Introdução até à última página da Conclusão)**

Índice de Tabelas

Tabela 1: Planeamento do projeto 4

Tabela 2: Requisitos de Usabilidade 20

Tabela 3: Requisitos de Confiabilidade 21

Tabela 4: Requisitos de Desempenho 21

Tabela 5: Requisitos de Suportabilidade 22

Tabela 6: Restrições de Desenho 22

Tabela 7: Restrições de Implementação 23

Tabela 8: Restrições de Interface 23

Notação e Glossário

|  |  |
| --- | --- |
| **CPU** | Processador |
| **PHP** | Linguagem de programação... |
| **SO** | Sistema operativo |
| **WS** | Serviço Web (do inglês *Web service*) |

«Esta secção (glossário de termos) apresenta os conceitos ordenados alfabeticamente, símbolos matemáticos e acrónimos **utilizados** no corpo do relatório.»

1. Introdução

O projeto final ou estágio curricular é um componente fundamental no processo da formação académica. De fato, no final de uma licenciatura, é crucial que seja dado ao estudante um problema que permita aplicar os conhecimentos adquiridos durante o percurso académico.

Este relatório descreve o projeto desenvolvido no âmbito Unidade Curricular Projeto/Estágio (PESTI) com o objetivo de criar para o Interdisciplinary Studies Research Center um software para a Aplicação de Modelos de Atenção em Campanhas de Marketing Digital e Realidade Virtual. Este tema será detalhado e contextualizado ao longo deste documento, incluindo os objetivos específicos, a abordagem, o planeamento e o desenvolvimento do projeto.

Este capítulo da introdução oferece assim uma visão geral do projeto e do seu contexto.

* 1. Enquadramento/Contexto

Este relatório realizou-se no âmbito da Unidade Curricular Projeto/Estágio (PESTI) do 3º ano do curso de Licenciatura em Engenharia Informática do Instituto Superior de Engenharia do Porto (ISEP).

Devido à minha frequência em regime pós-laboral da licenciatura e, também, devido ao meu interesse em áreas de investigação e desenvolvimento, particularmente nas áreas de interligência artificial, foi decidido realizar um projeto no Interdisciplinary Studies Research Center (ISRC).

O ISRC foca-se na integração de investigação científica com a área da educação em engenharia, preenchendo um vazio na investigação e desenvolvimento atual, tirando partido das sinergias interdisciplinares existentes na Instituição. Todos os avanços tecnológicos mundiais das últimas décadas implicam que todas as pessoas e, consequentemente, os seus países, estejam interligados numa sociedade cada vez mais global e multicultural. Este fenómeno da multiculturalidade faz parte do meu dia-a-dia, uma vez que atualmente desempenho as minhas atuais funções numa empresa americana multinacional.

* 1. Descrição do Problema

Um dos projetos atualmente a ser desenvolvido no ISRC é o PHYNHANCAI que propõe o desenvolvimento de modelos de Inteligência Artificial Multimodal e a exploração de abordagens de Realidade Virtual e Realidade Aumentada para melhorar as estratégicas de Marketing Omnicanal. [1]

Neste contexto do marketing digital e da realidade virtual, é essencial capturar e manter a atenção do público-alvo para garantir o sucesso de campanhas e experiências. No entanto, a análise manual de imagens e vídeos para identificar áreas de interesse pode ser morosa e suscetível a erros.

Para além destes problemas, não existem ferramentas específicas onde se apliquem modelos de atenção o que dificulta ainda mais a integração dessas técnicas nas estratégicas de marketing e em experiências de realidade virtual. Assim, surge a necessidade de um software que permita uma aplicação simples e rápida de modelos de atenção em imagens e vídeos, possibilitando a geração de mapas de saliência visual para orientar o design e a produção de conteúdo mais interessante para o utilizador.

* + 1. Objetivos

Este projeto tem como objetivo principal o desenvolvimento de um software para Aplicação de Modelos de Atenção que seja robusto e intuitivo, permitindo a aplicação de modelos de atenção em imagens de campanhas de email marketing e vídeos de realidade virtual. O software deverá ter uma utilização fácil e eficiente, proporcionando resultados precisos e informativos.

Subjacente ao desenvolvimento do software está uma validação experimental da eficiência do mesmo. Assim, irão realizar-se uma série de testes experimentais para avaliar a eficácia e precisão do software desenvolvido, que incluirão comparações entre os mapas de saliência gerados pelo software e avaliações humanas, garantindo a qualidade e relevância dos resultados para as estratégias de marketing e experiências de realidade virtual.

Um último objetivo é a implementação de funcionalidades de pré-processamento de dados para garantir a qualidade e consistência dos inputs fornecidos ao software. Isso incluirá técnicas de normalização, redimensionamento e tratamento de ruídos, preparando os dados para uma análise mais precisa e eficaz pelos modelos de atenção.

* + 1. Abordagem

Para o desenvolvimento de qualquer solução, a primeira ação a realizar é a leitura e compreensão de artigos publicados que tenham relação com o problema a resolver. Devido à exigência do cliente do uso da ferramenta Python, paralelamente a esta ação de recolha de informação publicada, foi aprendida esta tecnologia amplamente usada no mercado. [2], [3]

Posteriormente, realizou-se a implementação do software propriamente dito, recorrendo a um processo ágil, iterativo e incremental. Em todas as reuniões com o cliente foi apresentada uma versão cada vez mais completa do software, de modo a obter feedback o mais cedo possível nas várias fases do desenvolvimento.

Para acompanhar todo o processo de desenvolvimento, após ser definido e acordado com o cliente e com a orientadora o planeamento geral do trabalho, foi usada a ferramenta *Jira*. Esta ferramenta define uma hierarquia para diferentes tipos de tarefas, sendo que neste projeto apenas foi necessário recorrer ao nível 4, o nível da fase de projeto, que inclui tarefas to tipo: *New Feature / Task / Bug*.

* + 1. Contributos

O desenvolvimento deste projeto e consequente documentação permite, tal como referido nos objetivos, através do uso de uma Aplicação de Modelos de Atenção, obter um feedback praticamente instantâneo acerca da qualidade de imagens ou vídeos de publicidade, tendo em conta o objetivo do utilizador.

Deste modo, uma empresa criadora de conteúdos publicitários poderá facilmente, com o recurso a esta aplicação desenvolvida, determinar o sucesso ou não desses mesmos conteúdos antes mesmo de os lançar para o mercado. Isto fará com que exista um benefício claro da qualidade e, consequentemente, do retorno financeiro dos conteúdos publicitários gerados pelos utilizadores do software desenvolvido neste projeto.

* + 1. Planeamento do trabalho

O planeamento do trabalho foi realizado tendo em conta as várias tarefas a desempenhar, priorizando inicialmente um conhecimento mais aprofundado sobre o tema e sobre as tecnologias a usar, passando posteriormente para a fase de desenvolvimento do software até à obtenção da solução final.

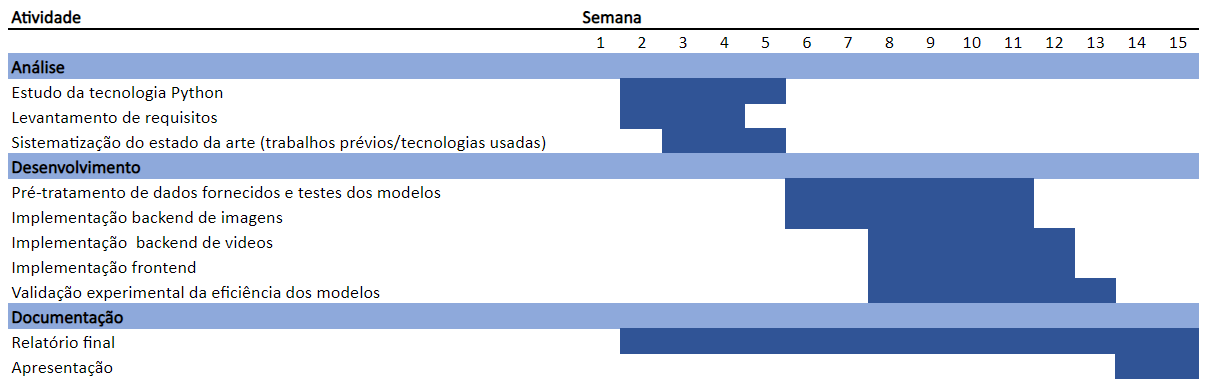


Tabela 1: Planeamento do projeto

* 1. Estrutura do relatório

O presente relatório está dividido em cinco capítulos, os quais se descrevem seguidamente.

No primeiro capítulo intitulado Introdução realiza-se uma contextualização do problema e explicitam-se os objetivos definidos. Posteriormente, apresenta-se a abordagem seguida e também o planeamento do projeto, evidenciando as várias fases do desenvolvimento.

No segundo capítulo, Estado da Arte, é mostrado o estado atual de conhecimento do problema, aborda trabalhos existentes e finalmente fornece uma visão geral das soluções existentes.

O terceiro capítulo, Análise e Desenho da solução, foca-se na análise do problema, explicando as diversas componentes e como o fluxo de informação é realizado. É também esquematizado o desenho da solução, dando resposta aos requisitos funcionais e não-funcionais.

No capítulo 4 é explicitada a Implementação da solução, seguindo o desenho da solução previsto e onde são avaliadas alternativas, vantagens e desvantagens recorrendo às métricas de validação.

Finalmente, no capítulo 5 são referidas as Conclusões, é combinando o conhecimento teórico com os resultados práticos, tendo em conta os objetivos propostos para o projeto. São também explicadas as limitações e futuras melhorias que podem ser realizadas.

1. Estado da arte

Neste capítulo vai-se explicitar a contextualização do tema e tecnologias estudadas ao longo do desenvolvimento deste estágio. O estudo teórico realizado permitiu ajudar e corroborar as decisões efetuadas na formulação da solução final desenvolvida. -> preciso melhorar esta parte

* 1. Marketing Digital

O Marketing Digital é hoje em dia uma ferramenta amplamente utilizada pelas empresas com a missão da continuidade do negócio empresarial. Através de um simples clique, coloca as empresas muito próximas do público-alvo, permitindo assim que as campanhas publicitárias sejam muito mais eficientes. Por outro lado, o Marketing Digital permite também um conhecimento mais aprofundado por parte das empresas do perfil de cada consumidor, pelo que as campanhas podem ser personalizadas e conseguir com isso maiores benefícios.

O conceito de Marketing de Atração, ou Inbound Marketing, foi inicialmente proposto por Brian Halligan em 2005, tendo sido usado como missão da sua empresa HubSpot, co-fundada por Dharmesh Shah. [4]

Este conceito baseia-se fundamentalmente em realizar ações em 4 níveis distintos por forma a conseguir aumentar vendas influenciando as decisões de compra dos utilizadores [5]:

* Atrair – usando diferentes técnicas de marketing, sejam marketing de conteúdo, difusão nas redes sociais, etc. com o objetivo único de aumento de tráfego no sítio web da empresa
* Converter – usando técnicas como conteúdos premium, webinars e outros, permite conseguir os dados dos utilizadores de modo a uma ser possivel uma posterior maior capacidade de personalização
* Vender – através de uma integração de uma equipa de vendas e uso de ferramentas de automatização de marketing ou de gestão de relação de clientes (CRM)
* Agradar – conseguir que os compradores repitam a compra e aumentar assim o vínculo com a empresa

Assim, num mundo cada vez mais virtual e mais informado, os utilizadores pretendem realizar compras de forma mais consciente e, por isso, as empresas, através do Marketing de Atração, adaptaram-se a essa nova realidade, aproveitando o maior conhecimento de cada cliente fornecendo campanhas de Marketing mais específicas e personalizadas.

