**Know More About Company & Job**

# **Introduction**

This project aims to address a series of challenges that individuals may face when searching for employment opportunities. In the current landscape, it can be daunting for talent to discern which companies align with their skills and aspirations. Additionally, when applying for a specific position, candidates often lack insights into how well their qualifications match the job requirements.

Our project seeks to provide solutions to these issues by leveraging AI technologies. We aim to revolutionize how company summaries are presented to make them more relevant and engaging for potential talent. By analyzing job descriptions and candidate profiles, we'll develop tools to provide detailed assessments, highlighting areas where candidates excel and identifying areas for improvement.

Furthermore, we're exploring the possibility of utilizing AI-generated content to craft personalized cover letters, streamlining the application process and increasing the chances of success for candidates.

In essence, our project endeavors to empower talent by providing them with the necessary information and tools to make informed decisions about their career paths and to enhance their chances of securing their desired positions.

# **Data Design**

I have compiled a dataset consisting of one resume and eight job descriptions sourced from LinkedIn. The resume encompasses essential details such as educational background, work experience, and a summary of the talent's skills and expertise focus on data science and machine learning. The job descriptions span various fields, including nursing, construction engineering, teaching, database administration and other roles relevant to data science.

Utilizing the OpenAI model, we aim to evaluate the degree of alignment between the provided resume and each job description. By leveraging advanced natural language processing techniques, we seek to assess the suitability of the talent for each job, identifying matches and areas of divergence. This process will enable us to provide valuable insights into which job roles best complement the talent's qualifications and aspirations within the diverse landscape of employment opportunities.

# **Application Development**

The entire application was developed using Python in Google Colab, incorporating various packages such as Langchain, PyPDF, and OpenAI. Throughout development, both GPT-3.5-turbo and GPT-4-turbo models were tested for their performance. Overall, the GPT-4-turbo model demonstrated superior performance compared to its predecessor, providing more accurate and effective results.

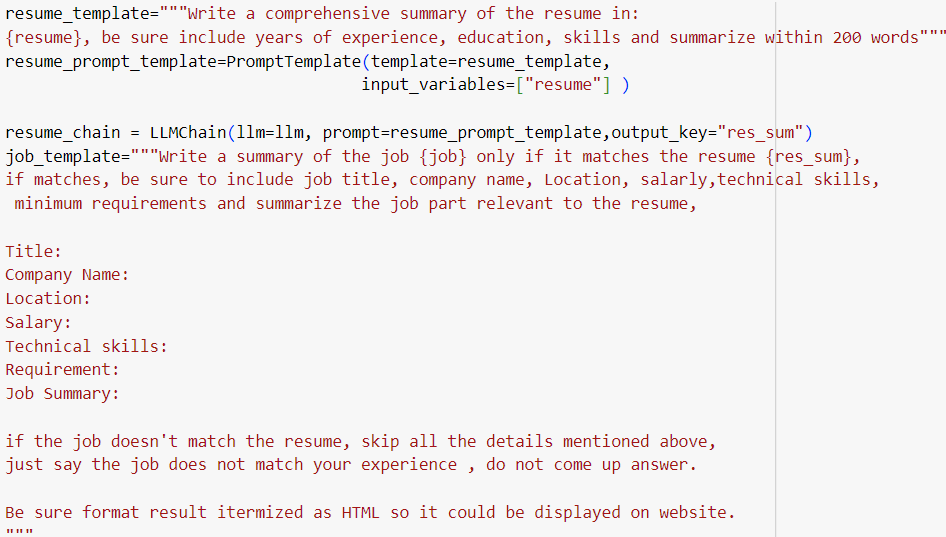
## **Company Summarization**

The objective of this task is to generate a summary from various job descriptions that match the background of the talent as outlined in their resume. The generated summary will encompass essential details including the job title, company name, location, salary information, required technical skills, job requirements, and a concise summary of the relevant portions of the job description. This summary will serve as a valuable resource for the talent, providing them with key insights into potential job opportunities that align with their qualifications and aspirations.

To accomplish this task, Langchain was utilized, and various prompts were experimented with to attain desirable outcomes. The challenge lay in determining the relevance of the job descriptions to the candidate's resume. To address this, a sequential chain approach was adopted, integrating the resume summary with each job description. The process began with a simple prompt such as "write a summary of the job where it is relevant to the resume," gradually incorporating additional details into the prompts. These details included the job title, company name, location, salary, required technical skills, job requirements, and a concise summary of the relevant aspects of the job description. Formatting requests were made within the prompts to ensure the results were presented in a clear and organized manner.

