SELEÇÃO INTELIGENTE DE RECURSOS HUMANOS COM MODELOS LLM

Dissertação apresentada à Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Bragança para obtenção do Grau de Mestre em Informática no âmbito da Dupla Diplomação com a Universidade Tecnológica Federal do Paraná.

Matheus Patriarca Santana - a61483

Trabalho realizado sob a orientação de

Prof. João Paulo Ramos Teixeira

Prof. Diego Bertolini Gonçalves

Mestrado em Informática

Outubro de 2025

SELEÇÃO INTELIGENTE DE RECURSOS HUMANOS COM MODELOS LLM

Matheus Patriarca Santana - a61483

Trabalho realizado sob a orientação de

Prof. João Paulo Ramos Teixeira

Prof. Diego Bertolini Gonçalves

Mestrado em Informática

Outubro de 2025

A Escola Superior de Tecnologia e de Gestão não se responsabiliza pelas opiniões expressas neste relatório.

Dedicatória

Dedico este trabalho a todas as pessoas que, de alguma forma, tornaram possível a realização desta jornada. Agradeço primeiramente a Deus, por me guiar, fortalecer e inspirar em cada etapa deste caminho. Ao meu orientador, PhD. João Paulo Teixeira, expresso minha profunda gratidão pela paciência, orientação e dedicação, fundamentais para a conclusão deste trabalho. Aos meus familiares, meu eterno agradecimento pelo apoio, incentivo e amor incondicional, que foram o alicerce para transformar este projeto em realidade.

Agradecimentos

Nesta etapa da minha vida acadêmica, marcada por esforço e dedicação, quero expressar minha profunda gratidão a todas as pessoas e instituições que tornaram esta jornada possível.

Agradeço, primeiramente, a Deus, por me guiar e me dar forças para superar os desafios encontrados ao longo do caminho. Ao meu irmão Alan, sou grato pelo incentivo e apoio que me motivaram a vir para Portugal, abrindo portas para novas oportunidades e experiências.

Minha sincera gratidão também vai à UTFPR, por me proporcionar esta oportunidade única de crescimento acadêmico e pessoal. Aos meus pais, agradeço pelo amor, apoio e sacrifícios que sempre me sustentaram e me deram segurança para perseguir meus sonhos.

Aos amigos que encontrei aqui em Bragança, meu muito obrigado por me fazerem sentir acolhido, trazendo companhia, risadas e momentos inesquecíveis em terras distantes.

Por fim, agradeço a todos que, direta ou indiretamente, contribuíram para que este projeto se tornasse realidade.

Resumo

O processo de triagem de currículos na área de vendas enfrenta desafios significativos devido à diversidade de formatos, terminologias e níveis de detalhamento. Para superar esta dependência da análise manual, este trabalho investigou a aplicação de modelos de linguagem natural (LLMs) para automatizar a extração e padronização de informações relevantes de currículos.

A metodologia utilizou modelos como GPT-4.1, GPT-4.1 Mini e Gemini 2.5 Pro, além da ferramenta ChatPDF e bibliotecas de apoio para a extração textual. Foram elaborados prompts específicos para estruturar atributos nominais e ordinais de forma consistente. O desempenho dos modelos foi avaliado com métricas como acurácia e erro médio aritmético. Posteriormente, o modelo GPT-4.1, que obteve o melhor desempenho, foi aplicado em um conjunto ampliado de 50 currículos para validação. Os dados extraídos foram submetidos a um modelo classificador, resultando em um Erro Médio Absoluto (MAE) de 0.76, numa escala de 10 pontos, na comparação com dados reais, o que valida a confiabilidade do método de extração para a classificação automática.

Os resultados demonstram que os modelos de linguagem natural são eficazes na extração de dados, destacando-se o modelo GPT-4.1. Conclui-se que o uso de LLMs é uma abordagem promissora para a triagem automatizada, pois reduz o esforço manual e aumenta a consistência das avaliações, tendo sido consolidada em um protótipo web integrador para demonstrar sua aplicabilidade prática.

Palavras-chave: linguagem natural; triagem de currículos; extração de dados; Ilm; engenharia de prompts.

Abstract

The process of resume screening in the sales field faces significant challenges due to the diversity of formats, terminologies, and levels of detail. To overcome the reliance on manual analysis, this study investigated the application of natural language models (LLMs) to automate the extraction and standardization of relevant information from resumes.

The methodology employed models such as GPT-4.1, GPT-4.1 Mini, and Gemini 2.5 Pro, as well as the ChatPDF tool and supporting libraries for text extraction. Specific prompts were designed to structure nominal and ordinal attributes consistently. The models' performance was evaluated using metrics such as accuracy and mean arithmetic error. Subsequently, the GPT-4.1 model, which achieved the best performance, was applied to an expanded set of 50 resumes for validation. The extracted data were submitted to a classification model, resulting in a Mean Absolute Error (MAE) of 0.76 compared to real data, which confirms the reliability of the extraction method for automatic classification.

The results demonstrate that language models are effective in data extraction, with GPT-4.1 standing out. It is concluded that the use of LLMs is a promising approach for automated screening, as it reduces manual effort and increases the consistency of evaluations , having been consolidated into an integrating web prototype to demonstrate its practical applicability.

Keywords: natural language; resume screening; data extraction; LLM; prompt engineering.

Conteúdo

Lista de Tabelas

Lista de Figuras

Siglas

API Application Programming Interface. 14

GPT Generative Pre-trained Transformer. 13, 18

LLM Large Language Model. 1, 2, 5, 8, 9, 13, 15, 17, 18

MAE Erro Absoluto Médio. 3, 16, 17, 19

PLN Processamento de Linguagem Natural. 1, 5, 11, 15, 18

RAG Retrieval-Augmented Generation. 15

Capítulo 1

Introdução

A análise de currículos é um passo fundamental nos processos de seleção de candidatos. Contudo, currículos apresentam diferenças significativas em formato, terminologia e detalhamento das informações (Barducci et al., 2022). Essa diversidade dificulta a extração automatizada de dados, exigindo interpretação contextual e técnicas de Processamento de Linguagem Natural (PLN) capazes de lidar com documentos variados (Jurafsky & Martin, 2025).

No contexto de recrutamento para cargos de vendas, a diversidade dos currículos se torna ainda mais evidente (Celsi, 2022). Alguns candidatos destacam resultados quantitativos e experiências diretas, enquanto outros enfatizam habilidades interpessoais ou experiências indiretas. Regras baseadas em expressões regulares e classificadores supervisionados frequentemente falham em lidar com documentos heterogêneos (Bhatia et al., 2019). Além disso, métodos tradicionais podem ter dificuldades em capturar a semântica completa dos currículos, o que impacta a correspondência entre candidatos e vagas (Kim et al., 2025).

Diante desse cenário, surge o problema central desta pesquisa: como extrair automaticamente informações de currículos despadronizados da área de vendas, garantindo precisão, consistência e aplicabilidade dos dados para classificação automática? A solução proposta visa preencher essa lacuna, utilizando modelos Large Language Model (LLM) capazes de interpretar textos não estruturados e capturar contextos variados (Gan

et al., 2024a). Ao aplicar técnicas de engenharia de prompts, espera-se guiar os modelos na identificação consistente de atributos-chave, transformando informações textuais em registros estruturados e codificados.

Este trabalho propõe desenvolver e avaliar um sistema automatizado baseado em LLM para extração e classificação de currículos, com potencial para reduzir esforço manual, aumentar a consistência na triagem e contribuir para processos seletivos mais eficientes.

1.1 Objetivos

Esta seção apresenta os objetivos que orientam o desenvolvimento do trabalho, descrevendo o propósito central da pesquisa e as metas específicas que direcionam sua execução.

1.1.1 Objetivo geral

Desenvolver e avaliar um método automatizado para a extração de atributos relevantes em currículos despadronizados da área de vendas, utilizando modelos de linguagem avançados e técnicas de engenharia de prompts. A abordagem visa transformar informações textuais em registros estruturados e codificados, facilitando sua aplicação em classificadores automáticos, aumentando a consistência das avaliações e reduzindo a necessidade de análise manual.

1.1.2 Objetivos específicos

Para atingir o objetivo geral, este trabalho propõe os seguintes objetivos específicos:

1. Implementar um pipeline de extração automatizada de dados para documentos em formatos PDF e DOCX, garantindo eficiência na captura de informações textuais.

2. Desenvolver prompts específicos, aplicando técnicas de engenharia de prompts para maximizar a precisão da extração.

3. Avaliar e comparar a performance dos modelos de linguagem utilizados, considerando métricas de acurácia para atributos nominais e Erro Absoluto Médio (MAE) para atributos ordinais, a fim de identificar os mais eficazes para cada tipo de dado.

4. Desenvolver uma plataforma web que consolide o processo de extração e classificação de currículos, permitindo a visualização dos dados estruturados, a aplicação automática de classificadores e a geração de notas dos candidatos.

Capítulo 2

Revisão da Literatura

Este capítulo apresenta os fundamentos teóricos que sustentam este trabalho, abordando conceitos, modelos e técnicas essenciais. Inicialmente, são discutidos os LLM, modelos de linguagem avançados que surgem a partir das técnicas clássicas de Processamento de Linguagem Natural PLN. Além disso, aborda-se a engenharia de prompts, destacando seu papel na interação com modelos de linguagem e na extração de informações de currículos. Por fim, são apresentados estudos de caso que ilustram a aplicação de métodos de Inteligência Artificial em tarefas de análise de dados textuais e automação de processos, incluindo a triagem de currículos, o desenvolvimento de frameworks multiagente e a engenharia de prompts para extração de informações estruturadas, contextualizando a relevância desta pesquisa para o domínio de seleção de Recursos Humanos.

2.1 Large Language Models

Os Large Language Models são modelos avançados de IA projetados para processar e gerar linguagem natural em larga escala. Eles surgem como uma evolução das técnicas clássicas de Processamento de Linguagem Natural, que buscam permitir que computadores compreendam e analisem textos automaticamente, identificando relações semânticas e estruturais entre palavras e frases (Jurafsky & Martin, 2025). Enquanto o PLN tradicional foca na extração de informações estruturadas e em tarefas como classificação

de textos e análise de sentimentos, os LLMs ampliam essas capacidades, lidando com textos complexos e heterogêneos, como currículos de candidatos. Baseiam-se na arquitetura Transformer, introduzida por Vaswani et al. (2017), que permite capturar relações contextuais complexas. A Figura 2.1 apresenta uma visão geral dessa arquitetura.

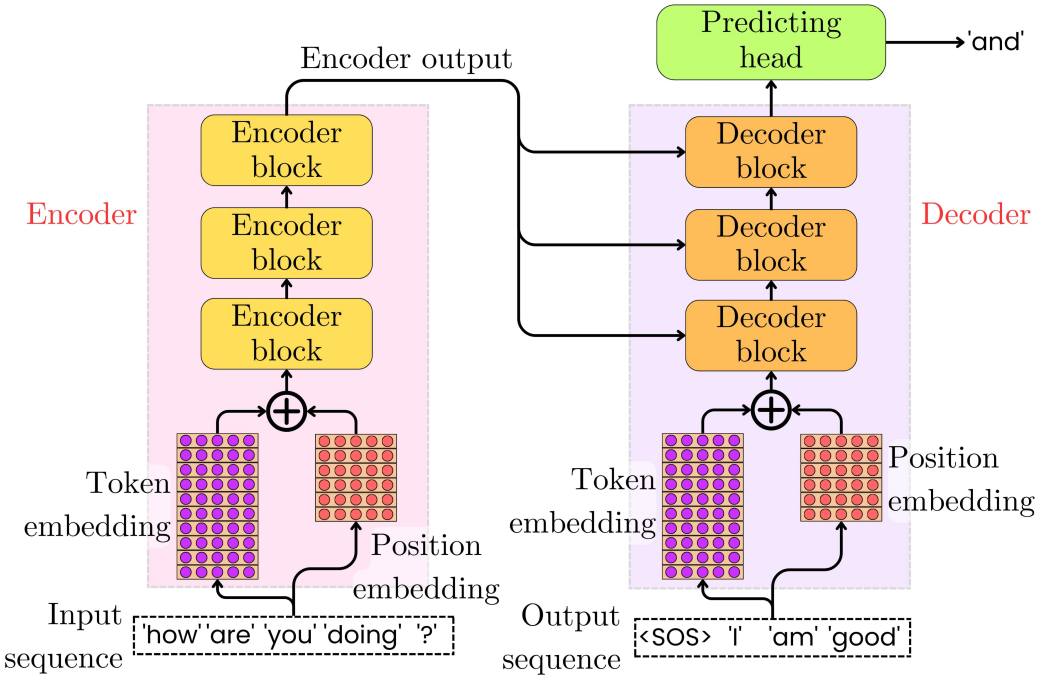


Figura 2.1: Arquitetura Transformer. Fonte: (Vaswani et al., 2017).