* 1. Machine Learning

A Aprendizagem Automática (Machine Learning), é uma das formas de aplicação da Inteligência Artificial (IA), definida como um conjunto de “métodos computacionais que usam a experiência a fim de sofisticar a performance ou efetuar previsões confiáveis”. [6]

Este campo de estudo tem o objetivo de imitar cada vez melhor a aprendizagem humana, permitindo que as máquinas melhorem ao longo do tempo, tornando-se cada vez mais precisas nas previsões ou classificações. Existem variadas aplicações usadas no nosso dia-a-dia e algumas que a maioria das pessoas nem reconhece que usa a Aprendizagem Automática. Como exemplo, pode-se explicitar algumas funcionalidades das redes sociais, permitindo uma personalização da experiência do utilizador; os assistentes virtuais, como a Siri da Apple, ou a Alexa da Amazon recolhem dados de cada utilização e melhoram a resposta nas utilizações seguintes ou ainda motores de busca com recomendações que são usados em diversos sites de comércio online, uma vez que recolhem dados das pesquisas e compras efetuadas e depois recomendam produtos com base nessa recolha. [7]

Existem 3 tipos de Aprendizagem Automática, a Aprendizagem Supervisionada (Supervised Learning), a Aprendizagem Não Supervisionada (Unsupervised Learning) e a Aprendizagem por Reforço (Reinforcement Learning). [8] [9]

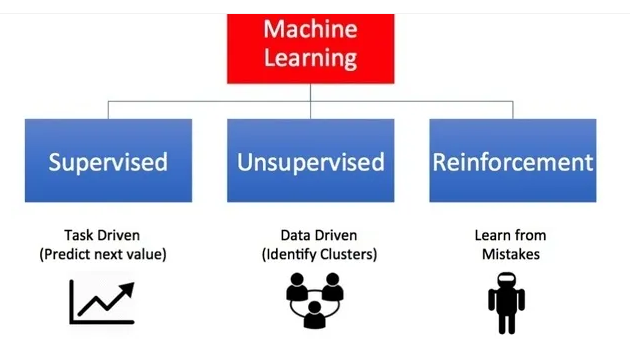


Figura 1: Os 3 tipos de Aprendizagem Automática [9]

* Aprendizagem Supervisionada

A Aprendizagem Supervisionada é definida pelo uso de dados rotulados para treinar algoritmos que classificam dados ou preveem resultados. De acordo com a maneira como os dados de entrada são inseridos no modelo, ele ajusta o peso de cada parâmetro até que o modelo seja ajustado corretamente, o que ocorre como parte do processo de validação cruzada. A Aprendizagem Supervisionada é a mais utilizada e ajuda organizações em todo o mundo a resolver uma variedade de problemas reais em grande escala e pode ser usada para criar modelos com elevada precisão. [10]

Este tipo de Aprendizagem usa um set de dados de treino para treinar o modelo, providenciando dados de entrada e respostas que se deveria obter. O algoritmo mede a sua eficácia e vai ajustando o erro até que tenha sido suficientemente minimizado. [10]

A Aprendizagem Supervisionada pode ser dividida em 2 tipos de problemas, problemas de classificação e problemas de regressão. Nos problemas de classificação é usado um algoritmo para assignar dados de teste a categorias específicas, sendo os modelos mais comuns as árvores de decisão, K-vizinhos mais próximos ou a Random Forest. Por outro lado, os problemas de regressão são usados para entender a relação entre variáveis dependentes e independentes, sendo os modelos de regressão mais comuns a regressão linear ou regressão polinomial. [10]

* Aprendizagem Não Supervisionada

A Aprendizagem Não Supervisionada usa algoritmos para analisar e agrupar conjuntos de dados não etiquetados. Estes algoritmos permitem descobrir padrões nos dados sem necessidade da interação com o ser humano, daí serem chamados de não supervisionados. [11]

Estes modelos de Aprendizagem Não Supervisionada são usados para 3 principais funções: agrupar, associar ou reduzir a dimensão dos dados:

- Agrupar é uma técnica para juntar dados não etiquetados baseando-se nas suas semelhanças ou diferenças. Por exemplo, K-vizinhos mais próximos agrupa pontos de dados semelhantes em grupos, em que K representa o tamanho do grupo ou granularidade. [11]

- Associação é outra técnica que usa diferentes regras para encontrar relações entre variáveis de um grupo de dados. Estes métodos são muito usados para motores de busca de recomendações. [11]

- Redução da dimensão é uma técnica usada quando o número de características (ou dimensões) de um grupo de dados é muito grande. Nestes casos, usa-se esta técnica reduzindo o número de dados de entrada para um valor aceitável preservando a integridade dos dados. [11]

* Aprendizagem Por Reforço

A Aprendizagem Por Reforço é uma técnica de Aprendizagem Automática que permite a um agente que aprenda com um ambiente interativo por tentativa e erro recebendo feedback das suas próprias ações e experiências. [9]

Embora a Aprendizagem Supervisionada e a Aprendizagem por Reforço usem um mapeamento entre dados de entrada e resultados obtidos, enquanto a Aprendizagem Supervisionada dá informação ao agente das ações a realizar para obter um resultado, no caso da Aprendizagem Por Reforço são dados recompensas ou punições como sinais de comportamento positivo ou negativo. [9]

Ao comparar a Aprendizagem Não Supervisionada e a Aprendizagem Por Reforço verifica-se que a grande diferença entre elas tem a ver com o objetivo. Enquanto que a Aprendizagem Não Supervisionada pretende descobrir semelhanças ou diferenças entre pontos de dados, na Aprendizagem Por Reforço o objetivo é encontrar ações que maximizem o valor de recompensa. [9]

* 1. Modelos de Atenção

A Aprendizagem Automática tem uma divisão transversal aos 3 tipos referidos acima que é chamada de Aprendizagem Profunda (Deep Learning).

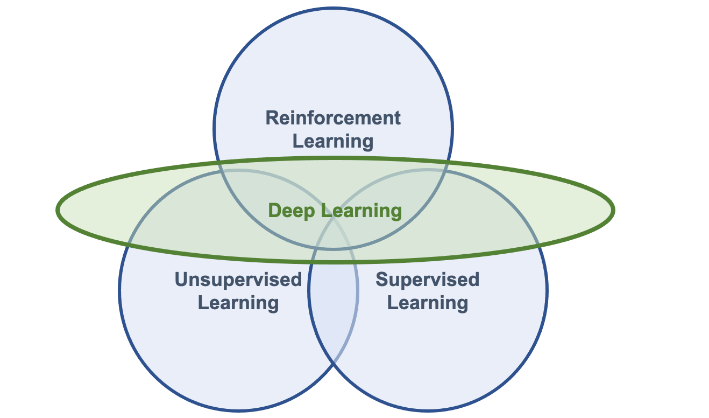


Figura 2: Os 3 tipos de Aprendizagem Automática e Aprendizagem Profunda [12]

Os Modelos de Atenção têm surgido como uma técnica poderosa da Aprendizagem Profunda, particularmente nos campos do Processamento da Linguagem Natural (Natural Language Processing – NLP) e Visão por Computador. Estes modelos foram desenhados para melhorar a performance das redes neuronais, usando a experiência adquirida da visão humana, permitindo um maior foco em partes específicas dos dados de entrada de modo a conseguirem obter um melhor resultado. [13]

Um Modelo de Atenção é um componente de uma rede neuronal que assigna um nível de importância, ou atenção, a diferentes partes dos dados de entrada. Os Modelos de Atenção podem ser integrados em vários tipos de redes neuronais, desde Redes Neuronais Recurrentes (Recurrent Neural Networks RNN), Redes Neuronais Convolucionais (Convolutional Neural Networks CNN) e também Modelos Transformacionais. [13]

Existem diferentes tipos de Modelos de Atenção que criam diferentes mapeamentos entre entradas e saídas, incluindo diferentes fontes, codificadores, descodificadores e pesos:

* Modelo de Atenção Global (ou Soft)

O Modelo de Atenção Global agrup entradas de todos os estados do codificador e do descodificador antes de avaliar o estado atual para determinar a saída. Este modelo utiliza cada etapa do codificador e cada etapa preliminar do descodificador para calcular os pesos de atenção ou pesos de alinhamento. Também multiplica cada etapa do codificador pelos pesos de alinhamento globais para determinar o valor de contexto a ser alimentado para as redes neurais recorrentes (RNN). Isso permite que o modelo encontre a saída do descodificador. [14]

* Modelo de Atenção Local

O Modelo de Atenção Local é semelhante ao modelo de atenção global, mas usa apenas algumas posições do codificador para determinar os pesos de alinhamento. O modelo calcula os pesos de alinhamento e o vetor de contexto usando a primeira posição de alinhamento única e uma seleção de palavras da fonte do codificador. O Modelo de Atenção Local é semelhante ao Modelo de Atenção Rígida, mas o Modelo de Atenção Rígida não é diferenciável na maioria dos pontos. Finalmente, o Modelo de Atenção Local combina aspetos de atenção rígida e suave. [14]

* Modelo de Própria Atenção

O Modelo de Própria Atenção foca-se em diferentes posições da mesma sequência de entrada. É possível utilizar os frameworks dos Modelos de Atenção Global e Local para criar este modelo. No entanto, este modelo inclui a mesma sequência de entrada em vez da sequência de saída alvo. É também um componente chave nos Modelos Transformacionais. [14]

A operação fundamental dum Modelo de Atenção envolve 3 componentes principais: consultas (queries), chaves (keys) e valores (values). Esses componentes são derivados dos dados de entrada e são usados para calcular resultados de atenção, que determinam quanto foco o modelo deve dar a cada parte dos dados. [13]

Num mecanismo de atenção típico, cada elemento da sequência de entrada é associado a uma chave e a um valor. Uma consulta, frequentemente relacionada ao estado atual do modelo, é comparada com todas as chaves usando uma função de compatibilidade, que pode ser um produto escalar simples ou uma rede neural mais complexa. O resultado dessa comparação é um conjunto de resultados de atenção, que são então normalizados, para criar uma distribuição de probabilidade conhecida como pesos de atenção. [13]

Esses pesos de atenção são usados para criar uma soma ponderada dos valores, que representa a informação agregada na qual o modelo deve focar. Esta soma ponderada é então tipicamente passada por camadas adicionais da rede neural para produzir a saída final. [13]

* 1. Trabalhos relacionados

O processamento de vídeos implica elevados níveis de processamento computacional, o que é um dos maiores problemas atualmente tendo em conta o aumento constante do peso e importância das redes sociais. Existem várias alternativas para tentar ultrapassar este problema, entre os quais realizar melhorias de hardware, otimização de algoritmos e a remoção de informação desnecessária. Gharahbagh, A. et all propõem no seu artigo uma solução para este problema, identificando e analisando apenas as partes críticas de cada vídeo o que melhora a performance do sistema de uma maneira geral. [15]

O algoritmo proposto por Gharahbagh, A. et all usa um algoritmo de registo de imagens (frames do vídeo), o método inicialmente começa por remover o movimento da câmara e, posteriormente, cada frame é usado para criar um mapa de saliências. Ao combinar o mapa de saliências com a informação de movimento derivado do fluxo ótico e da segmentação de cores é possível criar um mapa de saliências contendo informação do movimento e também informação espacial. Finalmente, é sugerida uma função não-linear para combinar os mapas de saliência temporais e espaciais, que foi otimizada usando um algoritmo genético de vários objetivos. [15]

A principal diferença deste método com o método usado neste projeto tem a ver com a não inclusão de mapa de saliências de movimento, apenas é analisado cada frame do vídeo. Por outro lado, cada frame é analisado como se fosse uma imagem estática onde não é feita nenhuma segregação de partes críticas da imagem, não se perdendo nenhum dado de cada frame. Este método é mais custoso a nível de processamento computacional, mas entendeu-se que, tendo em conta a aplicabilidade deste projeto, o processamento computacional não terá necessariamente de providenciar resultados instantâneos.

