

AutoCV: Transformando Perfis do LinkedIn em Currículos Inteligentes

Marcos Carvalho Corrêa Junior

¹Instituto de Informática – Universidade Presbiteriana Mackenzie (MACKENZIE)

10401844@mackenzista.com.br – RA: 10401844

Abstract. *The project proposes the development of an Artificial Intelligence-based tool capable of automatically transforming LinkedIn profiles into structured résumés. By extracting data via the LinkedIn API, the information will be processed, classified, and organized to create an optimized and standardized résumé. The proposal aims to streamline and simplify the résumé creation process, assisting professionals in their job search and companies in their recruitment processes.*

Resumo. *O projeto propõe o desenvolvimento de uma ferramenta baseada em Inteligência Artificial capaz de transformar perfis do LinkedIn em currículos estruturados de forma automática. A partir da extração de dados via API do LinkedIn, as informações serão tratadas, classificadas e organizadas para compor um currículo otimizado e padronizado. A proposta visa agilizar e facilitar o processo de criação de currículos, auxiliando profissionais em busca de oportunidades e empresas em seus processos seletivos.*

Introdução

Com o avanço das tecnologias digitais, a forma como profissionais se apresentam ao mercado de trabalho evoluiu. O LinkedIn tornou-se uma das principais plataformas para exposição de experiências, formações e habilidades. Apesar disso, o currículo em formato PDF continua sendo amplamente exigido por empresas e recrutadores. Muitos profissionais, no entanto, enfrentam dificuldades para transformar seus perfis em currículos bem estruturados, seja por falta de tempo ou conhecimento em formatação. Esse projeto propõe o desenvolvimento de uma ferramenta que automatize esse processo, utilizando a API do LinkedIn e técnicas de Inteligência Artificial para tratar e organizar os dados de forma clara e padronizada. A escolha do tema se deu pela aplicabilidade prática da solução e pela oportunidade de unir tecnologia à resolução de um problema comum no mercado de trabalho.

Descrição do Problema

Atualmente, não há uma forma prática e eficiente de converter perfis do LinkedIn em currículos bem estruturados. As ferramentas disponíveis geralmente apenas copiam e colam as informações, muitas vezes posicionando dados em locais inadequados, sem qualquer tratamento ou organização. Além disso, essas soluções não lidam bem com erros ou inconsistências nos perfis dos usuários, o que compromete a qualidade do currículo final. Isso gera documentos confusos e pouco profissionais, dificultando a avaliação por recrutadores. Diante disso, há uma clara necessidade de uma ferramenta mais inteligente, capaz de interpretar e organizar os dados de forma eficiente e automatizada.

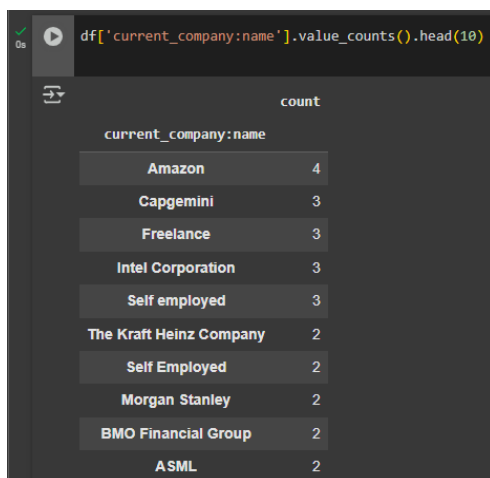
Aspectos Éticos do Uso da IA

É fundamental garantir a privacidade e a segurança dos dados dos usuários, especialmente por se tratar de informações pessoais e profissionais sensíveis extraídas do LinkedIn. O acesso à API deve seguir todas as diretrizes da plataforma, com o consentimento explícito do usuário. Além disso, a IA deve ser programada para tratar os dados com imparcialidade, sem distorcer ou excluir informações importantes, garantindo uma representação fiel do perfil profissional. Outro ponto importante é a transparência: o usuário deve saber como seus dados estão sendo utilizados e ter controle sobre o resultado gerado. Por fim, o desenvolvedor da solução tem a responsabilidade de garantir que a ferramenta não seja usada de forma indevida, como na criação de currículos falsos ou manipulados, promovendo sempre a ética e a integridade no uso da tecnologia.