A arquitetura Transformer, como ilustrado na Figura 2.1, é fundamentalmente composta por um Encoder e um Decoder. O processo começa com a sequência de entrada (Input sequence), onde cada palavra ou token é convertido em um embedding de token (Token embedding) e combinado com um embedding de posição (Position embedding). Esta combinação permite que o modelo compreenda não apenas o significado das palavras, mas também sua ordem na frase.

Os dados seguem para o Encoder, que consiste em múltiplos blocos de Encoder (Encoder block) empilhados. Cada bloco processa a sequência de entrada em paralelo, capturando relações contextuais complexas através de mecanismos de autoatenção. A saída final do Encoder, uma representação contextualizada da entrada, é então passada para o Decoder.

O Decoder, também composto por blocos de Decoder (Decoder block), recebe a saída do Encoder e a sequência de saída já gerada (Output sequence), que também passa por embeddings de token e posição. O Decoder gera a próxima palavra na sequência de saída de forma autoregressiva, um token por vez. O Predicting head usa as representações do Decoder para prever o próximo token mais provável, como ‘’and‘‘ no exemplo da figura.

Exemplos notáveis de LLMs incluem:

- **BERT (Bidirectional Encoder Representations from Transformers)**: Desenvolvido pelo Google (Devlin et al., 2018), o BERT é um modelo que utiliza aprendizado bidirecional para entender o contexto de uma palavra com base em todas as outras palavras em uma frase, tanto antes quanto depois.

- **GPT (Generative Pre-trained Transformer)**: Criado pela OpenAI (Radford et al., 2018), o GPT opera de forma autoregressiva, gerando texto palavra por palavra com base no contexto anterior. Modelos subsequentes como o GPT-3 e GPT-4 expandiram essa capacidade para realizar uma vasta gama de tarefas, como sumarização, tradução e geração de código.

Modelos mais recentes expandem essas capacidades para incluir dados multimodais, ou seja, são capazes de processar e entender não apenas texto, mas também outros tipos de dados, como imagens, áudio e vídeo.

2.2 Engenharia de Prompts

A engenharia de prompts consiste no desenvolvimento de entradas textuais que orientam os LLMs a gerar respostas mais precisas e consistentes. Técnicas como o few-shot prompting e restrições de formato são amplamente utilizadas para otimizar o desempenho desses modelos. No contexto da triagem de currículos, a engenharia de prompts é particularmente relevante, pois direciona o modelo para identificar atributos específicos, como idade e experiência profissional. Essa abordagem contribui para padronizar a extração de informações e aumentar a precisão das classificações.

A Figura 2.2 apresenta um ciclo iterativo de engenharia de prompts, destacando as etapas envolvidas na otimização da interação com modelos LLM.

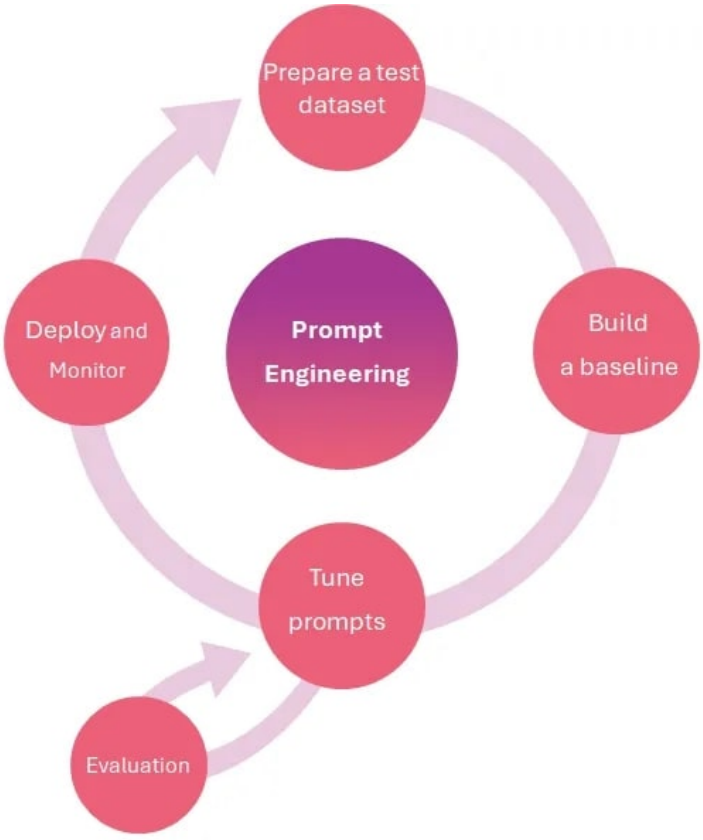


Figura 2.2: Ciclo da Engenharia de Prompts. Fonte: (Ruksha, 2024)

Conforme ilustrado na Figura 2.2, o processo de engenharia de prompts é iterativo e visa aprimorar continuamente a qualidade das interações com LLMs. Ele começa com a preparação de um conjunto de dados de teste e a construção de um baseline. Em seguida, os prompts são ajustados e refinados, seguido por uma etapa de avaliação para verificar a eficácia das modificações. Esse ciclo se repete, com a melhoria contínua dos prompts levando à implantação e monitoramento das soluções.

2.3 Estudos de Caso em Triagem Automatizada de Currículos

A seguir, apresentam-se estudos relevantes que exploram diferentes abordagens de Inteligência Artificial aplicadas à análise de dados textuais e automação de processos. Esses trabalhos destacam avanços na eficiência, na extração de informações estruturadas e na aplicação de modelos de linguagem em tarefas práticas, contextualizando a relevância desta pesquisa para o domínio de seleção de Recursos Humanos e outras aplicações relacionadas.

- O trabalho desenvolvido por Jatobá (2020) aplicou redes neurais do tipo Multi-Layer Perceptron para automatizar a triagem de currículos de candidatos à vaga de técnico de vendas. O estudo envolveu a construção de uma base de dados com 800 currículos, todos rotulados manualmente por um especialista, considerando atributos como idade, escolaridade, experiência profissional e cursos complementares. O modelo treinado apresentou um desempenho satisfatório, alcançando um MAE (Mean Absolute Error) de 0,292 no conjunto de teste.

- O trabalho desenvolvido por Gan et al. (2024b) propôs um framework multiagente para triagem automatizada de currículos, integrando LLMs a módulos especial propôs um framework multiagente para triagem automatizada de currículos, integrando LLM a um sistema de agentes. O sistema baseia-se em uma arquitetura de agente LLM com quatro componentes principais: Character, Memory, Planning, e Action. A triagem automatizada é composta por três etapas: 1) Classificação de Orações (para estruturar o texto do currículo), 2) Avaliação e Resumo (utilizando um HR Agent para atribuir notas e gerar resumos) , e 3) Tomada de Decisão (feita por um agente ou um profissional de RH). A eficácia do framework foi validada comparando as avaliações dos modelos LLM (GPT-4 e GPT-3.5-Turbo) com um conjunto de 50 currículos anotados manualmente. Os resultados mostraram que as avaliações e decisões dos LLMs se assemelham muito às de revisores humanos e que o framework

é 11 vezes mais rápido que os métodos manuais tradicionais.o.

- O estudo conduzido por Polat et al. (2025) investigou diferentes métodos de engenharia de prompts aplicados à extração de conhecimento textual de currículos. A pesquisa focou em como a formulação dos prompts impacta a capacidade de extrair informações estruturadas, como habilidades, experiências e qualificações, e vinculá-las a bases de conhecimento. Os resultados indicam que pequenas alterações na forma de escrever os prompts podem influenciar significativamente a precisão e a completude dos dados extraídos, ressaltando a importância da engenharia de prompts para otimizar a interação com LLMs em tarefas de extração de informações.

Capítulo 3

Metodologia

Este capítulo tem como objetivo descrever de forma geral os materiais utilizados e os métodos aplicados para alcançar os objetivos propostos neste trabalho. A seção de materiais aborda as ferramentas, tecnologias e ambientes de desenvolvimento empregados, enquanto a seção de métodos detalha os procedimentos adotados para a extração, processamento e análise dos dados.

3.1 Materiais

Para a execução deste trabalho, foram utilizadas tecnologias de PLN, ferramentas de análise de documentos e um ambiente de desenvolvimento que permitem experimentação prática e replicável.

3.1.1 Base de Dados

A base de dados utilizada difere daquela apresentada em (de S. Neto et al., 2025), que contém 600 instâncias organizadas em 15 atributos, mas não disponibiliza o mapeamento direto entre cada currículo original e seus respectivos registros estruturados. Essa limitação inviabiliza sua aplicação neste trabalho, já que o processo de extração automática proposto requer tanto o texto completo do currículo como entrada, quanto os atributos

correspondentes para fins de validação. Diante disso, foi construída uma base experimental composta por 50 currículos reais da área de vendas. Cada currículo foi analisado manualmente para identificar os atributos relevantes, permitindo criar um conjunto de referência.

3.1.2 Atributos de Análise

Os atributos considerados para a construção do dataset foram previamente definidos e organizados em categorias nominais e ordinais, com codificação específica para cada variável. Esses atributos serviram de base para a padronização das respostas e para a etapa de encoding aplicada nos prompts. A Tabela 3.1 apresenta a lista completa de atributos, tipos de variáveis e respectivos esquemas de codificação.

3.1.3 Extração do texto dos Currículos

Para viabilizar a leitura dos currículos em diferentes formatos e preparar os dados para a extração automática de atributos, foram empregadas bibliotecas especializadas, em Python, escolhidas pela sua robustez e capacidade de lidar com documentos heterogêneos:

- `pdfplumber`: utilizada para a extração de texto estruturado de arquivos PDF, especialmente aqueles que apresentam múltiplas colunas, tabelas ou formatação irregular. Essa biblioteca permite recuperar o conteúdo textual de forma precisa, preservando a sequência e a integridade das informações.

- `python-docx`: empregada para a leitura de arquivos no formato DOCX, garantindo a extração de parágrafos, listas e tabelas sem necessidade de conversão manual para outros formatos. Isso assegura que informações relevantes contidas em tabelas ou seções formatadas não sejam perdidas.

Além da extração, foi implementado um processo de padronização do texto, responsável por normalizar caracteres acentuados, remover símbolos especiais e uniformizar quebras de linha. Essa etapa é fundamental para que os textos dos currículos possam ser

utilizados de forma consistente como entrada para os LLMs. O pipeline de extração inclui tratamento de exceções, de modo que arquivos corrompidos ou incompatíveis não interrompam a execução. Mensagens de alerta são registradas para permitir auditoria e identificação de problemas.

Os textos extraídos de cada currículo são armazenados em uma estrutura organizada, mapeando o nome do arquivo ao conteúdo textual correspondente. Essa abordagem facilita o acesso aos dados para etapas subsequentes de processamento, como a aplicação de prompts para a extração de atributos e a validação dos resultados.

3.1.4 Modelos de Linguagem

Para a extração automática de informações dos currículos, foram selecionados LLMs capazes de lidar com textos não estruturados e gerar respostas contextualizadas. Foram utilizados Generative Pre-trained Transformer (GPT)-4.1 Mini e GPT-4.1, Gemini 2.5 Pro (Google) e ChatPDF.

GPT

Os modelos da família GPT, desenvolvidos pela OpenAI, são LLMs avançados capazes de compreender e gerar texto de forma contextualizada. Para este trabalho, foram utilizados GPT-4.1 Mini e GPT-4.1 completo, que diferem principalmente em capacidade de processamento e detalhamento das respostas.

O GPT-4.1 Mini é uma versão compacta, eficiente e de baixo custo, adequada para tarefas que exigem rapidez e menor consumo de recursos, enquanto o GPT-4.1 oferece maior capacidade de compreensão contextual, gerando respostas mais detalhadas e precisas.

Além das versões padrão, a OpenAI permite a criação de GPTs customizados, que podem ser adaptados para tarefas específicas, como priorização de certas habilidades ou experiências em currículos. Esses modelos customizados permitem ajustar o tom das respostas, incorporar conhecimento especializado e integrar o sistema com dashboards ou fluxos de recrutamento automatizados via Application Programming Interface (API). A

utilização via API baseia-se no consumo de tokens (entrada e saída), sendo pago, com custos variando conforme o modelo e a quantidade de tokens processados.