Em relação aos Modelos de Atenção, Fan et all avaliaram como alguns detalhes criados com o objetivo de evidenciar emoções nas imagens se relacionam com a atenção do ser humano. Devido aos limites de capacidade do cérebro humano, nem toda a estimulação a que o ser humano está sujeito, pode ser processada em paralelo e avaliada. De facto, o fenómeno conhecido como atenção seletiva corresponde ao estado em que todos os estímulos visuais competem para se tornar o foco dos olhos e são codificados na memória visual de curto prazo. Existe muita pesquisa que indica que a relevância emocional de um estímulo influencia a atenção seletiva. Por exemplo, as pessoas tendem a prestar atenção preferencialmente a estímulos que evocam emoções, como bebés fofos ou cenas eróticas. [16][17]

Fan et all criaram um conjunto de dados de Atenção Emocional (EMOd). Descobriram que a atenção humana está mais focada em cenas de veículos e animais que provocam admiração no EMOd. Com o objetivo de modelar computacionalmente o comportamento de atenção humana, projetaram uma rede neuronal profunda (CASNet II), que inclui uma sub-rede de ponderação de canais que prioriza objetos que provocam emoção e uma estrutura de Atrous Spatial Pyramid Pooling (ASPP) que aprende a importância relativa de regiões de imagem em múltiplas escalas. [17]

Neste projeto, foi usada a mesma rede neuronal profunda para análise do conjunto de dados fornecidos. Em vez de se procurar estabelecer o melhor método de análise de imagens e apreender a atenção do ser humano ao visualizar essas mesmas imagens, usou-se a rede CASNet II para avaliar imagens de marketing digital e comparar com dados empíricos obtidos pela análise de um conjunto de seres humanos.

Usou-se também a rede neuronal precedente do CASNet II, a CASNet I que se diferenciam essencialmente em 2 assuntos: 1- na rede CASNet II foi usada uma estrutura ASPP customizada para se poder codificar saliências de contextos e escalas múltiplas. Os campos recetivos ampliados do ASPP permitem que a CASNet II aprenda a importância relativa em regiões da imagem maiores, estendendo a prioritização baseada em objetos anteriores para uma área de imagem mais ampla. 2- substituíram a arquitetura de duplo fluxo no CASNet I por um único fluxo e reduziram a escala da imagem de entrada. Com essas mudanças, o CASNet II modela melhor o efeito de prioritização emocional humana e supera significativamente o CASNet I na previsão de saliência em todos os cinco conjuntos de dados de referência, melhorando também o tempo de processamento em quase 300%. [17]

* 1. Tecnologias existentes

Os trabalhos realizados com Aprendizagem Profunda estão ainda bastante dependentes das tecnologias implementadas, uma vez que exigem bastante trabalho computacional. Neste projeto utilizaram-se 2 modelos de redes neuronais diferentes, que já se conhecia à priori a sua grande diferença de performance, mas usaram-se outras tecnologias que se entenderam ser mais eficientes para a concretização do projeto.

**Python**

É considerada pela TIOBE [18], uma organização especializada em métricas de qualidade de software e pela PYPL [19], um índice criado pela Google destinado a analisar a popularidade das linguagens de programação, a linguagem de programação mais usada há vários anos. De facto, a linguagem que adquiriu o nome devido a um grupo de comediantes chamado Monty Python [20], tem-se mantido no top 3 destes rankings desde maio de 2019 e destacadamente número 1 desde junho de 2023.

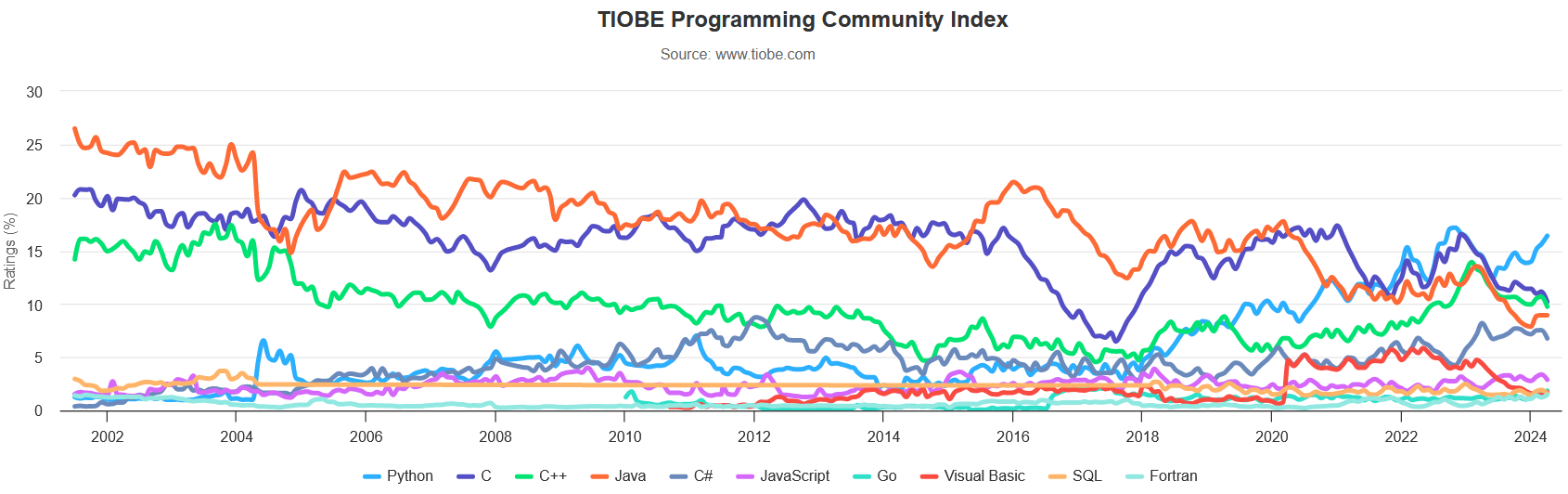
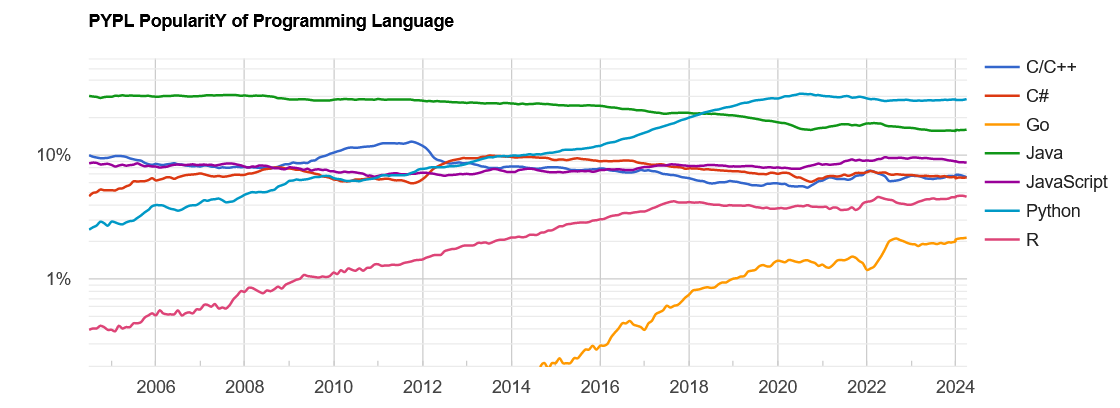
 

Figura 3: Ranking do TIOBE [18] Figura 4: Ranking do PYPL [19]

Python é uma linguagem de fácil manipulação e que, por ter um número significativo de utilizadores, existe bastante documentação. O fato de ser integralmente público (*open* *source*) facilita também a colaboração dos utilizadores desenvolvendo módulos e bibliotecas. [21]

O backend foi desenvolvido em Django que é uma framework web de alto nível para implementação de código Python. No Django, foi ainda necessário o uso de diferentes bibliotecas do Python, desde Pathilib, Datetime, Corsheaders, Rest\_framework, para nomear algumas.

**HTML**

Neste projeto foi necessária também a criação de uma página web de modo a mostrar o resultado obtido nos mapas de saliência do conjunto de dados fornecidos. HTML – Hyper Text Markup Language é a linguagem em que a maior parte das páginas web são desenhadas, normalmente complementada design mais apelativo codificado em CSS.

**Anaconda**

Conda é uma ferramenta de linha de comandos que serve para gerir diferentes ambientes e packages, usando diferentes versões de cada software. [22] No projeto em questão o código dos Modelos de Atenção utilizados exigia versões específicas de diferentes packages, pelo que esta ferramenta foi essencial para a concretização do projeto.

**MySQLFreeNetHosting**

Em relação à base de dados para persistir os dados dos utilizadores, bem como outros dados que pudessem ser pertinentes, foi usada uma base de dados relacional do tipo MySQL, neste caso a base de dados foi alojada em servidores do MySQLFreeNetHosting.net.

**React**

React é uma ferramenta muito utilizada para a criação de código *frontend*, ou seja, que permite a criação de interfaces para o utilizador, comunicando com o *backend* para obtenção de dados. Foi usado VITE para contruir este projeto de forma mais rápida e leve e também vários componentes que foram necessários para concretizar o projeto: axios, jwt-decode, mysqlclient são alguns dos componentes instalados.

1. Análise e desenho da solução

Tendo efetuado na secção anterior a contextualização do tema a abordar e quais os assuntos mais relevantes, torna-se agora importante começar a análise do problema que se pretende resolver.

A documentação do processo de análise e posterior desenho da solução será realizada através do levantamento de requisitos funcionais e não funcionais, de acordo com o modelo FURPS+.

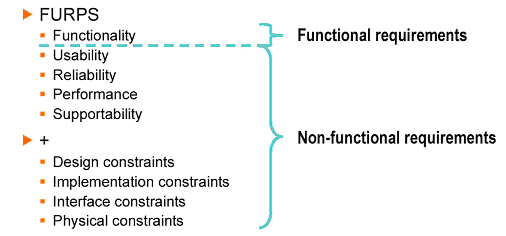
****

Figura 5: Modelo FURPS+ [23]

Tal como se pode verificar na figura 5, o modelo FURPS divide-se em requisitos funcionais: **F**uncionalidade, **U**sabilidade, Confiabilidade (***R****ealibility)*, Desempenho (***P****erformance*), **S**uportabilidade e em requisitos não funcionais: restrições de design, restrições de implementação, restrições de interface e restrições físicas. Cada um deste tipo de requisitos será detalhado e aprofundado, sendo devidamente enquadrado no problema em questão.

* 1. Domínio do problema

Antes de começar com a análise propriamente dita dos requisitos funcionais e não funcionais, é importante esclarecer o domínio em que o problema se vai centrar.

De facto, este projeto prende-se essencialmente com o desenvolvimento de uma ferramenta que irá ser útil na implementação da análise de imagens ou vídeos que, até hoje, teriam de ser manipulados um a um, tendo em conta qual o modelo de análise escolhido. Isto significa que, para cada dado de entrada, teríamos diferentes implementações para cada modelo de análise.

Tendo em conta o âmbito do Marketing Digital, essa ferramenta será muito importante para melhorar esta análise. Pretende-se assim que, numa ferramenta simples e intuitiva, qualquer utilizador possa escolher o modelo de análise numa página web e, ao clicar num botão, obter o resultado do mapa de saliências desse método, sendo esse mesmo resultado mostrado ao utilizador na mesma página web, paralelamente à imagem ou vídeo original.

Estes resultados obtidos através de vários métodos de Modelos de Atenção permitirão ao utilizador adaptar as imagens e obter o mapa de saliências respetivo, repetindo facilmente o processo quantas vezes forem necessárias, até obter o resultado que pretende.

Por outro lado, na análise de vídeos, foram usados vídeos retirados de mundos virtuais com o intuito de analisar a potencialidade de colocar publicidade no mundo virtual e perceber qual a atenção que o utilizador irá ter. Deste modo, também se pode escolher o melhor sítio para essa campanha, bem como adaptar a imagem da mesma maneira usada para a otimização de imagens estáticas.

Com o recurso a linguagens gráficas como o UML, pode-se criar artefactos que podem ajudar na compreensão do problema. Essas linguagens permitem visualizar as ligações entre os vários intervenientes no problema, bem como especificar as responsabilidades de cada uma delas.

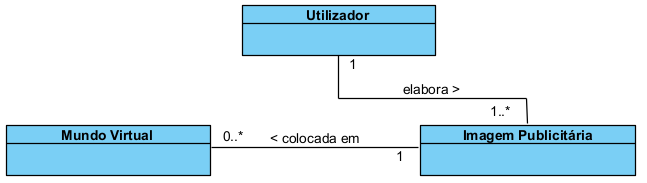


Figura 6: Modelo de Domínio

A figura 6 pretende demonstrar o domínio do problema em questão. No entanto, pode-se facilmente estender as entidades que poderão usar o software desenvolvido, uma vez que o utilizador poderá fazer parte de uma empresa publicitária, com clientes que irão adquirir serviços nessa área, neste caso, a criação da imagem com o melhor resultado de atenção dos seus clientes, bem como a possível colocação dessa imagem num mundo virtual, de modo a obter um melhor resultado da publicidade, apostando em vários canais de marketing.