Both GPT-3.5-turbo and GPT-4-turbo were evaluated for this task. However, the results from GPT-3.5-turbo were deemed unreliable, as it often misidentified non-relevant jobs as relevant. In contrast, GPT-4-turbo consistently provided accurate matches and clear summaries of relevant job descriptions. This difference in performance underscores the significant advancements made with GPT-4-turbo, particularly in discerning relevance and generating precise summaries.

### Use Case

The final prompt is:

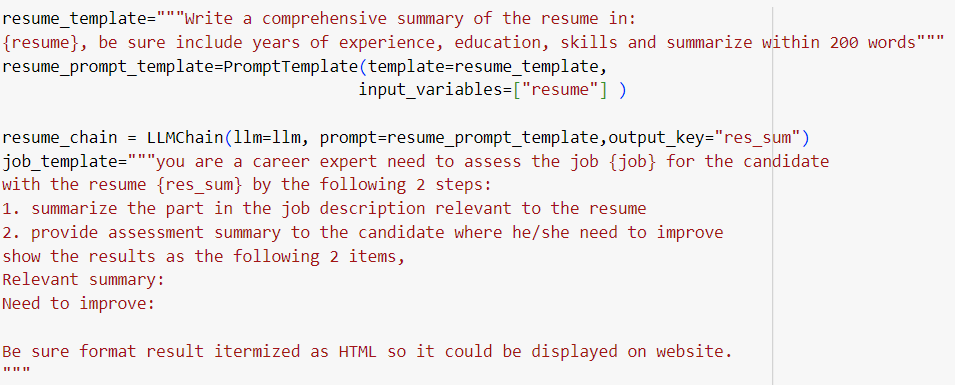
Here is part of the result from gpt-4-turbo for the 8 jobs.

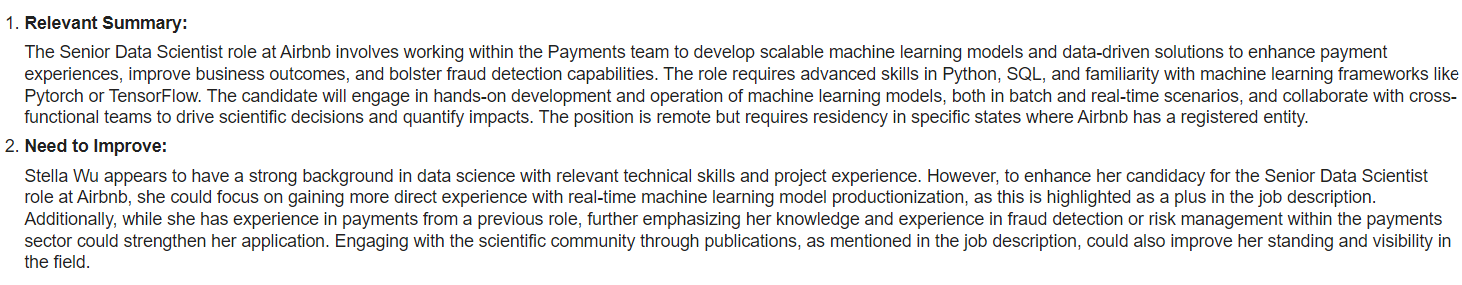
The first three jobs identified—nurse, construction engineer, and database administrator—clearly do not align with the resume, which is consistent with the results obtained from GPT-4-turbo. The matched jobs include Senior Data Scientist-Algorithms, Senior Data Analyst, and Business Analyst. Conversely, the two non-matched jobs, Science Teacher and 2024 GDIA Data Scientist/Supplemental, were correctly identified as non-matches. The latter may have been classified as such due to the limited mention of data science job content in the job description, primarily evident in the job title.

## **Job Assessment**

In assessing a specific job relative to the resume, GPT-4-turbo was employed, necessitating adjustments to the prompts to extract both the relevant job summary and areas for candidate improvement. Two crucial steps were incorporated into the prompts:

1. Summarize the portion of the job description relevant to the resume.
2. Provide a summary assessment to the candidate, outlining areas for improvement.