Gemini

Os modelos da linha Gemini, desenvolvidos pela Google, são modelos multimodais capazes de processar texto e outros tipos de dados de forma integrada. Para este trabalho, utilizou-se o Gemini 2.5 Pro, que oferece maior robustez e desempenho em tarefas complexas, como a análise de currículos despadronizados e documentos extensos.

O Gemini 2.5 Pro combina compreensão textual avançada com capacidade de raciocínio contextual, permitindo extrair informações estruturadas de forma eficiente. Sua arquitetura multimodal possibilita lidar com diferentes tipos de entrada e gerar respostas detalhadas, tornando-o adequado para pipelines de análise de dados que exigem integração com textos complexos ou grandes volumes de informação (DeepMind, 2025).

O acesso ao Gemini 2.5 Pro é possível via Google AI Studio, que oferece uso gratuito limitado, com restrições de quantidade de requisições e tokens processados por dia. Para volumes maiores ou uso contínuo, há necessidade de subscrição paga. Essa limitação deve ser considerada ao planejar experimentos que envolvam grandes quantidades de documentos ou consultas frequentes, uma vez que o modelo gratuito pode não suportar cargas extensas de dados sem throttling ou interrupções.

ChatPDF

O ChatPDF é uma ferramenta que possibilita a interação com documentos PDF de forma dinâmica, permitindo que o usuário faça perguntas sobre o conteúdo do documento sem precisar lê-lo linearmente. A ferramenta processa o PDF e gera embeddings do texto, possibilitando buscas semânticas rápidas dentro do documento. Isso permite identificar e recuperar informações relevantes de forma contextualizada.

O funcionamento do ChatPDF é baseado em técnicas de PLN e, em essência, segue uma abordagem similar à de Retrieval-Augmented Generation (RAG). O texto do PDF é indexado e armazenado em vetores semânticos, e quando o usuário faz uma pergunta,

o sistema recupera trechos relevantes do documento e fornece uma resposta gerada pelo modelo de linguagem com base nesses trechos. Dessa forma, o modelo não precisa memorizar todo o conteúdo do PDF, mas sim usar o contexto recuperado para gerar respostas precisas.

Apesar de sua praticidade, o ChatPDF apresenta a limitação de não realizar codificação direta dos atributos dos currículos conforme a tabela de análise. Ou seja, embora seja capaz de extrair informações do PDF, ele não converte automaticamente esses dados em registros estruturados ou nos valores ordinais/nominais previamente definidos, exigindo processamento adicional para padronização e integração com o restante do pipeline de análise.

3.1.5 Engenharia de Prompts

A engenharia de prompts foi uma etapa central para orientar os LLMs na extração precisa de atributos dos currículos. Consiste na construção de instruções cuidadosamente formuladas, capazes de direcionar o modelo a identificar, interpretar e padronizar informações específicas de forma consistente.

Cada atributo definido na Tabela 3.1 recebeu prompts específicos, considerando o tipo de dado (nominal, ordinal ou textual) e o formato de resposta esperado. Por exemplo, ao extrair a experiência em vendas, o prompt orienta o modelo a localizar menções temporais e descritivas de funções comerciais, convertendo-as em uma faixa ordinal de 0 a 8.

Os prompts foram refinados iterativamente para lidar com variações linguísticas, informações implícitas e omissões parciais, garantindo respostas mais determinísticas e uniformes. Essa prática aumenta a confiabilidade da extração, possibilita a padronização dos dados e favorece a reprodutibilidade do processo.

Embora eficaz, a engenharia de prompts apresenta limitações: informações ambíguas ou formatos inesperados nos currículos podem gerar respostas imprecisas, exigindo validação e pós-processamento. No contexto deste trabalho, os prompts funcionam como parte de um pipeline integrado, que inclui extração, padronização e codificação de atributos,

garantindo que os dados extraídos estejam prontos para a classificação automática.

3.1.6 Métricas de Avaliação

Para avaliar a precisão da extração automática dos atributos dos currículos, foram utilizadas duas métricas principais: acurácia e MAE. Esses indicadores permitiram comparar os modelos de linguagem e selecionar o mais adequado para aplicação em conjuntos de dados maiores.

Acurácia

A acurácia mede a proporção de atributos extraídos corretamente em relação ao total de atributos avaliados. Matematicamente, é definida como:

$$\text{Acurácia} = \frac{\text{Número de atributos corretamente extraídos}}{\text{Número total de atributos}} \times 100\%.$$

Essa métrica é especialmente adequada para atributos nominais ou categóricos, em que a correspondência exata entre o valor extraído e o valor correto é necessária.

Erro Médio Aritmético (MAE)

O MAE avalia a diferença média entre os valores extraídos e os valores reais, sendo útil para atributos ordinais ou numéricos. O MAE é calculado como:

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{N}|y_t - \hat{y}_t|,$$

onde N é o número total de atributos avaliados, y_t é o valor real do atributo e $\hat{y}_t$ é o valor extraído pelo modelo.

Aplicação das métricas

Inicialmente, a avaliação foi realizada em uma amostra de 20 currículos, permitindo a comparação dos modelos de linguagem com base na acurácia, MAE e visualização de

gráficos de barras para cada atributo. A partir dessa análise, o modelo com melhor desempenho foi selecionado.

Em seguida, o modelo selecionado foi aplicado ao conjunto completo de 50 currículos. Os atributos extraídos foram codificados e fornecidos como entrada ao classificador automático, responsável por gerar as notas dos candidatos a partir desses dados. Para avaliar o impacto da extração na classificação, a acurácia e o MAE foram calculados tanto em relação aos valores extraídos quanto aos resultados reais, garantindo uma avaliação consistente e abrangente do desempenho de todo o pipeline.

3.1.7 Modelo de Classificação de Currículos com Random Forest

Como parte do processo de análise, foi utilizado um modelo de Random Forest para classificação de currículos, conforme implementado em (de Souza Neto, 2025). Este modelo serviu como referência para comparar a extração das informações com o melhor modelo LLM selecionado.

O modelo recebeu como entrada os atributos codificados dos currículos (conforme descrito na Tabela 3.1) e produziu previsões de nota final para cada candidato. Seu uso permitiu validar o desempenho da classificação automática e avaliar a consistência das informações extraídas.

3.1.8 Ambiente de Desenvolvimento Web.

Para a integração final, foi desenvolvido um protótipo web integrador, que unifica os processos de extração e classificação em uma interface interativa. A arquitetura de desenvolvimento desse protótipo foi implementada utilizando as seguintes tecnologias:

- Backend (API de Extração e Classificação): O backend é desenvolvido utilizando o framework Django em Python. Esta camada é responsável por gerenciar a lógica do processo de ponta a ponta, incluindo a recepção dos arquivos de currículos, a comunicação com a API para extração e codificação dos atributos, e a alimentação do modelo de classificação para gerar a nota final.

- Frontend (Interface do Usuário): O desenvolvimento da interface de visualização, essencial para a análise centralizada e comparação dos resultados, foi realizado com a biblioteca ReactJS. Esta escolha tecnológica garantiu a construção de uma aplicação web moderna, interativa e dinâmica para a apresentação dos dados extraídos e das notas previstas.

3.2 Métodos

A metodologia adotada para a análise e extração de informações de currículos foi estruturada em oito etapas principais, conforme ilustrado na Figura 3.1:

1. A primeira etapa consistiu na coleta e extração dos dados, em que os textos foram extraídos dos arquivos de currículos.

2. Em seguida, na segunda etapa, os modelos de linguagem (LLM) foram configurados para gerar respostas determinísticas, garantindo consistência e confiabilidade na extração das informações. Foram utilizados GPT GPT-4.1-mini, GPT-4.1, Gemini 2.5 Pro e ChatPDF.

3. Na terceira etapa, foi realizada a engenharia de prompts, em que foram desenvolvidos comandos inteligentes para orientar os modelos na identificação e padronização dos atributos relevantes.

4. A quarta etapa consistiu na estruturação e codificação dos dados, transformando informações não estruturadas em dados consistentes e comparáveis, utilizando técnicas de PLN.

5. A quinta etapa envolveu a comparação dos resultados, permitindo avaliar a eficiência do processo e a qualidade da extração. Nessa fase, além da análise quantitativa, foram utilizadas visualizações gráficas, como gráficos de barras, de modo a facilitar a comparação entre os modelos e evidenciar diferenças de desempenho.

6. Na sexta etapa, após a comparação, o modelo que apresentou melhor desempenho foi selecionado e reaplicado ao conjunto ampliado de dados, possibilitando uma extração mais robusta e abrangente das informações.

7. Na sétima etapa, os dados extraídos e os dados reais dos currículos foram fornecidos como entradas ao modelo de classificação de notas, permitindo comparar os resultados obtidos a partir da extração com aqueles baseados nos dados reais, avaliando o quanto a extração se aproxima da realidade e identificando possíveis discrepâncias. Além disso, para medir a acurácia da extração, foi calculado o MAE entre os valores previstos e reais, garantindo uma análise quantitativa precisa.

8. Por fim, a oitava etapa consistiu no desenvolvimento de um protótipo web integrador, cujo objetivo foi unificar os processos de extração e classificação em uma interface interativa. Esse protótipo permitiu visualizar, comparar e analisar os resultados de forma centralizada, facilitando a interpretação dos dados e demonstrando a aplicabilidade prática da metodologia proposta.

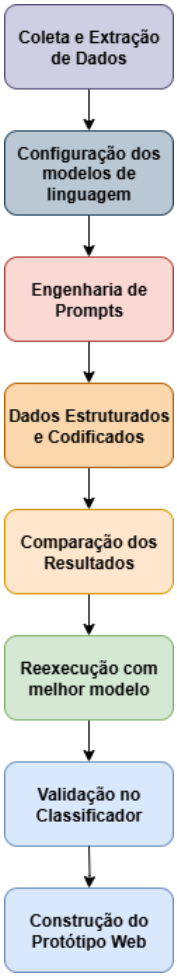


Figura 3.1: Método. Fonte: Autoria própria.

Tabela 3.1: Codificação dos atributos da base de dados original.

Atributo	Qualificação	Tipo de Variável	Codificação
Idade	Sem informação	Ordinal	0
	18 a 24 anos		1
	25 a 34 anos		2
	35 a 44 anos		3
	45 a 51 anos		4
	52 ou mais		5
Sexo	Masculino	Nominal	1
	Feminino		2
Nível Educacional	Menor que 2º grau	Ordinal	1
	2º grau completo		2
	Graduação		3
	Pós-graduação/MBA		4
	Mestrado		5
Viatura própria	Não Possui / Possui	Nominal	1 / 2
Experiência no setor de vendas	Não informado	Ordinal	0
	Sem experiência		1
	Até 6 meses		2
	Entre 6 meses e 1 ano		3
	De 1,5 a 2 anos		4
	2,5 a 4 anos		5
	5 a 9 anos		6
	10 anos ou mais		7
Experiência Tipo de função	Não se aplica	Nominal	0
	Técnico		1
	Consultor		2
	Gestor		3
Experiência em outros setores	Não informado	Ordinal	0
	Sem experiência		1
	Até 6 meses		2
	Entre 6 meses e 1 ano		3
	De 1,5 a 2 anos		4
	2,5 a 4 anos		5
	5 a 9 anos		6
	10 anos ou mais		7
Cursos em vendas	Nenhuma formação	Ordinal	1
	De 1 a 3 cursos		2
	De 4 a 6 cursos		3
	7 ou mais cursos		4
Cursos afins à área de vendas	Nenhuma formação	Ordinal	1
	De 1 a 3 cursos		2
	De 4 a 6 cursos		3
	7 ou mais cursos		4
Promoção no trabalho	Não / Sim	Nominal	1 / 2
Conhecimentos em Português	Não / Sim	Nominal	1 / 2
Conhecimentos em Inglês	Não / Sim	Nominal	1 / 2
Conhecimentos em Espanhol	Não / Sim	Nominal	1 / 2
Experiência em venda de serviços	Não / Sim	Nominal	1 / 2

Capítulo 4

Desenvolvimento

4.1 Extração de Dados

A primeira etapa consistiu na extração dos textos contidos nos currículos, os quais foram recebidos em diferentes formatos, como PDF e DOCX. Para essa finalidade, foram utilizadas distintas ferramentas e estratégias:

- **ChatPDF**: O ChatPDF é uma ferramenta que permite a análise direta de documentos, possibilitando o envio de arquivos e a interação com o seu conteúdo de forma dinâmica. A extração e o processamento do texto são realizados internamente pela API, oferecendo praticidade, mas com menor transparência sobre o funcionamento interno e dependência de uma solução externa.
- **Modelos de Linguagem (OpenAI/Gemini)**: nesses casos, os modelos não processam diretamente arquivos em PDF ou DOCX. Assim, a extração do texto foi realizada previamente com bibliotecas específicas, como pdfplumber (para PDF) e python-docx (para DOCX), antes do envio ao modelo. Essas bibliotecas permitem converter os documentos em texto bruto com maior controle do processo, embora não realizem padronização semântica nem interpretação do conteúdo. Cabe destacar que arquivos em PDF digitalizados, compostos por imagens, exigiriam técnicas de OCR, o que não foi abordado neste trabalho.