* 1. Requisitos funcionais e não funcionais

A engenharia de requisitos é um processo de análise e levantamento de necessidades do problema e permite retirar ideias antes da formulação da solução propriamente dira. Esses requisitos podem ser divididos em 2 grandes categorias, requisitos funcionais e não funcionais, que serão descritos de seguida.

* + 1. Requisitos Funcionais

Estes requisitos definem as unidades funcionais, os limites de cada funcionalidade e os limites dos seus efeitos nas operações individuais e conjuntas do sistema. [24]

O diagrama de casos de uso é o diagrama que melhor explicita as funcionalidades do sistema, tal como se pode verificar na figura 7.

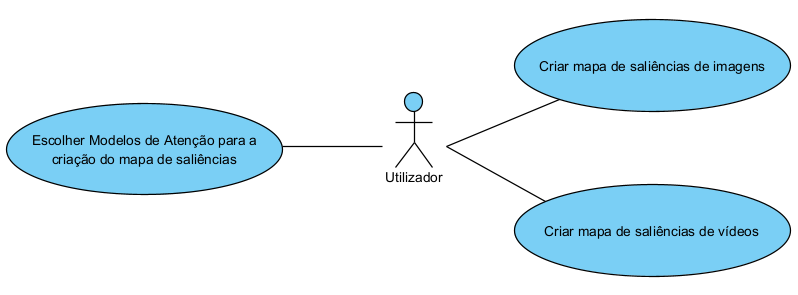


Figura 7: Diagrama de Casos de Uso

**RF001: Escolher Modelos de Atenção para a criação do mapa de saliências**

O primeiro requisito funcional do sistema desenvolvido é o fato do utilizador ter a opção de escolher qual o modelo de atenção que pretende usar na criação no mapa de saliências. Assim, o sistema deverá ter a possibilidade de escolha de mais do que um modelo de atenção de modo a ser possível comparar os mapas de saliências obtidos por cada um desses modelos.

Devo fazer diagramas SSD para cada US?

**RF002: Criar mapa de saliências de imagens**

O requisito funcional de criar o mapa de saliências de imagens tem a ver com a implementação propriamente dita dos Modelos de Atenção escolhidos no requisito anterior e obter o resultado previsto pelo modelo. Com este resultado, pretende-se ter uma previsão de como o utilizador irá interagir, ou seja, como vai percecionar e focar a sua atenção em cada parte da imagem fornecida como entrada de dados.

**RF003: Criar mapa de saliências de vídeos**

O requisito funcional de criar um mapa de saliências de vídeos vai ser usado para receber como dado de entrada um vídeo de realidade virtual, dividir esse vídeo em frames que correspondem a imagens e, para cada imagem, converter essa imagem num mapa de saliências à semelhança do RF002. Posteriormente, essas imagens geradas pelo modelo de atenção escolhido, serão agrupadas num vídeo de mapa de saliência para que o utilizador possa comparar esse resultado com o vídeo original. Deste modo, será possível verificar quais são as zonas do vídeo em que o utilizador presta mais atenção, de modo a que se possa usar esse aumento de atenção para, por exemplo, criar uma publicidade e assim conseguir mais foco do utilizador.

* + 1. Requisitos Não Funcionais do sistema

Os requisitos não funcionais do sistema prendem-se com caraterísticas do sistema que não têm a ver diretamente com a funcionalidade pretendida, mas sim com a maneira como essas funcionalidades são mostradas ao cliente.

**RNF001: Usabilidade**

A usabilidade de um sistema tem a ver com a facilidade de seguir intuitivamente os passos necessários para que um determinado dado de entrada produza um determinado resultado. Essencialmente, tem como objetivo capturar e avaliar requisitos baseados na interface com o utilizador. [25]

Tabela 2: Requisitos de Usabilidade

|  |  |
| --- | --- |
| RNF001.1 | O utilizador deverá ter uma curva de aprendizagem perante a solução desenvolvida bastante pequena |

No sistema a ser desenvolvido, pretende-se que a interface web com que o utilizador vai interagir seja simples de forma a que o utilizador entenda facilmente quais os passos a dar para que forneça um dado de entrada e receba o resultado que pretende.

**RNF002: Confiabilidade**

A confiabilidade de um sistema tem a ver com a disponibilidade, precisão, recuperabilidade e tempo médio entre falhas. [25]

Tabela 3: Requisitos de Confiabilidade

|  |  |
| --- | --- |
| RNF002.1 | O servidor deve-se manter operacional e funcional após falhas, devendo existir cópias de segurança para recuperar rapidamente o sistema. |
| RNF002.2 | O servidor deverá ter alta disponibilidade, ou seja, disponível no mínimo em 99% do tempo. |

Tal como é exposto na tabela 3, no projeto a ser desenvolvido pretende-se que o servidor se mantenha operacional e funcional após falhas. Por outro lado, o servidor deverá ter alta disponibilidade, ou seja, disponível no mínimo em 99% do tempo.

**RNF003: Desempenho**

O desempenho de um sistema diz respeito a rendimento, tempo de resposta do sistema, tempo de recuperação e tempo de inicialização. [25]

Tabela 4: Requisitos de Desempenho

|  |  |
| --- | --- |
| RNF003.1 | O sistema deverá produzir resultados, independentemente do Modelo de Atenção disponibilizado no tempo máximo de 1 minuto. |

Neste projeto, o desempenho do sistema é algo que poderá ser bastante crítico. Isto deve-se não necessariamente à base do sistema, mas sim ao tempo necessário para o cálculo do resultado dos modelos de atenção. Como já foi referido, um dos problemas da Aprendizagem Profunda tem a ver com o tempo de processamento e neste projeto esse problema também será uma questão importante. Sempre que necessário poderá ser importante a introdução de algum tipo de contador de tempo para que o utilizador possa acompanhar o tempo de processamento e não desistir do pedido ao sistema, aguardando pelo resultado pretendido. 🡪 é só uma ideia, tenho de ver como implementar

**RNF004: Suportabilidade**

A suportabilidade de um sistema está relacionada com limitações de linguagem, testabilidade, adaptabilidade, manutenibilidade, compatibilidade e escalabilidade. [25]

Tabela 5: Requisitos de Suportabilidade

|  |  |
| --- | --- |
| RNF004.1 | O sistema deverá permitir a inclusão simples de novas funcionalidades. |
| RNF004.2 | O sistema deverá permitir a implementação de outros modelos de atenção de forma fácil. |

O sistema deverá permitir a inclusão de novas funcionalidades, de maneira simples. O sistema deverá ser desenvolvido de maneira a poder incluir outros modelos de atenção mantendo o design da página web.

**RNF005: Restrições de desenho**

Nas restrições de desenho é limitada a forma como a aplicação é desenhada. [25]

Tabela 6: Restrições de Desenho

|  |  |
| --- | --- |
| RNF005.1 | O sistema deverá comunicar com a base de dados de forma assíncrona. |

No presente projeto, a comunicação com a base de dados deve ser assíncrona, uma vez que essa comunicação não será contínua e assim evita-se a latência do sistema. A base de dados a implementar deverá conter, numa primeira fase, apenas os dados necessários para registo e autenticação no sistema.

**RNF006: Restrições de implementação**

Num sistema implementado em páginas web, é importante a compatibilidade entre diferentes navegadores de internet.

Este é um exemplo de uma restrição de implementação, uma vez que afeta a forma como o sistema é concebido, causando por vezes impacto no design da solução. [25]

Tabela 7: Restrições de Implementação

|  |  |
| --- | --- |
| RNF006.1 | O sistema deverá poder ser utilizado em diferentes navegadores de internet. |

**RNF007: Restrições de interface**

No caso das restrições de interface, estas afetam a forma como certos componentes comunicam com outros componentes. [25] Na solução em questão, a ligação à internet permanente é fundamental para que não haja interrupções no serviço a disponibilizar.

Tabela 8: Restrições de Interface

|  |  |
| --- | --- |
| RNF007.1 | O sistema deverá manter a ligação à internet permanentemente. |

**RNF008: Restrições físicas**

O que pôr aqui???

* 1. Desenho

Tendo agora o problema corretamente clarificado e também devidamente analisados e definidos os requisitos funcionais e não funcionais, segue-se a proposta de desenho do sistema.

O desenho da solução é o processo de identificar a melhor estrutura e funcionamento do programa desenvolvido de forma a responder com eficiência aos requisitos funcionais e não funcionais identificados na secção anterior.

* + 1. Arquitetura do sistema

Tendo em conta as boas práticas de desenvolvimento de software e o domínio do problema apresentado na secção 3.1 deste relatório, foi feita a divisão do sistema por camadas.

O principal objetivo desta divisão prende-se com a responsabilidade de cada camada ser única e assim poder-se atuar facilmente aquando de processo de manutenção e escalabilidade, melhorando também a adaptabilidade e a segurança. Foram então definidas as seguintes camadas:

* *User Interface* codificado em React, onde vai estar desenhada a interface que o utilizador vai aceder. Esta interface irá receber os pedidos do utilizador, comunicar com o *Web Service* para solicitar dados e apresentar os mesmos ao utilizador no *Browser*
* *Web Service*, codificado em Django é o componente responsável por interagir com a base de dados, para posterior envio desses dados à *User Interface*. Irá também possuir um serviço para pedir obter os dados dos mapas de saliências, usando uma ligação aos módulos *Python*.
* Modelos de Atenção, codificados em *Python*, serão os módulos de criação dos mapas de saliências, realizando a previsão dos modelos de atenção.
* Base de dados será um componente para registo dos dados de utilizador. Este componente deverá comunicar com o componente *Web Service.*
* *Browser* é o componente que irá ser usado pelo utilizador para realizar pedidos ao sistema e obter informação de forma apelativa, através de uma página HTML com CSS.

Este desenho da solução é mais facilmente visível num diagrama de componentes, tal como mostra a figura 8.

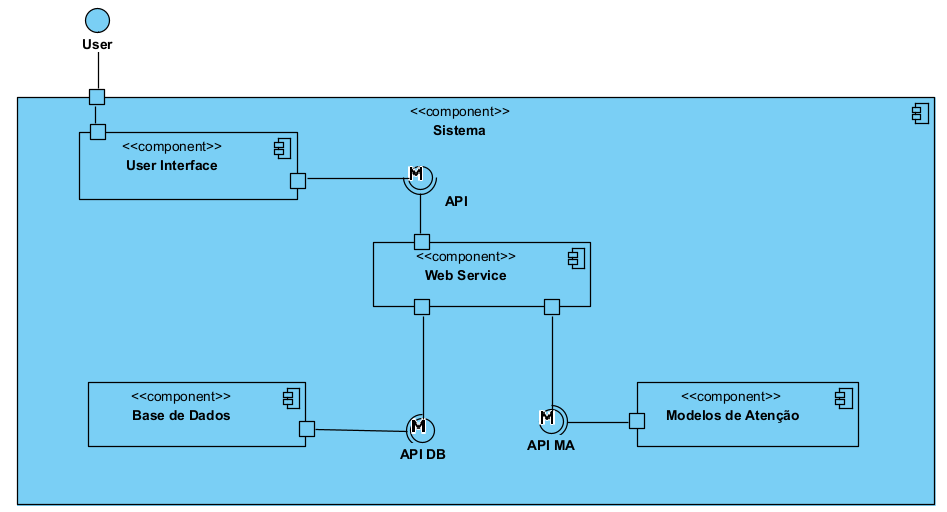


Figura 8: Diagrama de componentes do sistema

Tendo em conta os componentes acima explicitados é importante perceber também como eles estarão inseridos fisicamente nas estruturas disponíveis. De forma a validar esta implementação física, de seguida mostra-se na figura 9 o diagrama de implantação do sistema.