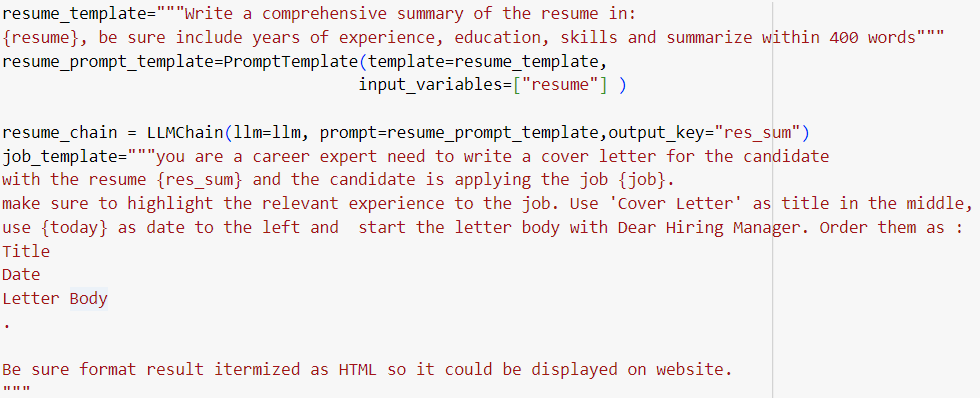
These steps were found essential for obtaining consistent and accurate results from the model. Additionally, formatting requests were made to enhance the clarity of the output.

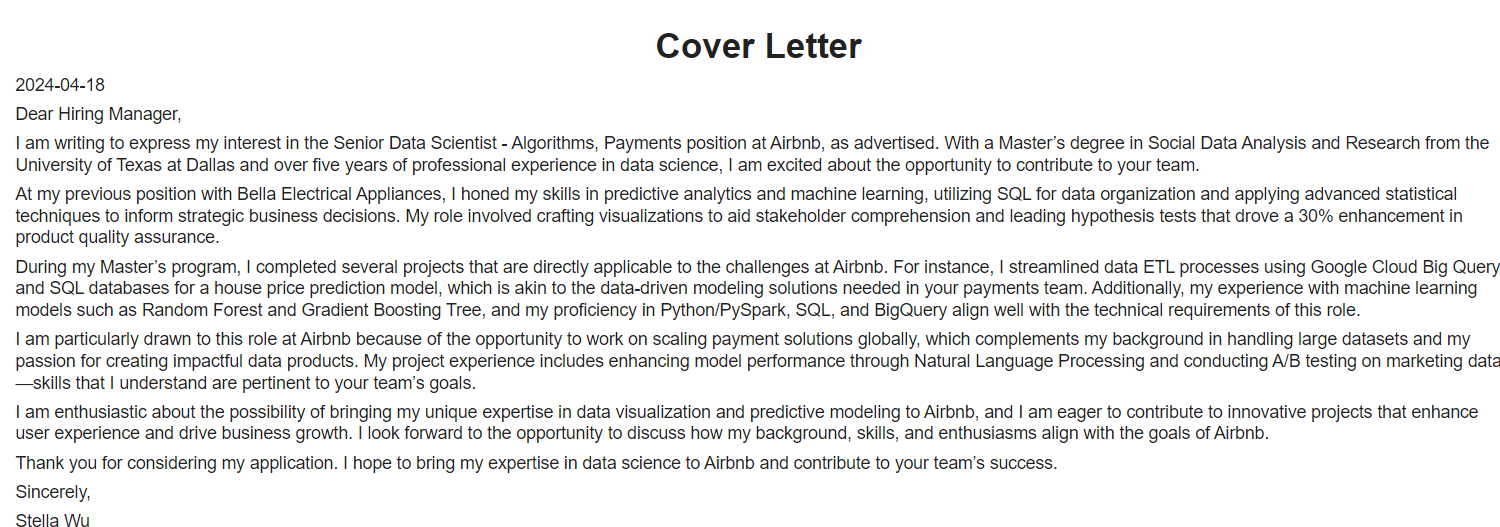
To showcase the effectiveness of this method, we employed the job listing for "Senior Data Scientist – Algorithms" to assess the resume, using the above adjusted prompts and the results shown below,

As assured, the final prompts yielded reasonable responses, as demonstrated above. The first section summarizes pertinent information aligned with the job requirements. The second part offers constructive feedback to the candidate, highlighting areas for improvement such as real-time machine learning model productionization, among others

## **Cover Letter Generation**

Text generation is a powerful tool for crafting personalized cover letters for candidates. Precisely tuning the prompt is essential to ensure the letter's specificity. One challenge lies in ensuring that the generated letter reflects the current date, rather than a date before the GPT-4-turbo training data cut-off in March 2023. To address this inconsistency, we need to incorporate a current date variable into the prompt.

Additionally, establishing a clear format and order for the cover letter can streamline the generation process. By specifying the desired format, including the layout of the header, date and body paragraphs, we can ensure consistency and professionalism in the generated letters

The generated cover letter for the job Senior Data Scientist – Algorithms at Airbnb shown as following,

## **Future Work**

Here are some future improvements,

1. Further tune the prompts to boost the performance.
2. Use RAG and LLM based retrieval to rank the relevant jobs for the candidates.
3. To generate detailed study plan for the candidates.