4.2 Configuração dos Modelos de Linguagem

Todos os modelos foram configurados para gerar respostas determinísticas, garantindo consistência e confiabilidade dos dados.

- **ChatGPT**: utilizado através da API da OpenAI, com `temperature = 0.0`, garantindo respostas objetivas e baseadas apenas em informações explícitas do texto.
- **Gemini 2.5 Pro**: acessado através da integração LangChain Google GenAI, também com `temperature = 0.0` para respostas determinísticas.
- **ChatPDF**: processa documentos PDF/DOCX via upload para a API, seguindo a lógica de extração estrita, sem interpretação adicional, capturando apenas dados explicitamente informados.

Essa configuração uniforme assegura que todos os atributos nominais e ordinais sejam extraídos de forma consistente, facilitando a codificação e análise subsequente.

4.3 Implementação dos Prompts

Com os textos extraídos e os atributos já definidos previamente conforme a Tabela 3.1, iniciou-se a elaboração e aplicação dos prompts. Esses prompts foram projetados para orientar os modelos de linguagem a identificar e retornar informações de interesse em um formato padronizado, reduzindo a necessidade de etapas adicionais de pós-processamento.

Foram desenvolvidos diferentes conjuntos de prompts de acordo com a ferramenta utilizada, considerando as particularidades de cada modelo:

- **Modelos LLM (OpenAI/Gemini):** os prompts foram elaborados para retornar os atributos diretamente em um formato estruturado, com o encoding aplicado no próprio prompt, conforme definido na Tabela 3.1 (por exemplo, 0, 1, 2, 3 para faixas de idade ou níveis educacionais). Esses dados estruturados e codificados foram posteriormente consolidados em um arquivo CSV contendo todos os registros extraídos.

- **ChatPDF:** devido à variabilidade e inconsistência dos resultados, não foi possível aplicar o encoding diretamente no prompt. A extração inicial foi realizada com o ChatPDF e, em seguida, os dados passaram por uma etapa de padronização e codificação separada, permitindo também a geração de arquivos CSV consolidados.

4.3.1 Técnicas de Engenharia de Prompts Utilizadas

Para garantir alta performance e consistência na extração dos dados, foram aplicadas as seguintes técnicas, com base em estudos recentes sobre prompting em modelos de linguagem:

- **Restrições de Formato (Formato rígido de saída):** Os prompts foram construídos para gerar saídas sempre no formato `Atributo: <Valor>`, facilitando a estruturação e análise dos dados, reduzindo ambiguidades e aumentando a consistência das respostas (S. S. et al., 2024).

- **Instruções de Inclusão e Exclusão (Valores booleanos e categóricos):** Cada prompt define claramente quais informações devem ser incluídas ou ignoradas, e valores booleanos ou categóricos são considerados apenas se explicitamente mencionados, evitando a geração de dados indesejados (S. S. et al., 2024).

- **Lógica de Raciocínio Implícita (Exemplos de cálculo ou interpretação):** Para tarefas que exigem múltiplos passos de interpretação ou cálculo, os prompts orientam o modelo a seguir uma sequência lógica. Exemplos podem ser fornecidos para padronizar classificações ou somar períodos de experiência, promovendo respostas coerentes (J. W. et al., 2022).

- **Few-Shot Prompting (Instrução explícita e completa):** Exemplos de entrada e saída são incluídos dentro do prompt para ajudar o modelo a compreender padrões complexos e generalizar melhor a tarefa, incluindo regras específicas para classificação (K. R. et al., 2025).

- **Escalas padronizadas para atributos numéricos ou temporais:** Para variáveis como idade ou experiência, os prompts aplicam escalas padronizadas, permitindo categorização uniforme mesmo quando os dados originais variam em formato (S. S. et al., 2024).

Nem todas as técnicas foram utilizadas em todos os prompts; a escolha dependia da complexidade do atributo a ser extraído e de testes prévios de eficácia. Atributos simples, como gênero ou posse de carro, demandaram apenas restrições de formato e instruções de inclusão/exclusão, enquanto atributos que exigem raciocínio, cálculo ou interpretação, como experiência em vendas, experiência em outros setores ou número de cursos concluídos, utilizaram também lógica de raciocínio implícita, escalas padronizadas e exemplos few-shot. Durante o desenvolvimento, algumas técnicas foram testadas em diferentes atributos e ajustadas quando se observava que não melhoravam a performance ou geravam inconsistências. Essa abordagem estratégica permitiu equilibrar precisão, consistência e eficiência na extração.

4.3.2 Prompts Específicos para Modelos LLM (OpenAI/Gemini)

desenvolvimento dos prompts para os modelos GPT-4.1, GPT-4.1 Mini e Gemini 2.5 Pro seguiu uma lógica de segmentação. Os prompts finais resultaram de um processo experimental e de otimização interativa (prompt engineering), sendo criados de forma modular e específica, um para cada categoria de atributo a ser extraída. Essa abordagem permitiu um controle mais preciso sobre cada tarefa, otimizando a acurácia para os atributos que exigiam raciocínio e cálculo.

- **Prompt para Múltiplos Atributos Nominais:** um único prompt para extrair simultaneamente vários atributos binários e nominais, como género, posse de veículo, promoções e habilidades linguísticas.
- **Prompts para Atributos Ordinais:** para atributos que exigem cálculos e classificação em escala (tempo de experiência em vendas, experiência em outros setores, quantidade de cursos, idade), um prompt individual foi elaborado para cada tarefa.

4.3.3 Prompts Utilizados

A seguir são apresentados os principais prompts elaborados para a extração dos atributos definidos na Tabela 3.1. Cada prompt foi desenvolvido considerando técnicas específicas de engenharia de prompts, de modo a garantir consistência e precisão na classificação.

Prompt para Idade Neste prompt, aplicou-se a técnica de restrição de formato (forçando a saída numérica), além de instruções de inclusão e exclusão e lógica de raciocínio implícita para calcular a idade a partir do ano de referência.

Listing 4.1: Prompt para extração da idade

```
Voce recebeu o texto de um curriculo.

Sua tarefa e:

1. Identificar a idade da pessoa explicitamente informada no
   texto, seja pela data de nascimento ou pela idade declarada.
2. Se encontrar uma data de nascimento, calcule a idade
   considerando o ano atual como dezembro de 2024.
3. Se nao encontrar nenhuma informacao clara sobre idade ou data
   de nascimento, responda com o numero 0.
4. Classifique a idade encontrada conforme a tabela abaixo e
   responda apenas com o codigo correspondente, sem qualquer outra
    explicacao:

Tabela de codificacao da idade:

0 = Sem informacao
1 = 18 a 23 anos
2 = 24 a 29 anos
3 = 30 a 35 anos
```

```
4 = 36 a 41 anos
5 = 42 a 50 anos
6 = 51 anos ou mais

Responda apenas com o numero do codigo (0 a 6).
```

Prompt para Gênero, Carro, Promoção e Idiomas Este prompt utiliza a técnica de instrução explícita e completa e saída estruturada, extraindo vários atributos em um único comando.

Listing 4.2: Prompt para genero, carro, promocao e idiomas

```
Extraia as seguintes informacoes do curriculo e responda
   exatamente no formato e opcoes abaixo, sem explicacoes
   adicionais.

Regras:
- Para cada atributo, responda SOMENTE com o numero
   correspondente as opcoes fornecidas.
- Use apenas os numeros permitidos para cada atributo. Nao use
   palavras, textos, simbolos ou outros numeros.
- Se a informacao nao estiver presente ou nao puder ser
   determinada, use o codigo para "Nao" daquele atributo (por
   exemplo: 1 para own_car e promotion_at_work; 0 para atributos
   de idioma).
- Para speak_portuguese, responda 1 se o texto do curriculo
   estiver em portugues. Caso contrario, responda 0.
```

```
- Para speak_english, responda 1 se o curriculo mencionar
   explicitamente proficiencia, fluencia, cursos, certificacoes,
   ou uso profissional do idioma ingles. Palavras-chave para
   identificar: "ingles", "English", "fluente", "avancado", "curso
    de ingles", "intermediario", "basico", "idioma", "comunicativo
   ". Caso contrario, responda 0.
- Para speak_spanish, responda 1 apenas se houver mencao
   explicita de proficiencia ou uso do espanhol.
- Caso o candidato tenha tido uma promocao no seu emprego na
   mesma empresa ou em outra empresa, classifique como 2. Caso
   contrario, classifique com 1.

Atributos e codigos:
gender: <1=Masculino, 2=Feminino>
own_car: <1=Nao, 2=Sim> - se nao houver informacao, use 1
promotion_at_work: <1=Nao, 2=Sim> - se nao houver informacao, use
    1
speak_portuguese: <0=Nao, 1=Sim>
speak_english: <0=Nao, 1=Sim>
speak_spanish: <0=Nao, 1=Sim>
experience_with_selling_services: <0=Nao, 1=Sim>

Formato de resposta (um atributo por linha, somente numero no
   valor):
gender: X
own_car: X
promotion_at_work: X
speak_portuguese: X
speak_english: X
speak_spanish: X
```

```
experience_with_selling_services: X
```

Prompt para Experiência em Vendas Aqui, a técnica aplicada foi lógica de raciocínio implícita com restrição de saída numérica, para mapear períodos de experiência em categorias pré-definidas.

Listing 4.3: Prompt para tipo de funcao

```
Voce recebera um curriculo. Sua tarefa e extrair e classificar o
   tipo de funcao do candidato baseado estritamente no que esta
   explicitamente escrito no curriculo. Nao faca suposicoes,
   inferencias ou deducoes que nao estejam claramente indicadas.

Instrucoes:
- Classifique a funcao com base somente nos titulos de cargos
   explicitos.
- Retorne apenas o codigo numerico correspondente.

experience_type_of_function:
0 = Nao aplicavel
1 = Tecnico
2 = Consultor
3 = Gerente

Exemplos:
- Titulos como "Vendedor", "Consultor de Vendas", "Representante
   Comercial", "Sales Advisor" -> classifique como 2 = Consultor
- Titulos como "Gerente de Vendas", "Sales Manager", "Supervisor
   de Equipe" -> classifique como 3 = Gerente
- Titulos como "Tecnico de manutencao", "Tecnico de informatica",
    "Field Technician" -> classifique como 1 = Tecnico
```

```
Se nenhum titulo aplicavel for encontrado ou a funcao estiver
   incerta -> classifique como 0 = Nao aplicavel
```

Prompt para Tipo de Função Neste prompt, utilizou-se **few-shot prompting implícito**, guiando o modelo com exemplos de cargos, além de saída restrita a código numérico.

Listing 4.4: Prompt para tipo de função

```
Voce recebera um curriculo. Sua tarefa e extrair e classificar o
   tipo de funcao do candidato baseado estritamente no que esta
   explicitamente escrito no curriculo. Nao faca suposicoes,
   inferencias ou deducoes que nao estejam claramente indicadas.

Instrucoes:

Classifique a funcao com base somente nos titulos de cargos
   explicitos.

A classificacao deve seguir a escala abaixo:

experience_type_of_function:
0 = Nao aplicavel
1 = Tecnico
2 = Consultor
3 = Gerente

Exemplos:

Titulos como "Vendedor", "Consultor de Vendas", "Representante
   Comercial", "Sales Advisor" ou "Consultor de Vendas" ->
   classifique como 2 = Consultor.
```

```
Titulos como "Gerente de Vendas", "Sales Manager", "Supervisor de
    Equipe" -> classifique como 3 = Gerente.

Titulos como "Tecnico de manutencao", "Tecnico de informatica", "
   Field Technician" -> classifique como 1 = Tecnico.

Se nenhum titulo aplicavel for encontrado ou a funcao estiver
   incerta -> classifique como 0 = Nao aplicavel.

Retorne apenas o codigo numerico.
```

Prompt para Experiência em Outros Setores A técnica aplicada foi **classificação categórica com escala ordinal**, similar ao prompt de experiência em vendas.