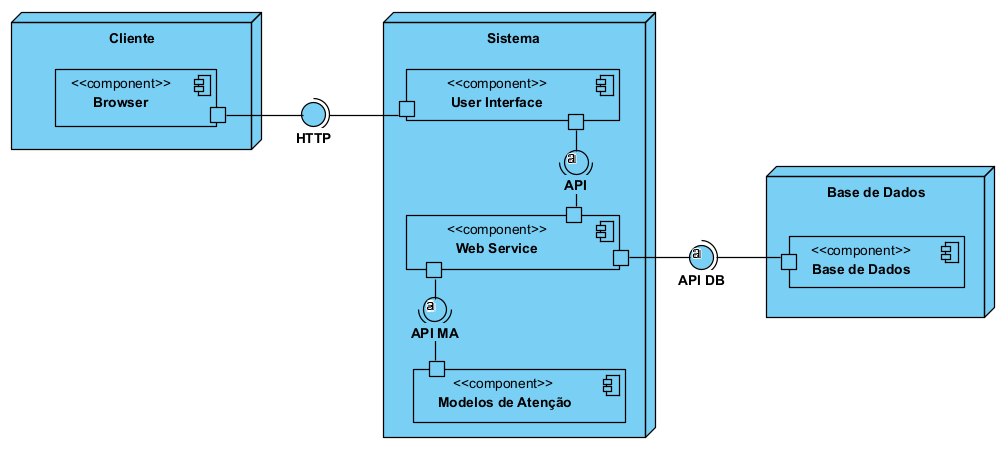


Figura 9: Diagrama de implantação do sistema

Devo fazer um Diagrama BPMN do processo que o utilizador percorre durante o pedido de criação do mapa de saliências??

* + 1. Solução proposta

Na solução proposta, temos os diferentes componentes com responsabilidades únicas e bem definidas, pelo que a solução é ótima no sentido de corresponder aos requisitos funcionais e não funcionais.

O facto de haver vários componentes permite uma melhor escalabilidade do sistema, bem como uma melhor adaptabilidade. De facto, se houver no futuro a necessidade de trocar a base de dados, por exemplo, apenas se terá de atualizar a ligação da API DB para que essa troca mantenha o sistema em funcionamento.

Nesta proposta de solução, existe também uma API MA que fará a ligação com o componente Modelos de Atenção que terá várias implementações, tantas quantas os Modelos de Atenção a serem implementados.

* + 1. Solução alternativa

Como solução alternativa, poder-se-ia pensar na ligação da base de dados diretamente ao componente da *User Interface,* em vez da ligação ao componente *Web Service*. Pode-se ver esta solução alternativa na figura 10 abaixo.

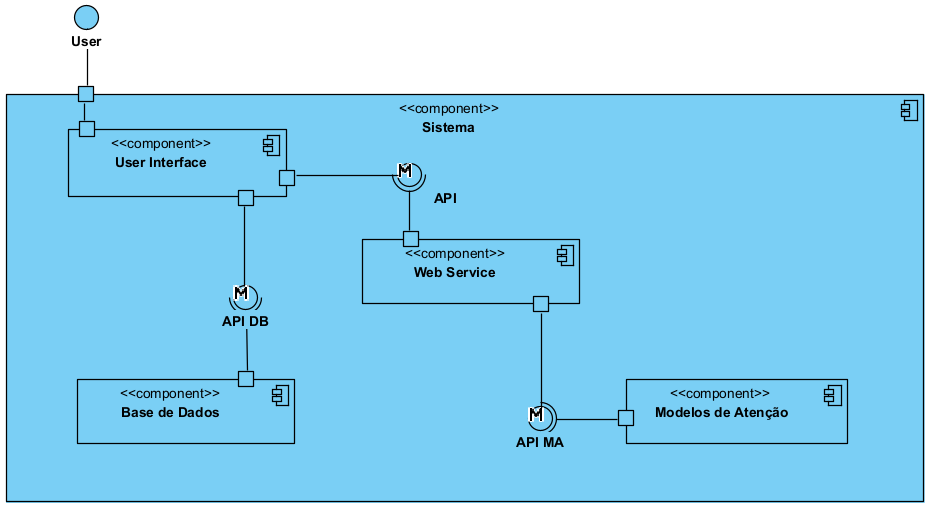


Figura 10: Diagrama de componentes alternativo do sistema

De facto, a base de dados tal como foi solicitada no âmbito deste projeto apenas irá guardar os dados do utilizador e o seu tipo de perfil. Nesse sentido, é uma base de dados simples e que não irá guardar dados que o *backend* irá produzir.

Deste modo, esta solução mantém os padrões de desenvolvimento de software, mas tem um risco acrescido em segurança e também na usabilidade do sistema uma vez que não permite a adição de novas funcionalidades a serem persistidas na base de dados.

Em termos de segurança, o fato dos dados dos utilizadores passarem pelo *Web Service* permite uma camada adicional de segurança, pelo que esta solução é menos suscetível a ataques de pirataria e de roubo de dados confidenciais dos utilizadores. Por outro lado, se no futuro houver a necessidade de guardar outros dados na base de dados, ter-se-ia de criar outra ligação entre o *Web Service* e a base de dados, uma vez que esta solução não contempla essa opção.

* + 1. Conclusões

1. Implementação da Solução

Neste capítulo irão ser apresentação os detalhes relacionados com o enquadramento e implementação das soluções explicitadas no capítulo anterior. Assim, o modo como os 4 principais componentes, *User Interface*, *Webservice*, Modelos de Atenção e a Base de Dados foram implementados e outros detalhes relevantes para a explicação do projeto serão detalhados neste capítulo.

* 1. Descrição da implementação

O primeiro passo num projeto deste género é criar, e posteriormente ativar, um ambiente virtual onde serão instalados os packages específicos para a execução desse mesmo projeto. Deste modo, nenhum outro projeto será afetado pela instalação ou não de certos packages, bem como se assegura que apenas os packages necessários neste projeto são os que estão instalados.

Como foi referido no capítulo anterior usou-se o Conda, versão mini-conda para essa instalação de packages específicos necessários para o código dos Modelos de Atenção funcionar corretamente.

Tendo em conta a especificidade e complexidade de usar os Modelos de Atenção exigidos na proposta inicial deste projeto, foi necessário o uso do *Jupyter Lab*. O *Jupyter Lab* é um software de desenvolvimento gratuito, muito flexível e com base web que pode ser usado para vários fins, entre os quais *machine learning*.

4.1.1. Web Service

O módulo WebService, ou seja, o backend foi desenvolvido em Python através da framework Django. Para além de ser uma linguagem muito utilizada, o fato do código dos Modelos de Atenção também ser desenvolvido em Python ajudou à comunicação entre os vários componentes. Por outro lado, uma vez que era desconhecida esta linguagem à partida para este projeto, fez com que o contato com Python não tivesse sido apenas no código dos Modelos de Atenção, mas sim para todo o desenvolvimento do backend. Deste modo, foi possível aprofundar mais o conhecimento desta linguagem de programação.

Após a ativação do ambiente virtual do projeto, é então necessária a instalação de todas as dependências para a execução do projeto. Já era sabido que seria necessária a instalação do Django, ou djangorestframwork, ou ainda PyJWT para nomear algumas. Depois dessa instalação, é necessária a criação de um novo projeto Django. Com esta criação é criada uma pasta com a informação genérica deste projeto, como por exemplo os settings ou urls.

De seguida, deve ser criada uma Django app onde se podem personalizar os diversos componentes e, assim, organizar o código logicamente. Deverá ser criada uma nova app para cada tipo de funcionalidade da aplicação, de modo a ser mais fácil a sua gestão e manutenibilidade.

Depois de serem personalizados alguns dos settings genéricos do projeto, como por exemplo, os dados de acesso à base de dados ou definir os tempos de acesso ou atualização de token dos JWT, o próximo passo é a criação dos end-points que teremos no projeto. Os end-points podem ser criados genericamente na app que contém a informação genérica do projeto, ou individualmente em cada app criada. No presente caso, foram criados end-points de autenticação no ficheiro genérico do programa e end-points específicos em cada app que, posteriormente, são chamados do ficheiro genérico url.py.

Prints dos 2 urls.py

Ao serem criados os end-points, isto é, os endereços da internet específicos para cada caso de uso, definem-se as views a serem executadas. Em Django, as views são funções que recebem um pedido web e retornam uma resposta web. No projeto em questão, foram criadas várias funções na app “api”: uma para criar utilizadores, uma para criar e listar notas, uma para apagar notas. No caso da app “casnet”, que deu suporte ao cálculo dos mapas de saliências obtidos, as funções criadas prendem-se com o processamento de imagens e outra com o processamento de vídeos.

Prints do views.py

O Django já possui um módulo de autenticação pelo que se fez uso desse módulo para tratar todo o acesso à aplicação, não sendo necessária nenhuma implementação específica. Por outro lado, foi necessária a criação de um modelo de Nota, na app “api”. Os modelos representam as classes a serem persistidas. Cada atributo do modelo representa um campo na base de dados e cada modelo mapeia uma tabela na base de dados.

Prints do modelos.py

O módulo de autenticação do Django faz uso do JWT, Json Web Token. Sempre que um website é acedido, o JWT vai atuar como as permissões ou uma autenticação a esse website. Uma vez que, sempre que fazemos um pedido web, o backend tem de saber quem fez o pedido e que permissões tem, é incluído no pedido um token JWT que será descodificado de modo a perceber quais são as permissões que esse pedido tem. Inicialmente, um utilizador inicia o login e, ao colocar os seus dados de acesso, o frontend envia essas credenciais ao backend. O backend confirma a veracidade dessas credenciais e concede 2 tokens, um token de acesso e um token de atualização que serão guardados temporariamente no frontend. O token de acesso será o usado em todos os pedidos e, para aumentar a segurança, este token tem uma validade curta, no caso foi definido 30 minutos. Assim, é definido um token de atualização, que vai atualizar o token de acesso frequentemente, mas que também tem um tempo de vida (24 horas) após o qual as credenciais de acesso terão de ser fornecidas novamente. Este processo e tipo de acesso aumenta consideravelmente a segurança do website. É, assim, necessária a configuração dos tempos limites dos tokens de acesso e de atualização, que foram executados no ficheiro genéricos de configurações do projeto settings.py.

Prints do settings.py

4.1.2. Modelos de Atenção

A implementação dos Modelos de Atenção CasNET1 e CasNET2 foi sem dúvida a parte mais complexa deste projeto. Apesar do código fonte destes modelos estar disponível online, o seu uso é bastante complexo devido a vários fatores. O que é provavelmente mais difícil prende-se com as versões específicas de packages necessários à sua execução que cada modelo necessita. Foi necessário obter essa informação, instalar cada package na versão correta, sendo que a maior parte das vezes não era a versão mais atual, de modo a poder executar o código, de outro modo era impossível a obtenção de qualquer resultado. Convém relembrar que esta instalação foi realizada num ambiente virtual criado especificamente para todo o projeto.

Após ser ultrapassado esse problema, foram criados 2 end-points no backend, mais concretamente na app “casnet” específicos para cada um dos modelos.

Falta acabar esta parte quando o código estiver ok

4.1.3. User Interface

A *User Interface* foi desenvolvida em *React*. Tal como o *Python*, o *React* é uma ferramenta não lecionada no Curso de Engenharia Informática e também bastante utilizada no mercado de trabalho, pelo que foi adotada no sentido de aprofundar a aprendizagem de diferentes tecnologias.

Tal como referido anteriormente usou-se a *framework* Vite.js em vez de outras, como por exemplo Next.js. Apesar da framework Next ter algumas caraterísticas interessantes como o ISR, que permite a atualização de páginas estáticas sem ser necessária o build da página completa ou ainda uma componente de otimização de imagens, como mudança de tamanho ou recorte das mesmas que poderia ser interessante no projeto em desenvolvimento, optou-se por usar a framework Vite.js.

De fato, esta framework permite uma divisão do código que faz com que diversas partes da página apenas vão carregadas à medida que são necessárias, bem como aceita diversos plugins como pré-processadores CSS ou outros que melhoram a rapidez do arranque e atualizações da página. Sendo o tempo um fator crucial, tal como referido anteriormente, em projetos de Aprendizagem Profunda, a escolha da framework recaiu na que mais rápido desempenho teria.