Dataset e Análise Exploratória

Para uma primeira análise exploratória dos dados, o projeto utiliza o LinkedIn Professional Profiles Dataset, disponível publicamente na plataforma Kaggle. Esse conjunto de dados reúne informações extraídas de perfis profissionais do LinkedIn, como nome, cargo atual, empresa empregadora, histórico de experiências, formação acadêmica e localização geográfica. Quando necessário, os dados são anonimizados para garantir a privacidade dos indivíduos e atender às normas de proteção de dados.

Na etapa inicial da análise exploratória do dataset, foi possível identificar as empresas mais recorrentes entre os perfis analisados, com destaque para Amazon, Capgemini, Intel Corporation e freelancers, que aparecem com maior frequência como empregadores atuais. Também foi constatado que o conjunto de dados abrange profissionais de 83 países diferentes, demonstrando uma ampla diversidade geográfica. Essas informações são fundamentais, pois reforçam a riqueza e a representatividade do dataset, que será utilizado para o treinamento da Inteligência Artificial responsável por transformar perfis do LinkedIn em currículos estruturados de forma automatizada.



```
df['current_company:name'].value_counts().head(10)
```

current_company:name	count
Amazon	4
Capgemini	3
Freelance	3
Intel Corporation	3
Self employed	3
The Kraft Heinz Company	2
Self Employed	2
Morgan Stanley	2
BMO Financial Group	2
ASML	2



```
[12] # LISTA PAÍSES DIFERENTES NO DATASET
df['country_code'].unique()

array(['CA', 'NL', 'UK', 'IT', 'ES', 'CL', 'US', 'GH', 'GB', 'IN', 'MA',
       'CD', 'AM', 'FR', 'IE', 'AU', 'SG', 'BE', 'TR', 'SE', 'ZA', 'CR',
       'AE', 'CN', 'CO', 'JP', 'PK', 'LV', 'SD', 'MY', 'IR', 'BR', 'ID',
       'QA', 'UY', 'DE', 'VN', 'VE', 'MX', 'CI', nan, 'SA', 'PL', 'GE',
       'DK', 'GR', 'AT', 'BS', 'SK', 'IL', 'DZ', 'NZ', 'CZ', 'LK', 'EG',
       'BA', 'NO', 'SY', 'SN', 'GH', 'PH', 'HK', 'AR', 'KR', 'RO', 'ZM',
       'TW', 'BH', 'BG', 'NG', 'RU', 'OM', 'PE', 'TH', 'BN', 'PT', 'GM',
       'RS', 'IQ', 'KM', 'FI', 'MO', 'MK'], dtype=object)

# QTDE PAÍSES DIFERENTES NO DATASET
len(df['country_code'].unique())

83
```

Metodologia e Resultados Esperados

A metodologia do projeto será dividida em três etapas principais. A primeira consiste na integração com a API do LinkedIn, por meio da qual serão coletados os dados do perfil do usuário, como experiências profissionais, formação acadêmica, habilidades e certificações. Em seguida, será aplicada uma camada de tratamento com técnicas de Inteligência Artificial e Processamento de Linguagem Natural (PLN), responsável por classificar, corrigir inconsistências e organizar as informações com base em um modelo de currículo ideal. Por fim, os dados processados serão formatados automaticamente em um modelo de currículo estruturado e personalizável, pronto para ser exportado em formatos como PDF. Os resultados esperados incluem a criação de uma ferramenta

funcional, que gere currículos organizados, coerentes e profissionais com base nos dados do LinkedIn, reduzindo significativamente o tempo e o esforço dos usuários, além de melhorar a qualidade dos currículos apresentados em processos seletivos.