Listing 4.5: Prompt para experiencia em outros setores

```
Voce recebera um curriculo. Sua tarefa e rastrear, analisar e
   classificar a experiencia em outros setores (nao vendas)
   seguindo criterios abaixo:

Instrucoes:
- Considere todos os empregos em tempo integral, exceto aqueles
   totalmente relacionados a vendas
- Ignore apenas cargos cujo objetivo principal e vender produtos
   ou servicos, incluindo titulos como: "Vendedor", "Consultor de
   Vendas", "Representante Comercial", "Sales Advisor" e "
   Consultor"
- Inclua todas as outras experiencias, mesmo que envolvam
   atendimento ao cliente, operacoes, servicos ou suporte, desde
   que nao sejam vendas diretas
```

```
- Considere somente experiencias com datas de inicio e termino
   explicitas
- Some a duracao total (em meses ou anos) de todos os trabalhos
   nao relacionados a vendas
- Classifique a duracao total da experiencia em outros setores
   conforme a codificacao abaixo

experience_in_other_sectors:
0 = Nao informado
1 = Sem experiencia
2 = Ate 6 meses
3 = Entre 7 meses e 1 ano
4 = 1,5 a 2 anos
5 = 2,5 a 3 anos
6 = 3,5 a 5 anos
7 = 5,5 a 9 anos
8 = 10 anos ou mais

Pense passo a passo e avalie cuidadosamente cada experiencia.
Faca todos os calculos e justificativas internamente, mas nao
   exiba o raciocinio na resposta final.

Retorne SOMENTE o codigo numerico final (0 a 8), sem qualquer
   texto adicional.
```

Prompt para Cursos Concluídos Técnica aplicada: **restrição de escala padronizada** para garantir uniformidade na codificação do número de cursos concluídos.

Listing 4.6: Prompt para cursos concluidos

```
Voce recebera um curriculo. Sua tarefa e extrair e classificar
   apenas o numero de cursos concluidos relacionados a vendas,
   baseado estritamente no que estiver explicitamente escrito no
   texto. Nao faca suposicoes, inferencias ou adicoes de
   informacoes que nao estejam claramente indicadas.

Instrucoes:
- Considere apenas treinamentos ou cursos explicitamente
   relacionados a vendas (exemplos: tecnicas de vendas, negociacao
   , relacionamento com cliente, televendas, vendas no varejo,
   estrategia de vendas, CRM, etc.)
- Conte somente os cursos concluidos. Nao inclua cursos em
   andamento ou mencionados sem confirmacao de conclusao.
- Nao conte treinamentos gerais que nao sejam relacionados a
   vendas.
- Se nao houver cursos concluidos relacionados a vendas
   explicitamente listados, considere nenhum treinamento.
- Classifique o numero de cursos concluidos conforme:

sales_courses_completed:
1 = Nenhum treinamento
2 = De 1 a 3 cursos
3 = De 4 a 6 cursos
4 = 7 ou mais cursos

Retorne somente o codigo numerico.
```

Prompt para Cursos Relacionados à Área de Vendas Mesma técnica do anterior (**escala padronizada**), porém considerando apenas cursos diretamente relacionados à área de vendas.

Listing 4.7: Prompt para cursos relacionados a area de vendas

```
Voce recebera um curriculo. Sua tarefa e extrair e classificar
   apenas o numero de cursos concluidos relacionados a area de
   vendas, com base estritamente no que estiver explicitamente
   escrito no texto. Nao faca suposicoes, inferencias ou adicoes
   de informacoes que nao estejam claramente indicadas.

Instrucoes:
- Considere apenas cursos concluidos que sejam relevantes para a
   area de vendas, incluindo:
  Tecnicas de vendas, negociacao, relacionamento com o cliente,
     vendas no varejo ou por telefone
  Ferramentas de CRM, estrategia comercial, prospeccao, marketing
     , atendimento ao cliente
- O curso deve estar explicitamente listado e marcado como
   concluido
- Nao inclua cursos em andamento ou mencoes informais
- Se nenhum curso concluido relacionado a area de vendas estiver
   explicitamente listado, retorne "1"

Classifique o numero de cursos concluidos relacionados a area de
   vendas usando a escala abaixo:

sales_area_related_courses_completed:
1 = Sem treinamento
2 = De 1 a 3 cursos
3 = De 4 a 6 cursos
4 = 7 ou mais cursos

Retorne apenas o codigo numerico.
```

4.3.4 Prompt Específico para o ChatPDF

O prompt utilizado para o ChatPDF exemplifica a estratégia de extração de dados adotada. Ele foi desenvolvido como uma instrução direta e completa, listando todos os atributos desejados e definindo precisamente o formato e as regras para cada um. O prompt solicita a extração de todos os atributos previamente definidos na Tabela 3.1.

Uma característica relevante do prompt é a rigidez exigida no formato da resposta. O modelo deve retornar as informações no formato `Atributo: <Valor>`, sem explicações adicionais, permitindo que um script transforme automaticamente as respostas em um *DataFrame* para análise posterior.

Prompt Utilizado

Listing 4.8: Prompt utilizado para extracao dos atributos com o ChatPDF

```
Please extract the following information from the resume and
   respond in English in the exact format and options below, with
   no additional explanations.

Instructions:
- For language skills (speak_portuguese, speak_english,
   speak_spanish), respond 'TRUE' only if the resume explicitly
   mentions proficiency in the respective language; otherwise,
   respond 'FALSE'.
- For the age, classify into: 18 to 23 years old, 24 to 29 years
   old, 30 to 35 years old, 36 to 41 years old, 42 to 50, 51 or
   more; otherwise, respond 'No information'.
- For the sales experience, if the date was not mentioned in the
   experiences, classify as Not informed.
- For sales_experience, sum the duration of all explicitly
   mentioned sales-related jobs and classify accordingly.
```

Example: From July 2022 to December 2024 equals 2 years and 6 months.
Add all durations together to determine the total time and classify it using the scale below.

Attributes to extract:
Age: <No information, 18 to 23 years old, 24 to 29 years old, 30 to 35 years old, 36 to 41 years old, 42 to 50, 51 or more>
gender: <Male or Female>
educational_level: <Less than high school, 2nd degree completed, Bachelor's Degree, Postgraduate, MBA, or Master's Degree>
own_car: <TRUE or FALSE>
sales_experience: <Not informed, No experience, Up to 6 months, Between 7m and 1 year, 1.5 to 2 years, 2.5 to 3 years, 3.5 to 5 years, 5.5 to 9 years, 10 or more>
experience_type_of_function: <Technician, Consultant, Manager, or Not applicable>
experience_in_other_sectors: <No experience, Up to 6 months, Between 7m and 1 year, 1.5 to 2 years, 2.5 to 3 years, 3.5 to 5 years, 5.5 to 9 years, 10 or more>
sales_courses_completed: <No training, From 1 to 3 courses, From 4 to 6 courses, or 7 or more courses>
sales_area_related_courses_completed: <No training, From 1 to 3 courses, From 4 to 6 courses, or 7 or more courses>
promotion_at_work: <TRUE or FALSE>
speak_portuguese: <TRUE or FALSE>
speak_english: <TRUE or FALSE>
speak_spanish: <TRUE or FALSE>
experience_with_selling_services: <TRUE or FALSE>

Capítulo 5

Resultados e Discussão

Este capítulo apresenta os resultados da aplicação dos modelos de linguagem e ferramentas de extração de dados nos currículos analisados. A avaliação foi dividida em duas partes principais, seguindo as métricas definidas na Subseção 3.1.6 do trabalho: a acurácia para os atributos nominais e o erro médio aritmético para os atributos ordinais.

5.1 Desempenho dos Modelos na Extração de Atributos

A seguir, são detalhados os resultados obtidos por cada um dos modelos utilizados (GPT-4.1, GPT-4.1 Mini, Gemini 2.5 Pro e ChatPDF), com a performance de cada um na extração dos atributos. A Tabela 5.1 apresenta um comparativo consolidado do desempenho dos quatro modelos na extração dos diferentes atributos.

5.1.1 Resultados do GPT-4.1

O modelo GPT-4.1, demonstrou alta precisão na extração dos atributos. Os resultados para as variáveis nominais indicaram um desempenho robusto, com 100% de acurácia em atributos como Gênero, Conhecimento de Português e Inglês, e Experiência em Venda de Serviços. Para os atributos ordinais, o modelo apresentou os menores erros médios,

Tabela 5.1: Desempenho de extração por atributo para todos os modelos

Atributo	**Escala**	**GPT-4.1**		**GPT-4.1 Mini**		**Gemini 2.5 Pro**		**ChatPDF**	
		MAE	Acc (%)	MAE	Acc (%)	MAE	Acc (%)	MAE	Acc (%)
Gender	–	–	100	–	100	–	90	–	100
Educational level	–	–	94.12	–	100	–	65	–	95
Own car	–	–	94.12	–	85	–	100	–	95
Promotion at work	–	–	88.23	–	100	–	70	–	90
Speak portuguese	–	–	100	–	100	–	100	–	100
Speak english	–	–	100	–	70	–	90	–	90
Speak spanish	–	–	100	–	90	–	90	–	95
Experience with selling services	–	–	100	–	100	–	85	–	95
Age	6	0.118	–	0.35	–	0.70	–	0.75	–
Sales Experience	8	0.647	–	0.65	–	1.15	–	1.10	–
Experience type of function	4	0	–	0	–	0.65	–	1.25	–
Experience in other sectors	8	0.882	–	0.77	–	1.10	–	1.75	–
Sales courses completed	4	0.118	–	0.6	–	0.45	–	0.80	–
Sales area related courses completed	4	0.059	–	0.6	–	0.45	–	0.85	–

destacando-se na extração de idade, tempo de cursos de vendas e cursos relacionados, o que demonstra a eficácia da engenharia de prompts na interpretação precisa de informações complexas.

5.1.2 Resultados do GPT-4.1 Mini

A versão compacta do modelo, o GPT-4.1 Mini, manteve uma alta acurácia na maioria dos atributos nominais, alcançando 100% em Gênero, Nível Educacional, Promoção e experiência com venda de serviços. No entanto, houve uma queda perceptível na precisão para o reconhecimento de proficiência em inglês (70%) e espanhol (90%) e na posse de carro (85%). Em relação aos atributos ordinais, o modelo apresentou um Erro Médio Absoluto ligeiramente superior ao modelo completo, o que pode estar associado à sua capacidade de interpretação contextual ser mais limitada.

5.1.3 Resultados do Gemini 2.5 Pro

O modelo Gemini 2.5 Pro demonstrou uma performance variada. Embora tenha obtido 100% de acurácia na identificação da posse de carro e conhecimento de português, o desempenho na extração de género (90%), nível educacional (65%) e promoção no trabalho (70%) foi inferior aos modelos da OpenAI. Nos atributos ordinais, o Erro Médio Absoluto

do Gemini foi consistentemente mais alto, especialmente na experiência em vendas e em outros setores, indicando uma maior dificuldade em realizar cálculos de tempo precisos a partir dos dados não estruturados dos currículos.

5.1.4 Resultados do ChatPDF

O ChatPDF, também foi avaliado. Embora tenha alcançado alta acurácia nos atributos nominais, com 100% para Género e Conhecimento de Português, apresentou o maior Erro Médio Absoluto nos atributos ordinais. A performance em Tipo de função exercida, Tempo de experiência em outros setores e Cursos concluídos relacionados à área de vendas sugere que, apesar de ser eficiente na extração de dados simples, a ferramenta enfrenta dificuldades em tarefas que exigem raciocínio complexo ou cálculos contextuais, como a soma de períodos de tempo.

5.2 Discussão dos Resultados

A análise comparativa entre os modelos evidencia diferenças claras no desempenho de extração de atributos, tanto nominais quanto ordinais, e permite identificar fatores determinantes para a acurácia obtida. Conforme ilustrado na Figura 5.1, o GPT-4.1 se destacou como a abordagem mais robusta, apresentando 100% de acurácia em diversos atributos nominais, incluindo gênero, proficiência em português e inglês, e experiência com venda de serviços. Além disso, apresentou os menores erros médios absolutos para atributos ordinais, como idade e número de cursos concluídos, indicando que a combinação de prompts cuidadosamente estruturados com capacidade avançada de raciocínio contextual permite uma interpretação precisa de informações complexas e cálculos temporais.