Assim que foi criada a pasta com o React instalado, foi necessária a instalação de algumas dependências, tais como Axios, Router-dom e jwt-decode.

O projeto template do *React* não tem divisões lógicas, pelo que foram criadas diversas pastas para organizar o código, tais como as páginas, estilos e componentes.

Inicialmente, foi criado um intercetador de pedidos web. O que este código fará é, para cada pedido, verificar se existe um token de acesso válido e, se houver, anexá-lo no cabeçalho do pedido. Deste modo, em todo o código a ser escrito posteriormente não haverá a necessidade de continuamente verificar se existe esse token, este código permitirá verificar isso em cada pedido.

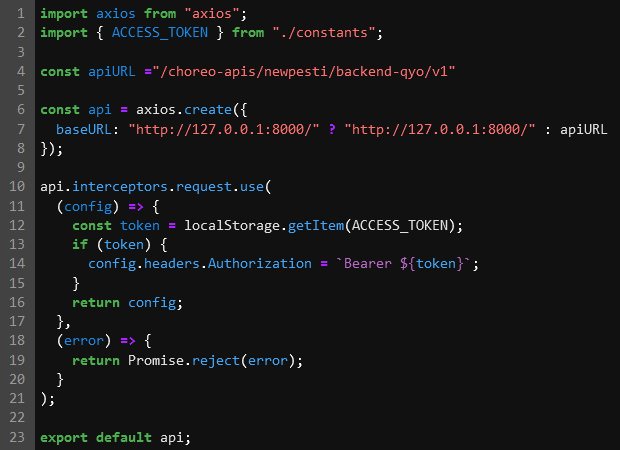


Figura 11: Intercetador de pedidos web para inclusão de token de acesso

Para além desta configuração, foram criados componentes React que são funções javascript que podem ser chamadas em diversas partes do código frontend para compor o que finalmente é mostrado ao utilizador no browser. Para além da criação do componente Form, que serviu de base às páginas de registo e de login do utilizador, sendo ambas muito semelhantes, talvez o componente mais importante de realçar é o componente de ProtectedRoute. A função deste componente é atualizar automaticamente o token de acesso, para que o utilizador possa continuar a ter acesso à aplicação, sem ter que se preocupar com a atualização do login. No caso de ter havido algum problema, ou ainda no caso do token de atualização ter expirado, então esta função vai pedir um novo login para que seja gerado um novo token de acesso.

A seguir à criação dos componentes, passou-se então à criação das páginas codificadas em javascript, para que depois sejam interpretadas no browser. Foram então criadas 4 páginas, uma para registo do utilizador, outra para login do utilizador, uma página genérica de erro e a página principal do projeto, Home.jxs. Nesta página, foram criados botões específicos para a renderização de imagens e vídeos, um outro botão para escolha dos Modelos de Atenção, dois espaços para serem mostradas as imagens/vídeos escolhidos e o output dos modelos para melhor comparação. Finalmente, cada utilizador tem também a sua própria área de comentários/notas que ficam unidos na base de dados à imagem e ao modelo utilizado no cálculo do respetivo mapa de saliências.

Print do home.jsx?

4.1.4. Base de dados

Neste projeto a base de dados foi desenhada de uma maneira simplista, uma vez que os dados a serem persistidos são dados simples, ou seja, apenas dados de login para acesso à ferramenta e, por outro lado, dados de texto simples para ser possível ao utilizador fazer corresponder um comentário ao mapa de saliências obtido. Assim, a pesquisa destes dados será também algo simples, pelo que a implementação de bases de dados mais modernas, do tipo NoSQL, como Redis ou do tipo documentais, como MongoDB não trariam vantagens para o projeto em questão.

Foi então necessário criar 2 classes para persistir os dados, uma para os utilizadores, a classe UserSerializer e a classe NoteSerializer.

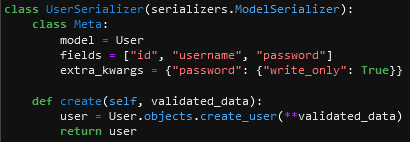


Figura 13: Classe User Serializer

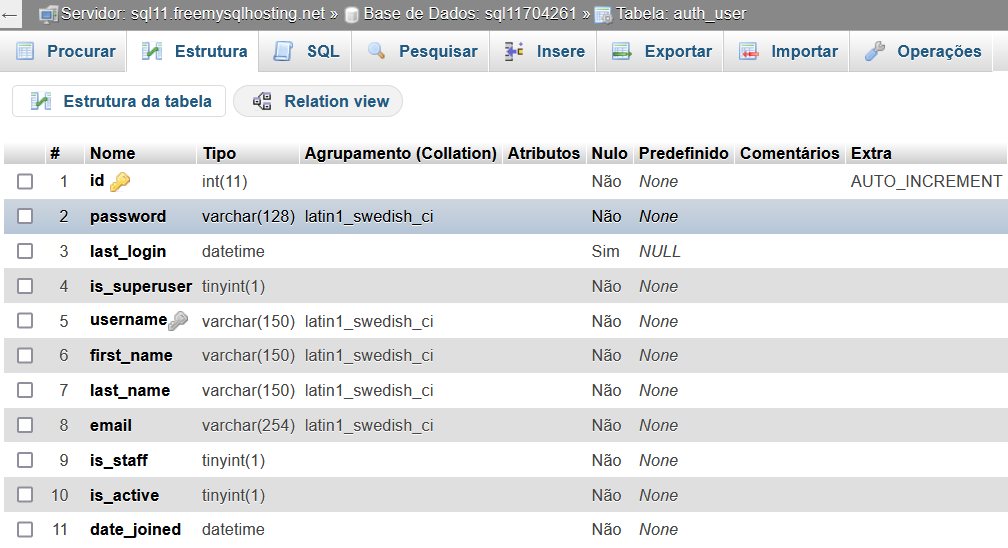


Figura 14: Tabela da base de dados da Classe User Serializer

Como é possível ver nas duas figuras anteriores, os campos codificados para serem persistidos na base de dados não coincidem com os campos da tabela na base de dados. De fato, esta diferença é devida ao fato do Django possuir um sistema de autenticação incorporado que contém todos os campos que ficam guardados na base de dados. Uma vez que a classe User Serializer apenas tem como campos a persistir o id, username e a password, apenas estes campos têm dados nas tabelas da base de dados. Uma possível melhoria futura poderia ser a introdução de mais campos no registo de cada utilizador, se tal fosse necessário por algum motivo. Caso contrário, poder-se-ia apagar os campos não usados das tabelas para melhorar a performance da base de dados.

A classe NoteSerializer, para além do comentário do utilizador, tem também um campo relativo ao URL da imagem escolhida, um campo para o método de atenção e também uma chave estrangeira para associar estes dados ao utilizador que os executou, para além da data/hora da criação. De notar que no campo de utilizador vemos uma chave cinzenta, que tem o significado de chave estrageira, ou seja, é o campo que liga esta tabela à tabela de utilizadores. As chaves douradas significam chaves primária, ou seja, o campo único da tabela que permite a identificação unívoca do registo.

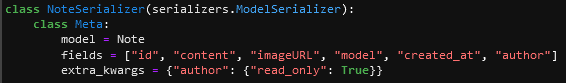


Figura 15: Classe Note Serializer

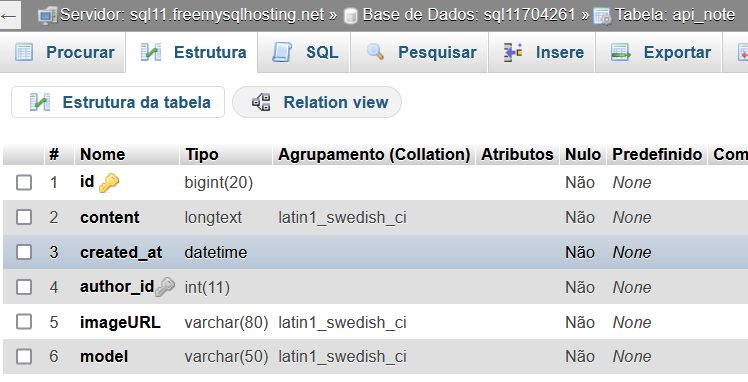


Figura 16: Tabela da Classe Note Serializer

Posteriormente, foi necessário configurar os dados de acesso à base de dados nos settings do projeto e, assim, os dados passaram a ser persistidos na base de dados.

Salienta-se que o provedor da base de dados funciona bem, mas por vezes tem alguns problemas na latência de rede, o que foi problemático após o *deployment* do *frontend* e *backend*. A consequência foi haver vários logins que não foram bem sucedidos, essencialmente devido ao excessivo tempo de resposta da base de dados.

4.1.5. Controlo de versões

O controlo de versões é a prática de rastrear e gerir alterações no código desenvolvido. Em todos os projetos realizados no âmbito da Licenciatura, sempre foi usado o *Bitbucket*, pelo que, mais uma vez procurando a melhoria a aprofundar de conhecimento, foi decidido usar o *GitHub*. O *GitHub* e o *Bitbucket* são ambos *softwares* de controlo de versões, sendo que o *Bitbucket* é mais focado em repositórios privados e apenas permite 5 colaboradores nesses repositórios gratuitamente, ao passo que o *GitHub* é focado em repositórios públicos.

Comparando os dois *softwares* em termos de usabilidade, ambos são simples e permitem a integração com outras plataformas também bastante usadas em desenvolvimento de *software*.

De referir que no início do projeto estava a ser usado um repositório gerido no *Bitbucket.* Quando foi necessária a integração do código dos Modelos de Atenção no projeto Django entretanto desenvolvido, ocorreram diversos erros de conflito de versões de packages, pelo que foi necessário criar um ambiente novo, adicionar os modelos e posteriormente integrá-los noutro projeto Django. Na sequência destas alterações, foi necessário criar um novo repositório e, sabendo da boa integração do *Choreo* com o *GitHub*, optou-se por criar um novo repositório no *GitHub*.

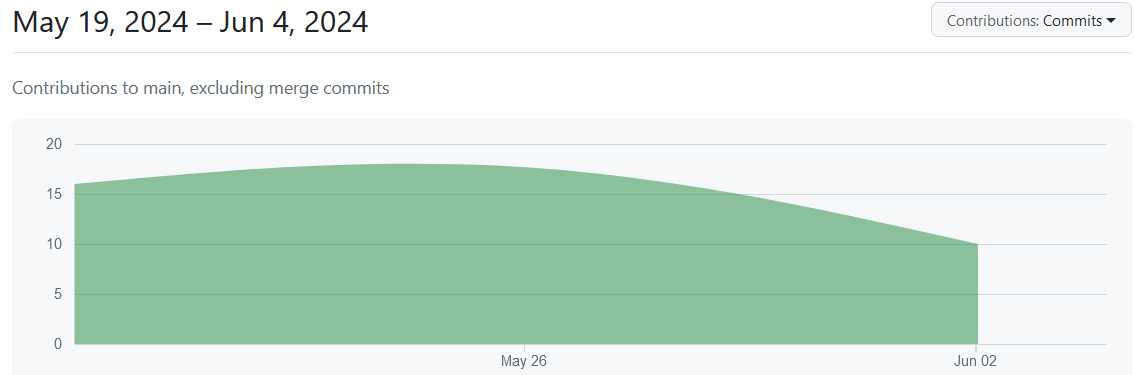


Figura 17: Commits realizados no repositório GitHub

Tal como se pode verificar na figura acima, existiram sempre *commits* regulares, apesar do repositório *GitHub* não demonstrar a totalidade do tempo de desenvolvimento do projeto.

4.1.6. CI/CD

4.1.7. Controlo de tarefas

No início do projeto tinha sido pensada a criação de um projeto na ferramenta *Jira* para ajudar no controlo de tarefas a realizar. Apesar desta ferramenta ser uma ferramenta muito poderosa no sentido em que permite, para além do controlo das tarefas e o estado em que cada uma está no processo de desenvolvimento, permite ainda a ligação de cada uma dessas tarefas a *commits* no repositório e permite a ligação com outras ferramentas de documentação como o *Confluence*. Para além disso, permite a extração de diversos tipos de métricas que são particularmente importantes no desenvolvimento de *software* em equipas.

Neste projeto em específico, entendeu-se que uma ferramenta tão vasta e completa não seria necessária, uma vez que se trata apenas de um elemento a realizar todo o projeto. Assim, optou-se por uma ferramenta mais simples, *Trello*, que funciona praticamente da mesma maneira, embora não permita a extração de métricas de uma maneira simples como o *Jira*, como por exemplo, o *burndown chart*.

Assim, foi possível usar esta ferramenta de modo a poder planear as várias tarefas do projeto e ir acompanhando o seu estado.

Meter aqui board do trello

Descrever a implementação da solução proposta no capítulo anterior, podendo ser dadas explicações e evidências de soluções intercalares. Devem também ser descritas as tecnologias e metodologias utilizadas (software, sistemas de operação, linguagens, dispositivos ou outras ferramentas) e perspetiva crítica sobre as mesmas.

Esta secção descreve a implementação da solução proposta no capítulo anterior. Alguns dos diagramas referidos na secção anterior podem aparecer nesta secção. Por exemplo, diagramas de classes ou diagramas de módulos, sendo detalhadas as operações de cada classe ou as funções de cada módulo (diagramas de atividades). Devem também ser descritas as especificidades de implementação de acordo com o ambiente de desenvolvimento, plataforma e linguagem escolhida para o desenvolvimento e deve ser claro que o desenho apresentado anteriormente foi, de facto, adotado na implementação

* 1. Testes

No desenvolvimento de uma aplicação são sempre necessários diferentes tipos de testes unitários: os testes unitários para verificação e validação dos cálculos e registos realizados no backend, os testes verificação e validação dos pedidos de frontend realizados ao backend e ainda os testes end 2 end, ou seja, a verificação do uso da página no browser que gera os pedidos corretos no frontend.

Os testes de backend prendem-se com a verificação da criação do utilizador (com utilizador e respetiva password) e com a criação da nota (com conteúdo, URL da imagem, método escolhido e utilizador que a criou). Ao correr o comando de testes de python no backend, obteve-se então a seguinte resposta:

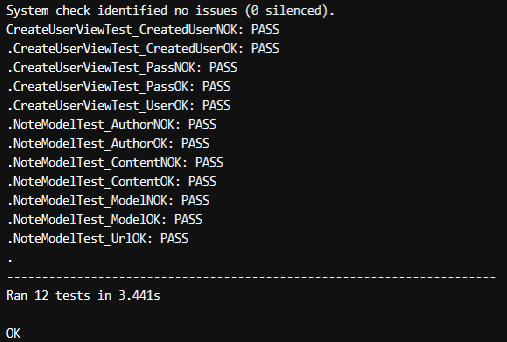


Figura 17: Testes unitários de domínio

Convém referir que foram realizados testes em que o resultado é positivo e também em que o resultado é negativo, como a criação de uma nota sem ter utilizador associado, por exemplo. Ir-se-á agora detalhar um dos testes de cada tabela da base de dados, um relativo aos utilizadores e um teste relativo às notas.

Em relação ao utilizador, como se pode ver na figura abaixo, primeiro define-se o nome de utilizador e a password, de seguida realiza-se um pedido POST à base de dados onde esses dados são registados e finalmente faz-se uma verificação através do comando self.assertEqual que a criação na base de dados foi bem sucedida, ou seja que o código HTTP da resposta ao pedido POST é 201.

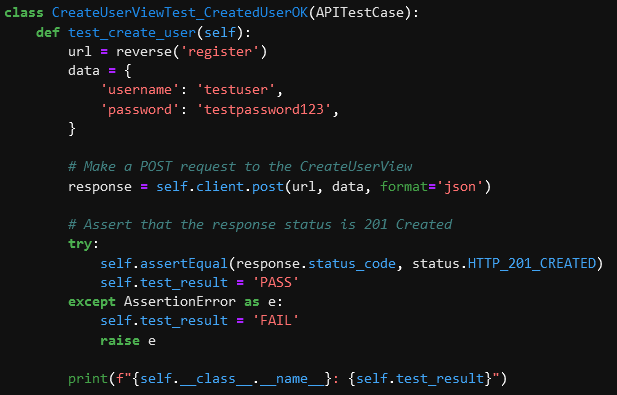


Figura 18: Teste unitário de criação de utilizador

Em relação à nota, como se pode ver na figura abaixo, primeiro define-se um autor. Depois, ao executar o comando Note.objects.create e detalhando os vários campos necessários, o conteúdo, o URL da imagem, o modelo escolhido e o autor escolhido, a nota é gravada na base de dados. De seguida realiza-se um pedido GET à base de dados pelo id da nota criada e finalmente faz-se uma verificação através do comando self.assertEqual que a conteúdo da nota é o mesmo que foi indicado originalmente. Se assim for, o resultado do teste unitário é PASS.

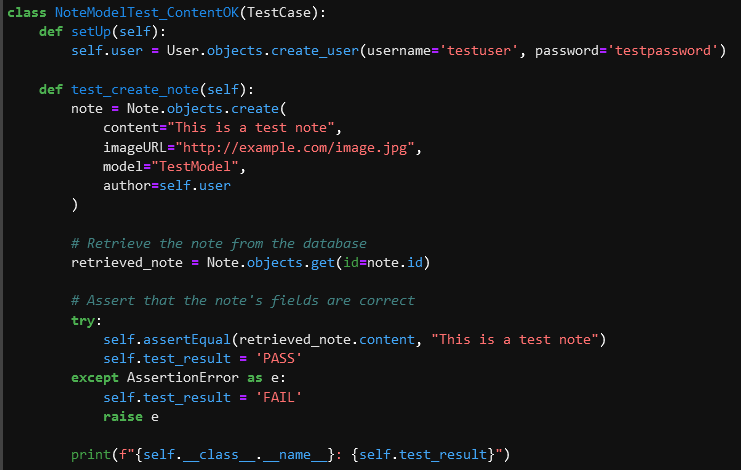


Figura 19: Teste unitário de criação de nota

Os testes de frontend necessários neste projeto têm a ver com o registo/login do utilizador e também com o pedido ao backend de criação do mapa de saliências de imagens ou vídeos.

Para tal, a função que se vai usar para validação do pedido ao backend é o cy.intercept que é uma função que vai verificar se o pedido POST foi feito e se o status do resultado é 200, ou seja, que o pedido POST foi concretizado.

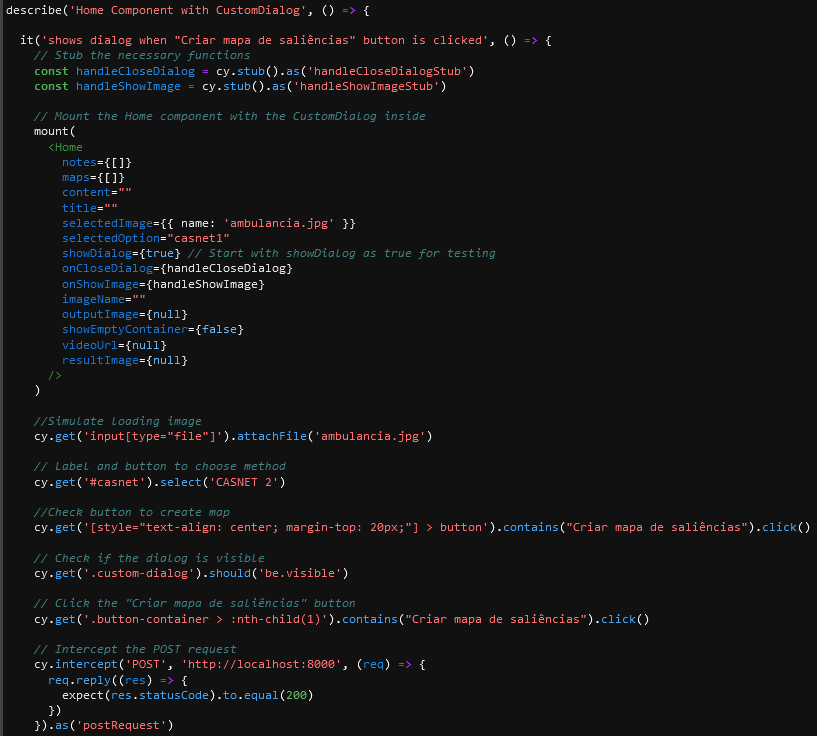


Figura 20: Teste de validação de criação do mapa de saliências

Finalmente, é necessário a verificação proveniente dos testes e2e, que são os testes que validam que as opções escolhidas pelo utilizador são corretamente enviadas para o frontend e o devido tratamento é realizado, seja apenas um routing para uma nova página ou o envio de dados para o backend.

Seguindo o mesmo raciocínio adotado anteriormente, foram criados diversos casos correspondentes a diferentes opções que o utilizador pode escolher, sendo que o resultado final desses testes foi o que está demonstrado na figura abaixo.

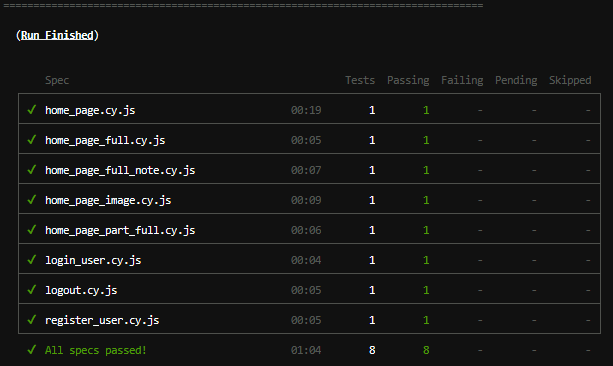


Figura 21: Resultado de Testes e2e

Será agora apresentado um teste e2e que representa o caso em que o utilizador decide criar o mapa de saliências a partir de uma imagem, mas que decide não escrever nenhuma nota acerca deste processo.

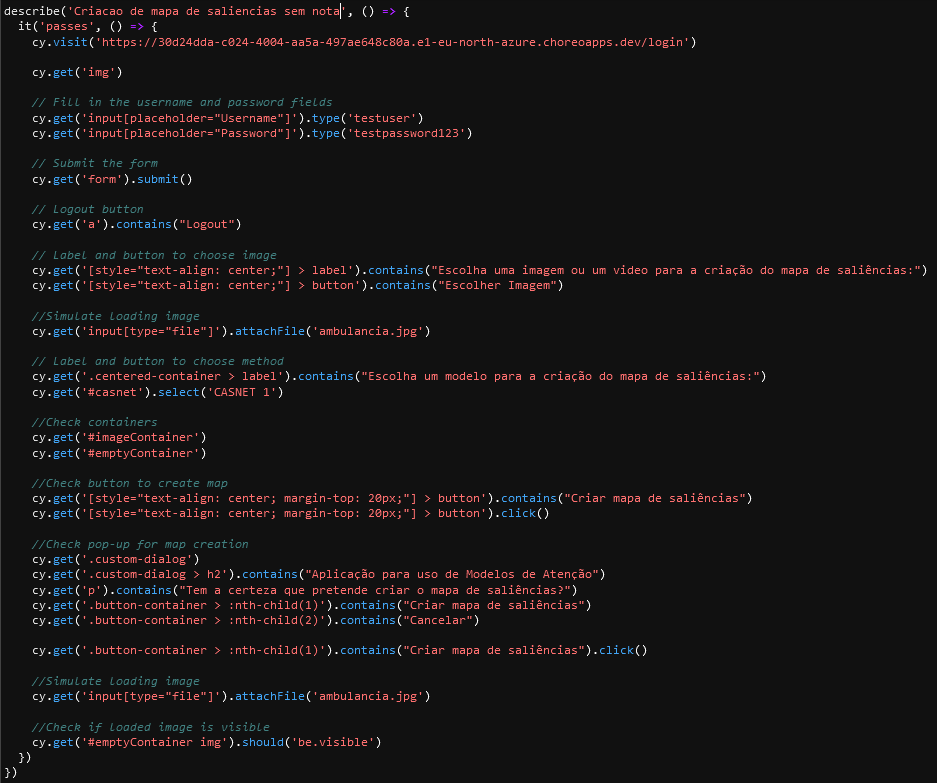


Figura 22: Testes e2e relativo à criação de um mapa de saliências de uma imagem sem criação de nota associada

Como é possível verificar na imagem acima, inicialmente o teste começa por visitar a página de login e colocar dados de utilizador de teste, caso contrário a criação de mapas de saliências não é permitida. Depois é adicionada uma imagem de teste e escolhido um Modelo de Atenção. Depois é feita a confirmação do pedido a ser enviado ao backend e finalmente é verificada a existência de uma imagem no espaço em que o output tem de ser renderizado. Fica assim verificado que o pedido foi feito ao backend, que o mesmo foi processado e renderizado no browser a mostrar ao utilizador.