A versão compacta, GPT-4.1 Mini, manteve alto desempenho na maioria dos atributos, porém apresentou quedas pontuais em reconhecimento de proficiência em idiomas e na posse de veículo. Isso sugere que, embora a versão Mini seja eficiente para tarefas de extração simples, sua capacidade de interpretação contextual e processamento de informações mais detalhadas é ligeiramente inferior à do modelo completo.

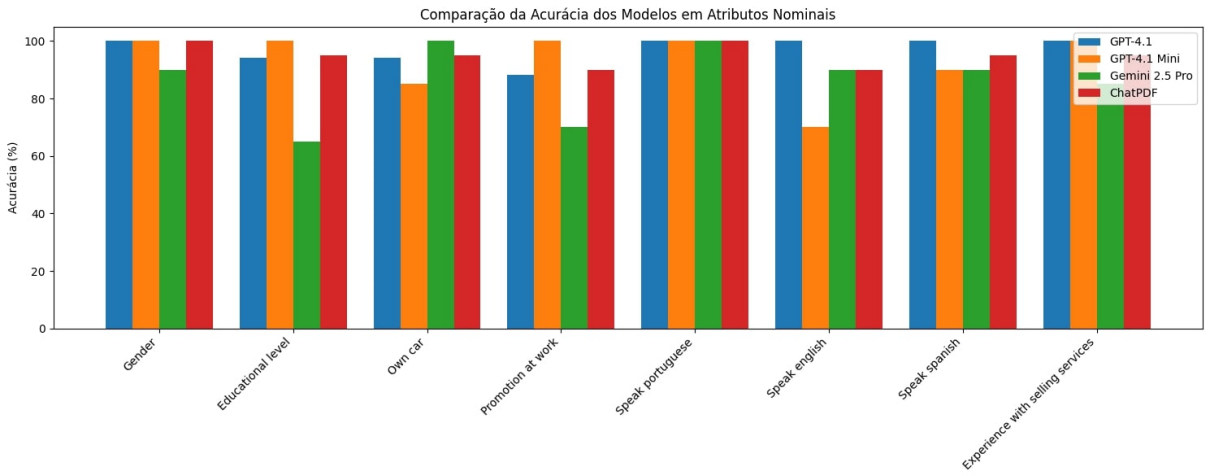


Figura 5.1: Comparação da Acurácia dos Modelos em Atributos Nominais. Fonte: Autoria própria.

O Gemini 2.5 Pro apresentou resultados mais heterogêneos. Apesar de obter alta acurácia em atributos como posse de carro e conhecimento de português, o desempenho em gênero, nível educacional e promoção no trabalho foi inferior aos modelos da OpenAI. Nos atributos ordinais, os erros médios absolutos mais elevados do Gemini são visíveis na Figura 5.2, que compara o desempenho dos modelos. Esses resultados indicam dificuldade na interpretação de informações temporais e cálculos de somatórios de experiência, possivelmente devido à menor capacidade de lidar com instruções complexas ou à necessidade de prompts mais específicos para guiar corretamente a extração.

O ChatPDF, por sua vez, mostrou-se eficiente para a extração de atributos nominais simples, alcançando alta acurácia em gênero e proficiência em português, mas apresentou os maiores erros médios absolutos nos atributos ordinais, especialmente em experiência em outros setores e cursos relacionados à área de vendas. Isso evidencia que ferramentas orientadas à análise conversacional podem ser eficazes para tarefas diretas de extração de dados, mas apresentam limitações quando a tarefa requer raciocínio contextual, cálculos ou agregação de informações de diferentes partes do documento.

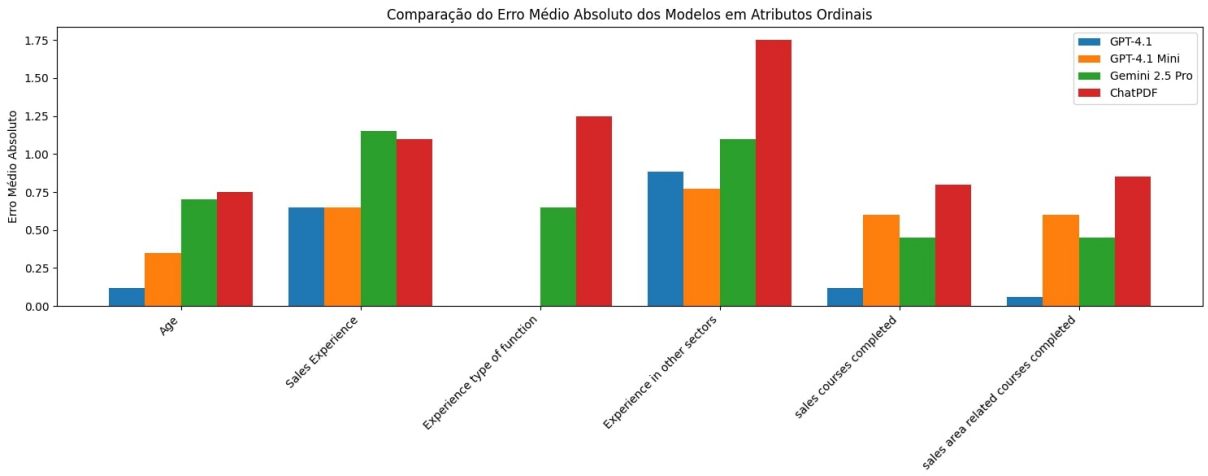


Figura 5.2: Comparação do Erro Médio Absoluto dos Modelos em Atributos Ordinais. Fonte: Autoria própria.

5.3 Validação do Modelo de Extração em Conjunto Ampliado e Classificação

Com base nos resultados apresentados na seção anterior, o modelo que demonstrou o melhor desempenho na extração de informações dos currículos foi o GPT-4.1. Para validar a precisão e a confiabilidade desta abordagem em uma escala maior, o modelo foi aplicado a um conjunto expandido de 50 currículos, de onde extraiu os atributos previamente definidos. Os resultados dessa etapa são detalhados na Tabela 5.2.

5.4 Validação Final com o Modelo Classificador

Após a validação do processo de extração com o conjunto de 50 currículos, a próxima etapa do estudo foi a de avaliar o impacto da extração automatizada na classificação final dos candidatos. Para isso, os atributos extraídos pelo modelo GPT-4.1 foram submetidos a um modelo classificador. O dataset de currículos utilizado como referência nesta análise é proveniente do trabalho de (Jatobá, 2020). Esta base de dados é composta por dados secundários de candidatos ao cargo de técnico de vendas, tendo sido construída por especialistas de Recursos Humanos no período compreendido entre janeiro de 2014 e

Tabela 5.2: Desempenho de extração por atributo

Atributo	**Escala**	**Métrica**	**Desempenho**
Gender		Acurácia (%)	100%
Educational level		Acurácia (%)	90%
Own car		Acurácia (%)	98%
Promotion at work		Acurácia (%)	92%
Speak portuguese		Acurácia (%)	100%
Speak english		Acurácia (%)	100%
Speak spanish		Acurácia (%)	100%
Experience with selling services		Acurácia (%)	90%
Age	6	Erro Médio Absoluto	0.2200
Sales Experience	8	Erro Médio Absoluto	0.8400
Experience type of function	4	Erro Médio Absoluto	0.1200
Experience in other sectors	8	Erro Médio Absoluto	0.6800
Sales courses completed	4	Erro Médio Absoluto	0.2000
Sales area related courses completed	4	Erro Médio Absoluto	0.2000

dezembro de 2018.

A classificação dos candidatos é feita numa escala de valores inteiros de 0 a 10, onde 10 representa o currículo perfeitamente adequado e de maior valor para o cargo em questão, e 0 representa o currículo de menor valorização. Nesta escala, currículos com uma pontuação igual ou superior a 7 são geralmente considerados para avançar para uma segunda etapa do processo de seleção, como, por exemplo, a entrevista.

Para a classificação, foi empregado o modelo de Random Forest implementado em (de Souza Neto, 2025). Este modelo recebeu como entrada os atributos codificados extraídos pelo GPT-4.1 e produziu previsões da nota final para cada candidato, permitindo validar o desempenho da classificação automática. O resultado da classificação gerada por este modelo, utilizando os dados extraídos, foi então comparado com a classificação de referência obtida a partir dos dados reais dos currículos.

A comparação demonstrou que o Erro Médio Absoluto (MAE) foi de 0,76. Considerando a escala de 0 a 10, isso significa que, em média, a diferença entre as notas de classificação geradas pelo sistema de extração automatizada e as notas reais foi muito pequena (inferior a 1 ponto). Este valor de MAE de 0.76 é superior ao MAE de 0,292 pontos obtido pelo modelo original de Rede Neuronal Artificial (RNA) de (Jatobá, 2020),

que utilizou esta mesma base de dados. No entanto, o erro obtido é considerado aceitável e valida a confiabilidade e a eficácia do método de extração, comprovando que ele é capaz de fornecer dados de boa qualidade para a classificação automática de currículos, com uma margem de erro que não compromete significativamente a decisão de seleção.

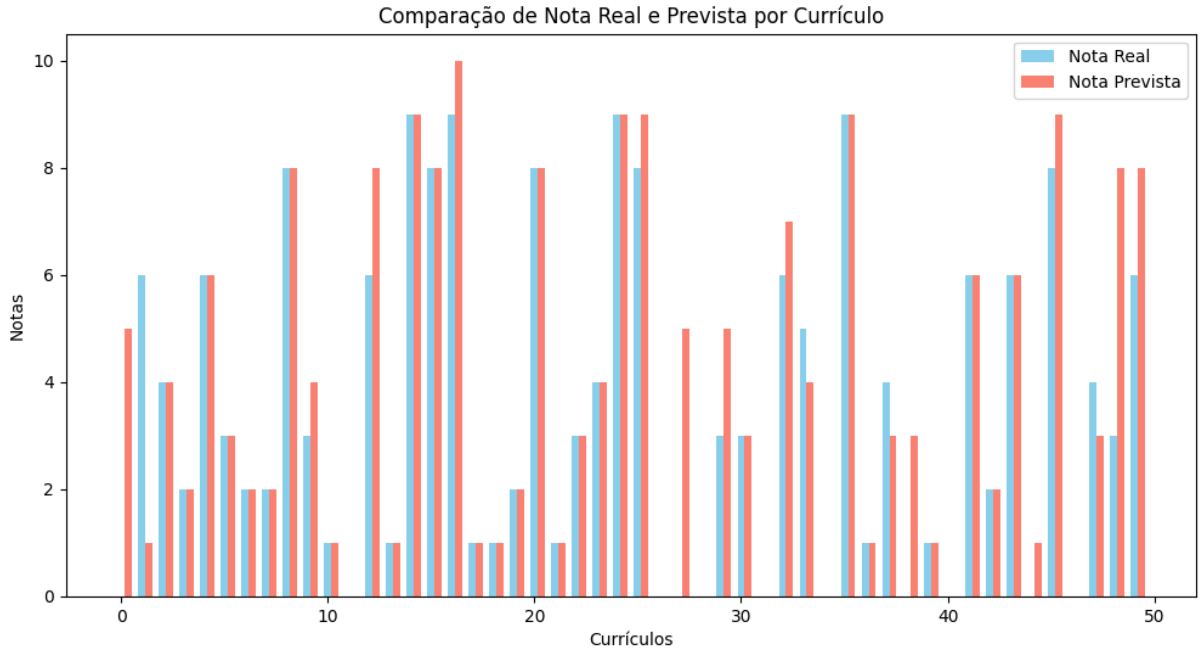


Figura 5.3: Comparação entre as notas reais e previstas pelo modelo para os 50 currículos. Fonte: Autoria própria.

Análise Visual do Gráfico de Barras

O gráfico de barras mostra, de forma clara, o motivo pelo qual o MAE é baixo e por que o modelo é considerado confiável. Para a grande maioria dos 50 currículos, a barra azul (nota real) e a barra laranja (nota prevista) apresentam alturas quase idênticas.

Essa proximidade visual entre as notas é a evidência de que, na maioria dos casos, o modelo fez uma previsão muito precisa. Embora existam alguns pontos de desvio, onde uma barra é notavelmente mais alta ou mais baixa que a outra, eles são a exceção. Isso valida visualmente a conclusão de que o sistema de extração é eficaz e que os dados extraídos são de boa qualidade para a classificação.