* 1. Avaliação da solução

Nesta secção deve ser descrita a abordagem seguida para avaliar a solução, ou parte da solução, nomeadamente um ou mais requisitos de qualidade (e.g. desempenho, usabilidade). São descritas experiências efetuadas e apresentados os dados/modelos utilizados, bem como os resultados obtidos. Devem ser descritos e avaliados os resultados obtidos. Deve ser feita uma discussão sobre a adequação dos resultados obtidos relativamente aos planeados anteriormente.

Esta secção poderá não existir em alguns relatórios de projeto/estágio, mas nesse caso deverá ser dada uma justificação para tal.

1. Conclusões

O capítulo de conclusões é um dos mais importantes do relatório, devendo ser apresentado um resumo dos resultados do trabalho efetivamente desenvolvido. ***As conclusões devem basear-se nos resultados realmente obtidos***. Devem enquadrar‑se os resultados obtidos com os objetivos enunciados e procurar extrair conclusões mais gerais, eventualmente sugeridas pelos resultados. Podem acompanhar as conclusões incluindo recomendações apropriadas resultantes do trabalho, nomeadamente sugerindo e justificando eventuais extensões e modificações futuras.

* 1. Objetivos concretizados

Nesta secção devem ser repetidos os objetivos apresentados no capítulo de introdução e para cada um deles deve ser descrito o seu grau de realização. Recomenda-se o uso de uma lista ou tabela, dado que facilita a leitura e compreensão.

* 1. Limitações e trabalho futuro

Nesta secção devem ser identificadas as limitações do trabalho realizado, fazendo uma análise autocrítica do trabalho realizado, bem como extrapolar eventuais direções de desenvolvimento futuro.

* 1. Apreciação final

Esta secção deve fornecer uma opinião pessoal sobre o trabalho desenvolvido.

Referências

- https://github.com/data-science-projects-and-resources/Data-Science-EBooks/tree /main

[1] PHYNHANCAI, “No Title,” *ISEP*. https://www2.isep.ipp.pt/isrc/index.php?page=phynhancai

[2] W. Mckinney, *Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter*, 3rd editio. 2022. [Online]. Available: https://wesmckinney.com/book/

[3] A. Subasi, *Practical Machine Learning for Data Analysis Using Python*. 2020.

[4] D. Halligan, B. & Shah, *Inbound marketing: get found using Google, social media and blogs*. 2010.

[5] ComunicaWeb, “Inbound Marketing.” https://comunica-web.com/blog/marketing-digital/inbound-marketing-que-es-que-significa/

[6] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*, 2nd ed. MIT Press, 2018.

[7] C. Staff, “Machine Learning,” 2024. https://www.coursera.org/articles/types-of-machine-learning

[8] M. Rimol, “Understand 3 Key Types of Machine Learning,” 2020. https://www.gartner.com/smarterwithgartner/understand-3-key-types-of-machine-learning

[9] S. Bhatt, “Reinforcement Learning 101,” *Towars Data Science*. https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292

[10] IBM, “What is Supervised Learning.” https://www.ibm.com/topics/supervised-learning

[11] J. Delua, “What is unsupervised learning?,” *IBM*. https://www.ibm.com/blog/supervised-vs-unsupervised-learning/

[12] J. Torres, “A Gentle Introduction to Deep Reinforcement Learning,” *Towards Data Science*. https://towardsdatascience.com/drl-01-a-gentle-introduction-to-deep-reinforcement-learning-405b79866bf4

[13] DeepAi, “Understanding Attention Models in Deep Learning.” https://deepai.org/machine-learning-glossary-and-terms/attention-models

[14] I. E. Team, “Attention Model: Definition and When to use One.” https://www.indeed.com/career-advice/career-development/attention-model

[15] A. Gharahbagh, V. Hajihashemi, M. Ferreira, J. Machado, and J. Tavares, “Hybrid time-spatial video saliency detection method to enhance human action recognition systems,” *Multimed. Tools Appl.*, vol. 83, no. 14, 2024, [Online]. Available: https://link.springer.com/article/10.1007/s11042-024-18126-x

[16] R. Marois and J. Ivanoff, “Capacity limits of information processing in the brain,” *Trends Cogn. Sci*, vol. 9, no. 6, pp. 296–305, 2005.

[17] S. Fan, Z. Shen, M. Jiang, B. Koenig, and M. Kankanhalli, “Emotional Attention: From Eye Tracking to Computational Modeling,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 2, pp. 1682–1699, 2023, [Online]. Available: https://ncript.comp.nus.edu.sg/site/ncript-top/sentiment/pdf/Emotional attention From eye tracking to computational modeling.pdf

[18] “TIOBE Index.” https://www.tiobe.com/tiobe-index/ (accessed Apr. 14, 2024).

[19] “PYPL Index.” https://pypl.github.io/PYPL.html (accessed Apr. 14, 2024).

[20] H. Bhasin, *Python Basics - A Self-Teaching Introduction*. 2019.

[21] “Python.org.” https://www.python.org/

[22] “Conda,” *Conda.io*. https://conda.io/projects/conda/en/latest/user-guide/getting-started.html (accessed May 01, 2024).

[23] “FURPS+.” https://qualidadebr.wordpress.com/2008/07/10/furps/

[24] A. Lamsweerde, *Requirements Engineering: From System Goals to UML Models to Software Specifications*. Wiley, 2009.

[25] R. Pressman and MaximB., *Engenharia de Software*. Mc Graw Hill, 2011.

1. Conteúdo em anexos

Esta parte do relatório deve conter informação adicional organizada por capítulos que, embora seja interessante, não faz parte do material estritamente necessário ao relatório. Documentos importantes produzidos ou utilizados durante o estágio que, pela sua dimensão, não sejam colocáveis no corpo principal do relatório podem ser incluídos em anexos.

Referências

O capítulo de Referências apresenta a lista de Referências consultada para a execução do projeto/estágio. A lista de Referências deve estar ordenada por autor (último nome) e em seguida por ano.

No texto, sempre que utilizem dados ou afirmações de terceiros devem indicar a fonte desses dados ou a afirmação, colocando entre parêntesis curvos o nome da referência, por exemplo (Sousa, 2002).

A formatação de cada entrada bibliográfica é diferente consoante o tipo de documento em questão:

* *Para um livro:* nome(s) do(s) autor(es), ano da edição entre parêntesis, título do livro em itálico, nome da editora, local da edição, país da edição;
* *Para um artigo em revista:* nome(s) do(s) autor(es), ano da edição entre parêntesis, título do artigo em itálico, nome da revista, volume da edição a negrito, número da edição, páginas;
* *Para uma comunicação em conferência:* nome(s) do(s) autor(es), ano da conferência entre parêntesis, título da comunicação entre aspas, nome da conferência em itálico, local da conferência, país da conferência, mês da conferência;
* *Para uma tese:* nome do autor, ano da tese entre parêntesis, título da tese, tipo de tese, universidade da tese, local da universidade, país da universidade;
* *Para um relatório interno:* nome(s) do(s) autor(es), ano do relatório entre parêntesis, título do relatório, origem do relatório, referência do relatório, instituição de acesso ao relatório, local da instituição, país da instituição, mês do relatório (abreviado com 3 letras, exceção aos meses com 4 letras);
* *Para um documento extraído da Internet:* o endereço entre parênteses, incluindo a data da última visita ao documento;
* *Para um endereço da Internet:* o endereço e a data da última visita ao endereço.

Na Figura 2 é apresentada uma lista de referências com um exemplo de cada tipo referido.

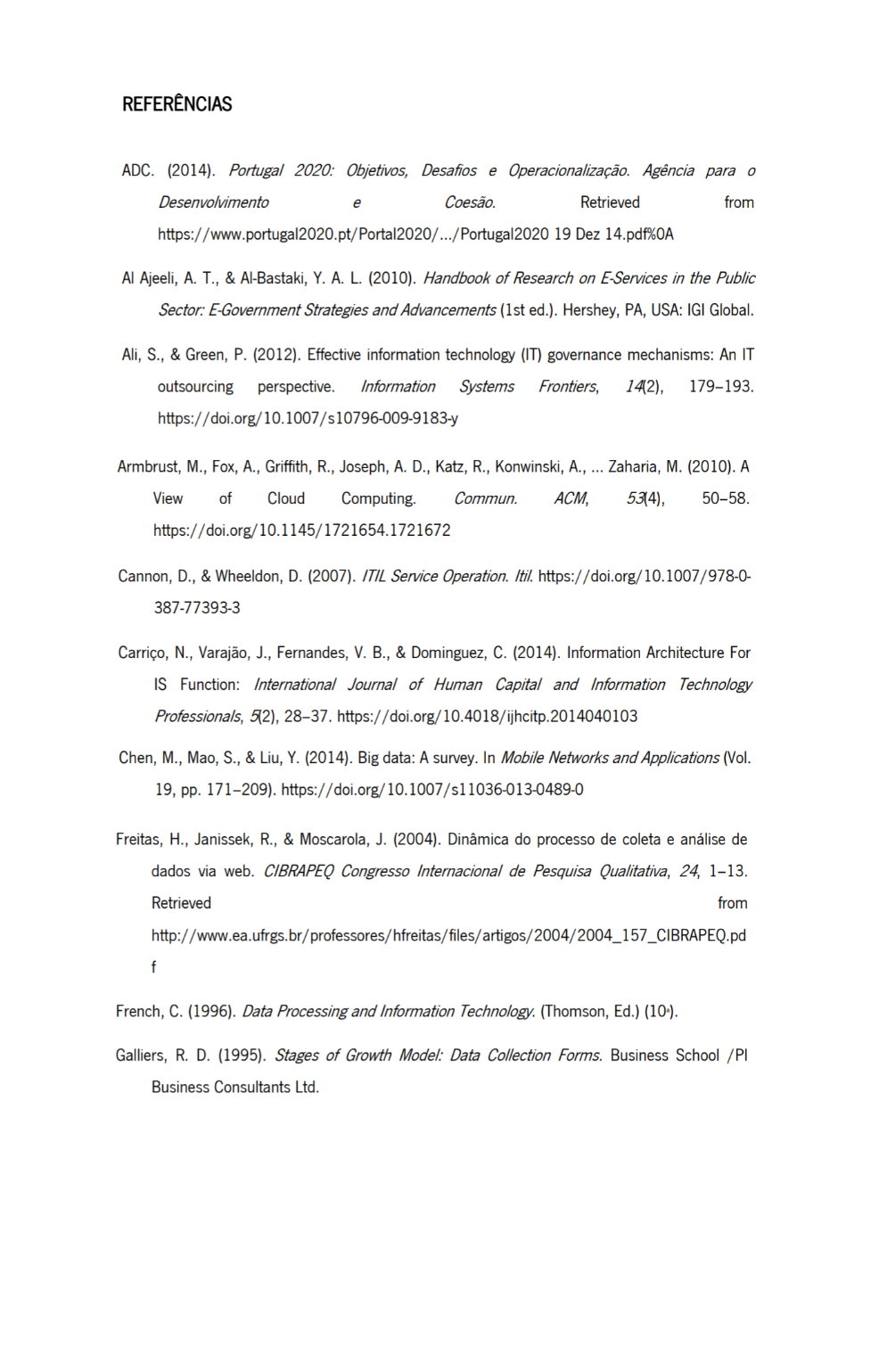


Figura 2 - Exemplo de lista de Referências.