5.5 Plataforma de Seleção de RH Online

Após a validação da eficácia do modelo GPT-4.1 e do classificador Random Forest, o pipeline de processamento foi consolidado em um protótipo web integrador, que serve para demonstrar a aplicabilidade prática da metodologia. O desenvolvimento deste sistema ocorreu sequencialmente, após a seleção do modelo com melhor desempenho, e utiliza a seguinte arquitetura:

- Backend: Implementado com o framework Django em Python, esta camada gerencia a lógica de processamento de ponta a ponta. O sistema é acessível localmente para fins de demonstração, e a rota principal de processamento é definida em https://thesis-frontend-theta.vercel.app/. Esta rota, é responsável por:

  1. Receber o arquivo de currículo (PDF/DOCX).
  2. Orquestrar a extração do texto e a chamada ao modelo GPT-4.1 para extração e codificação dos atributos.
  3. Submeter os atributos codificados ao modelo de classificação Random Forest.
  4. Retornar a nota de classificação prevista e os atributos extraídos.

- Frontend (Interface do Usuário): O desenvolvimento da interface de visualização, essencial para o upload de currículos e a exibição interativa dos resultados, foi realizado com a biblioteca ReactJS. O frontend utiliza a biblioteca Axios para realizar requisições HTTP ao backend, enviando os arquivos de currículo e recebendo os dados processados para exibição.

O sistema retorna o resultado da classificação em formato JSON. O retorno é enriquecido, pois fornece tanto os valores codificados (utilizados como input para o classificador) quanto os valores descritivos (para interpretação humana), juntamente com a Nota Prevista. Esta estrutura de retorno dupla reforça a validade do pipeline, permitindo a rastreabilidade de como cada atributo foi extraído e codificado pelo modelo GPT-4.1. A Figura 5.4 ilustra um exemplo da estrutura de dados retornada pelo sistema.

Figura 5.4: Exemplo de estrutura JSON retornada pelo sistema, contendo atributos codificados e valores descritivos. Fonte: Autoria própria.

O protótipo web representa a concretização do sistema validado, cumprindo a oitava e última etapa da metodologia proposta. A interface do usuário, desenvolvida em ReactJS, visa facilitar a interpretação e a aplicabilidade da classificação automática por parte do profissional, oferecendo diversas funcionalidades.

A tela principal da Figura 5.5 do sistema permite a Classificação de Currículos via upload de um ou mais arquivos, acionando o pipeline completo de extração e classificação. Adicionalmente, ela suporta a visualização da última classificação realizada e o upload de

um CSV com dados previamente classificados.

Figura 5.5: Tela Home. Fonte: Autoria própria.

Após o processamento dos currículos, o sistema oferece a Exibição de Notas Individuais (Figura 5.6), listando a Nota Prevista final gerada pelo classificador para cada currículo. Nesta interface, é possível filtrar os resultados para visualização segmentada: pode-se optar por ver apenas os currículos com Nota Suficiente, que são aqueles que atingiram ou superaram a nota de corte (nota maior ou igual a 7), ou os currículos com Nota Insuficiente, que ficaram abaixo do critério mínimo (nota menor que 7). Além dessas opções, o usuário pode visualizar todos os currículos, obtendo uma visão geral de todos os resultados processados. Esta metodologia de classificação de currículos está alinhada com os conceitos apresentados em Jatobá (2020).

Figura 5.6: Seção de Exibição de Notas Individuais. Fonte: Autoria própria.

A Visualização Detalhada dos Resultados (Figura 5.7) em formato de tabela permite ao usuário alternar a exibição dos atributos: em formato descritivo para fácil leitura ou em formato codificado para análise técnica dos dados de entrada do classificador. Como pode ser visto também, o sistema permite a Exportação de Dados em formato CSV, o que é crucial para viabilizar auditorias e análises externas dos atributos extraídos e da Nota Prevista.

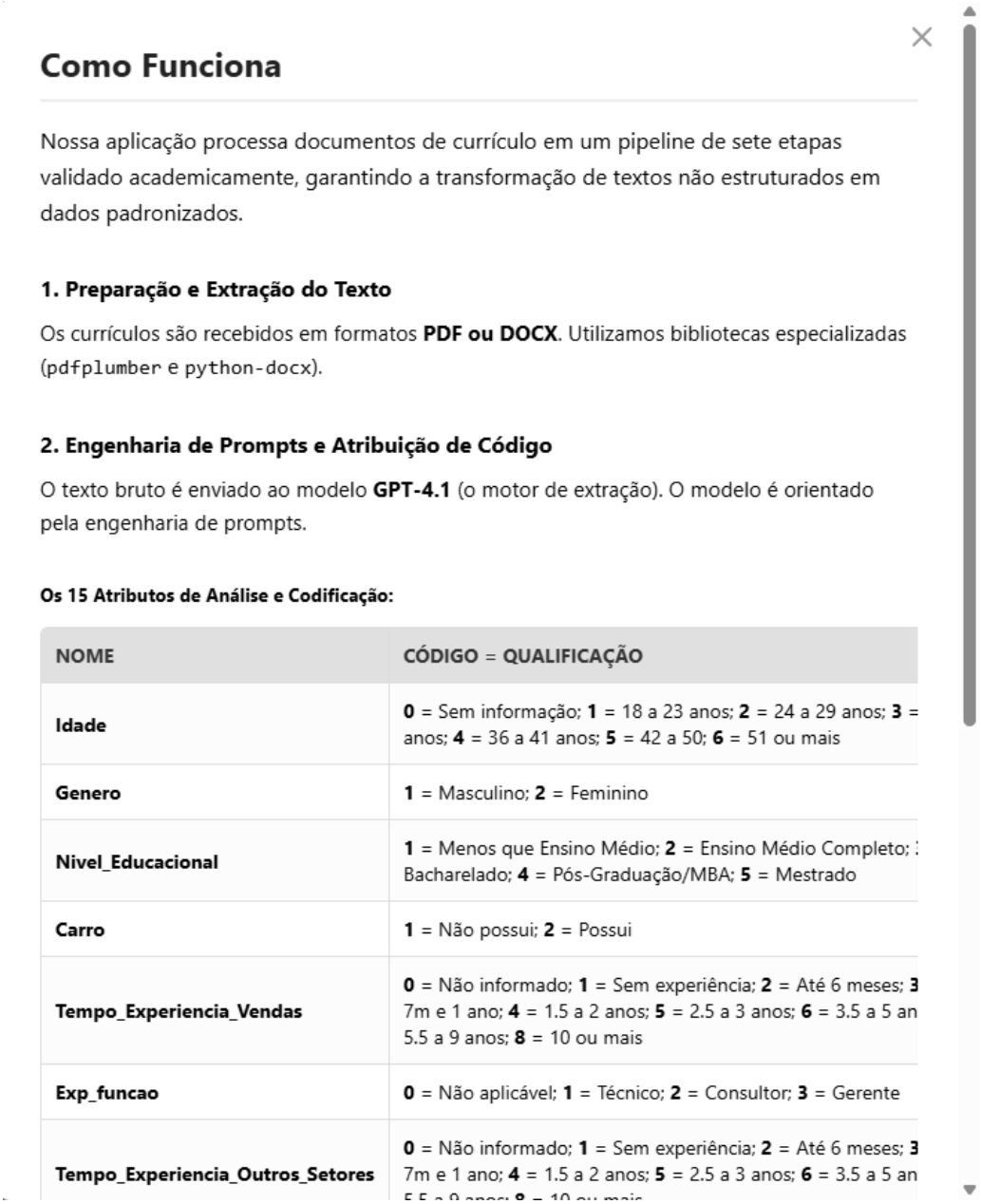
Para auxiliar na triagem e na análise do pool de candidatos, foram implementadas funcionalidades de Visualização Gráfica (Figura 5.8). O Histograma de distribuição de notas facilita a triagem rápida e a contagem de currículos por faixa de classificação. Adicionalmente, o Gráfico de Setores (Classificação do Pool de Candidatos) segmenta o grupo de currículos em faixas de classificação (ex: 'Fraco', 'Médio'), visando agilizar a análise do perfil geral do grupo de candidatos.

Por fim, também foram implementadas duas funcionalidades complementares: Info e Sobre, ambas acessíveis diretamente no menu principal do sistema. A funcionalidade de

Figura 5.7: Visualização Detalhada dos Resultados. Fonte: Autoria própria.

Figura 5.8: Funcionalidades de Visualização Gráfica. Fonte: Autoria própria.

Info apresenta informações institucionais e contextuais sobre o projeto, conforme ilustrado na Figura 5.9. Nessa seção, o usuário tem acesso a uma explicação detalhada sobre o propósito da aplicação, sua base científica e a equipe responsável pelo desenvolvimento. Já a funcionalidade Sobre fornece uma visão técnica do funcionamento interno do pipeline de classificação, como mostrado na Figura 5.10.



Como Funciona

Nossa aplicação processa documentos de currículo em um pipeline de sete etapas validado academicamente, garantindo a transformação de textos não estruturados em dados padronizados.

1. Preparação e Extração do Texto

Os currículos são recebidos em formatos **PDF ou DOCX**. Utilizamos bibliotecas especializadas (`pdfplumber` e `python-docx`).

2. Engenharia de Prompts e Atribuição de Código

O texto bruto é enviado ao modelo **GPT-4.1** (o motor de extração). O modelo é orientado pela engenharia de prompts.

Os 15 Atributos de Análise e Codificação:

NOME	CÓDIGO = QUALIFICAÇÃO
Idade	**0** = Sem informação; **1** = 18 a 23 anos; **2** = 24 a 29 anos; **3** = anos; **4** = 36 a 41 anos; **5** = 42 a 50; **6** = 51 ou mais
Genero	**1** = Masculino; **2** = Feminino
Nivel_Educacional	**1** = Menos que Ensino Médio; **2** = Ensino Médio Completo; Bacharelado; **4** = Pós-Graduação/MBA; **5** = Mestrado
Carro	**1** = Não possui; **2** = Possui
Tempo_Experiencia_Vendas	**0** = Não informado; **1** = Sem experiência; **2** = Até 6 meses; **3** 7m e 1 ano; **4** = 1.5 a 2 anos; **5** = 2.5 a 3 anos; **6** = 3.5 a 5 an 5.5 a 9 anos; **8** = 10 ou mais
Exp_funcao	**0** = Não aplicável; **1** = Técnico; **2** = Consultor; **3** = Gerente
Tempo_Experiencia_Outros_Setores	**0** = Não informado; **1** = Sem experiência; **2** = Até 6 meses; **3** 7m e 1 ano; **4** = 1.5 a 2 anos; **5** = 2.5 a 3 anos; **6** = 3.5 a 5 an

Figura 5.9: Seção de Informações do Sistema (*Info*). Fonte: Autoria própria.

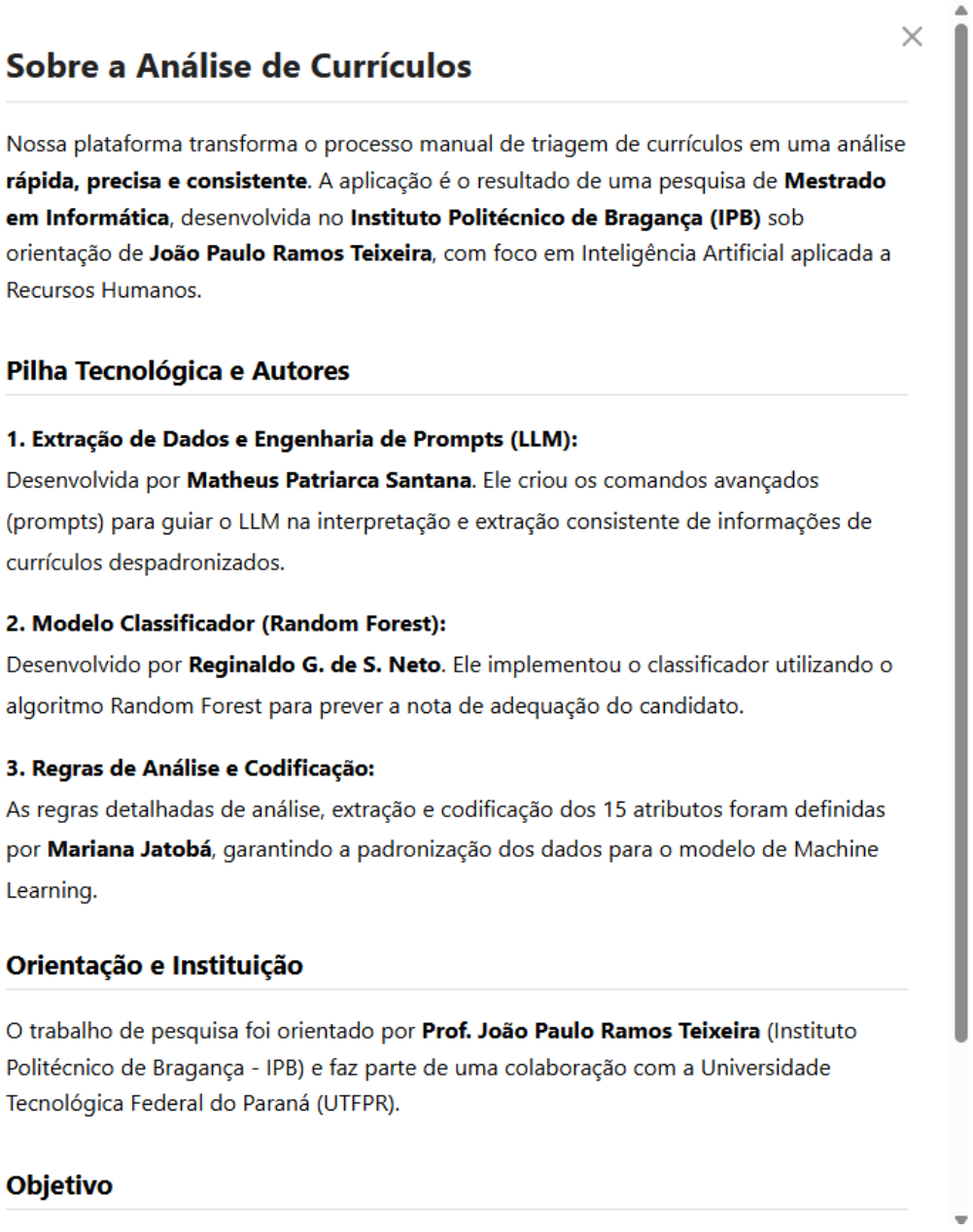


Figura 5.10: Seção Técnica de Funcionamento (*Sobre*). Fonte: Autoria própria.

Capítulo 6

Conclusões

O presente trabalho teve como objetivo desenvolver um sistema automatizado para triagem e análise de currículos na área de vendas, utilizando modelos de linguagem avançados (GPT-4.1-mini, GPT-4.1 e Gemini 2.5 Pro) e ferramentas de extração de dados de documentos (ChatPDF), todos integrados em um ambiente de execução controlado no Google Colab. A pesquisa fundamentou-se em conceitos de processamento de linguagem natural, engenharia de prompts e análise de métricas de desempenho, aplicando-os a um contexto real de seleção de candidatos.

Durante o desenvolvimento, foi possível observar que o uso de modelos de linguagem LLMs proporcionou maior eficiência na extração de informações de currículos despadronizados, reduzindo a dependência de avaliação manual e permitindo a padronização dos dados para classificação automática. A aplicação de métricas como acurácia e erro médio aritmético possibilitou avaliar a performance dos modelos de forma objetiva, identificando vantagens e limitações de cada abordagem. Entre os pontos positivos, destacam-se a capacidade de lidar com diferentes formatos de documentos, a escalabilidade do processo e a reprodutibilidade proporcionada pelo ambiente Colab. Por outro lado, limitações como a necessidade de ajustes finos nos prompts e pequenas inconsistências na extração de informações mais complexas foram enfrentadas e mitigadas por meio de testes iterativos e ajustes de codificação.

Após a avaliação inicial, o modelo GPT-4.1 foi selecionado por demonstrar o melhor

desempenho na extração de atributos. Em seguida, este modelo foi reaplicado a um conjunto expandido de 50 currículos, extraindo os mesmos atributos definidos anteriormente. As notas previstas pelo classificador construído a partir desses dados foram então comparadas com as notas reais dos candidatos, obtendo-se um Erro Médio Absoluto (MAE) de 0,76. Esse resultado evidencia que o modelo conseguiu prever as notas com alta precisão na grande maioria dos casos, validando a eficácia do sistema de extração e classificação automatizada. Para demonstrar a aplicabilidade prática, o modelo de melhor desempenho foi implementado em uma plataforma web. Nessa plataforma, é possível inserir currículos em diversos formatos e obter, automaticamente, a extração dos atributos e a classificação final, consolidando a solução desenvolvida.

O trabalho contribui para a comunidade acadêmica e profissional ao demonstrar a aplicabilidade prática de técnicas de PLN e modelos LLMs em tarefas de triagem de currículos, mostrando que conceitos aprendidos em sala de aula podem ser aplicados em problemas reais de forma eficiente. Em síntese, os objetivos propostos foram alcançados, demonstrando que a combinação de modelos de linguagem avançados, ferramentas de extração de dados e métricas de avaliação constitui uma abordagem eficaz e prática para a triagem automatizada de currículos.

Como trabalhos futuros, é sugerido aprimorar a generalização do sistema e a flexibilidade da plataforma. A metodologia de extração baseada em LLMs e engenharia de prompts é potencialmente aplicável a qualquer área de recrutamento; para ser universal, exigiria a definição de um novo conjunto de atributos e a criação de prompts específicos para cada contexto. Além disso, visando aumentar a robustez da plataforma, a implementação de uma funcionalidade para que o profissional de RH possa realizar a entrada manual e a codificação dos atributos de currículos, ou mesmo editar os dados extraídos, seria extremamente valiosa. Isso garantiria a utilidade da plataforma mesmo em casos raros de falha na extração automática e facilitaria a integração de dados provenientes de outras fontes, conferindo maior controle e adaptabilidade ao processo de triagem.

Bibliografia

et al., J. W. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. Em: Em *Advances in Neural Information Processing Systems. 35*. Curran Associates, Inc., 2022, 24824–24837. https://doi.org/10.48550/arXiv.2201.11903

et al., K. R. (2025). Prompt Engineering Guidelines for Using Large Language Models in Requirements Engineering. *arXiv preprint arXiv:2507.03405*. https://doi.org/10.48550/arXiv.2507.03405

et al., S. S. (2024). The Prompt Report: A Systematic Survey of Prompting Techniques. *arXiv preprint arXiv:2406.06608*. https://doi.org/10.48550/arXiv.2406.06608

Barducci, L., et al. (2022). Information extraction from heterogeneous CVs: Approaches and evaluation. *Expert Systems with Applications*, *200*, 117057. https://doi.org/10.1016/j.eswa.2022.117057

Bhatia, V., Rawat, P., Kumar, A., & Shah, R. R. (2019). End-to-End Resume Parsing and Finding Candidates for a Job Description using BERT. *arXiv preprint arXiv:1910.03089*.

Celsi, L. R. (2022). HR-Specific NLP for the Homogeneous Classification of Hard Skills. *Journal of Organizational Computing and Electronic Commerce*, *32*(1), 1–19. https://doi.org/10.1080/08839514.2022.2145639

de S. Neto, R. G., Jatobá, M. N., Santana, M., Fernandes, P. S., Ferreira, J. J., Foleis, J. H., & Teixeira, J. P. (2025). Human Resources Sptimization with MultiLayer Perceptron: An Automated Selection Tool [CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social

Care Information Systems and Technologies]. *Procedia Computer Science*, *256*, 238–245. https://doi.org/https://doi.org/10.1016/j.procs.2025.02.117

DeepMind, G. (2025). *Gemini 2.5 Pro* [Acesso em: 15 out. 2025]. https://deepmind.google/models/gemini/pro/

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.

Gan, C., Zhang, Q., & Mori, T. (2024a). Application of LLM Agents in Recruitment: A Novel Framework for Resume Screening. *arXiv preprint arXiv:2401.08315*.

Gan, C., Zhang, Q., & Mori, T. (2024b). Application of LLM Agents in Recruitment: A Novel Framework for Resume Screening. *arXiv preprint arXiv:2401.08315*.

Jatobá, M. N. (2020). *Inteligência artificial no recrutamento & seleção: inovação e seus impactos para a gestão de recursos humanos* [Tese de Mestrado]. APNOR e UNIFACS-Universidade Salvador [Orientadores: Paula O. Fernandes, João Paulo Teixeira, Daniela Moscon.]. http://hdl.handle.net/10198/21805

Jurafsky, D., & Martin, J. H. (2025). *Speech and Language Processing (3rd Edition)* (3ª ed.). Stanford University. https://web.stanford.edu/~jurafsky/slp3/

Kim, T., Zhang, Y., & Chen, L. (2025). Resume2Vec: Transforming Applicant Tracking Systems with Intelligent Resume Embeddings. *Electronics*, *14*(4), 794. https://doi.org/10.3390/electronics14040794

Polat, F., Tiddi, I., & Groth, P. (2025). Testing prompt engineering methods for knowledge extraction from text [Mencionado como Tiwari et al. no contexto original]. *Semantic Web*.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training* (rel. téc.). OpenAI. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

Ruksha, K. (2024). *Prompt Engineering: Classification of Techniques and Prompt Tuning* [Acesso em: 30 set. 2025]. https://www.godeltech.com/blog/prompt-engineering-classification-of-techniques-and-prompt-tuning/

de Souza Neto, R. G. (2025). *Explicabilidade em Modelos de IA Aplicados à Seleção de Recursos Humanos* [Dissertação de Mestrado]. Escola Superior de Tecnologia e de Gestão de Bragança [Trabalho realizado sob a orientação de Prof. João Paulo Ramos Teixeira e Prof. Juliano Henrique Foleis].

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. Attention is all you need. Em: Em *Advances in Neural Information Processing Systems. 30.* 2017. https://arxiv.org/abs/1706.03762

Apêndice A

Proposta Original do Projeto

Proposta de tema para
Dissertação/Estágio/Projeto - Trabalho de Conclusão de Curso

Orientador da Instituição onde se realiza o trabalho:

João Paulo Teixeira	joaopt@ipb.pt

Instituição do orientador:

Instituto Politécnico de Bragança	Escola Superior de Tecnologia e Gestão

Co-orientador da Instituição parceira:

Diego Bertolini	diegobertolini@gmail.com

Instituição do co-orientador:

Universidade Tecnológica Federal do Paraná	Campus Campo Mourão

Curso ou cursos da Instituição do orientador onde se propõe que o trabalho seja realizado:

Mestrado em Informática
Aluno: Matheus Patriarca Santana (matheussantanapatriarca@gmail.com; msantana@alunos.utfpr.edu.br)

Título do trabalho:

Intelligent Human Resources Selection with Large Language Models

Palavras chave:

Inteligência Artificial, Sistemas de Classificação Inteligentes, Seleção de Recursos Humanos, LLM

Objetivos:

Explorar modelos de IA generativa baseados em Large Language Models (LLM), nomeadamente a sua arquitetura, plataformas onde possam ser desenvolvidos, e modelos disponíveis para re-treinar. Pretende-se explorar em que forma podem ser usados para a classificação/seleção de recursos humanos com base na análise de CVs.

Descrição adicional:

O aluno será integrado num ambiente de investigação com investigadores experientes na área de recursos humanos e na área de inteligência artificial nos centros de investigação CeDRI e UNIAG, do IPB. Terá disponível um dataset de CV de candidatos ao cargo de consultor de vendas. Pretende-se que o aluno explore e experimente modelos LLM para classificação dos CVs.
Observar o eventual desbalanceamento da base de dados.
Para a análise dos resultados deverá explorar métricas como a matriz de confusão, curvas ROC e Scatter Plots, e outras medidas de desempenho.
No final do trabalho espera-se que o aluno tenho um domínio abrangente sobre modelos LLM.

Metodologia/Plano de trabalhos:

1. Estudo do estado da arte sobre LLM (2 month)
2. Tomada de conhecimento dos objetivos do trabalho pelo contacto com o dataset disponível e dos objetivos do trabalho. (1 month)
3. Exploração da utilização dos modelos e sua adaptação para o objetivo de classificar CVs (2 months)
4. Desenvolvimento dos modelos (3 months)
5. Aplicação de diversas métricas para avaliação dos resultados, como exatidão, precisão, matriz de confusão, curvas ROC e Scatter Plots. (1 month)
6. Escrita da dissertação. (1 month)

Recursos necessários:

Computador
Dataset
(todos disponíveis para o trabalho)