# PROBLEM STATEMENT

You are required to match resumes to a job posting and identify the most relevant resumes using Gen-AI tools in Python. This assignment aims to simulate a real-world scenario where recruiters need to filter through multiple resumes to find the best candidates for a given job.

# OBJECTIVES:

## JOB POSTING ANALYSIS

I chose Job Posting 1 as my job scenario which labels the need of a Senior Software Engineer.

The LLM utilized was “Gemini-Pro”.

The prompt provided was:

Read the following job description and provide a summary. The summary must contain information about the required degree, the required amount of experience. Mention the Required Qualifications as mandatory and Preferred Qualifications. Please provide answers in this format with proper line breaks and formatting. Do not mention any details twice. Keep the summary as short as possible

This prompt ensures that we receive a summary which includes all the requirements such as Degree, Experience, Necessary skills and Bonus skills. The output provided by this prompt was following:-

***Job Title:*** *Senior Software Engineer*

### Summary:

Tech Innovators Inc. seeks a Senior Software Engineer with a Bachelor's degree in Computer Science or a related field and 5+ years of experience in software development.

### Degree:

* *Bachelor's degree in Computer Science, Engineering, or a related field*

### Experience Needed:

* *5+ years of experience in software development*

### Required Qualifications:

* *Proficiency in programming languages such as Python, Java, or C++*
* *Experience with web frameworks (e.g., Django, Flask, Spring)*
* *Strong understanding of database technologies (e.g., SQL, NoSQL)*
* *Familiarity with version control systems (e.g., Git)*
* *Excellent problem-solving skills and attention to detail*

### Preferred Qualifications:

* *Master's degree in Computer Science or a related field*
* *Experience with cloud platforms (e.g., AWS, Azure, GCP)*
* *Knowledge of front-end technologies (e.g., HTML, CSS, JavaScript)*
* *Understanding of Agile development methodologies*
* *Strong communication and collaboration skills*

## RESUME

The resumes were loaded using the PyPDF2 library.

The resumes were converted into text format and splitted for better readability and then stored in a dictionary for further evaluation by our LLM model

## MATCHING PROCESS

The following prompt was provided to the LLM to provide each resume with a score ranging from 1 to 500 based on the matching skills and other criterias:

You are an experienced Application Tracking System. Evaluate the score of resumes from 1 to 500 based on the Job description summary provided.

Also mention the reason for the evaluated scores and highlight any matching skills or qualifications. Rank these resumes based on their scores.

Resumes which have a higher number of matching Required Qualification should receive a higher score. Years of experience and matching degrees have a high priority.

If the resume does not have a matching Bachelor's degree give them a very low score and avoid mentioning most of their skills

Resumes with a low amount of matching Required Qualifications should also receive a low score even if they they have matching degrees or the required amount of years of experience

Resumes with a high number of matching Preferred Qualifications but a low number of matching Required Qualification should receive a low score Matching degree and years of experience is a must

Mention that the resume with the highest score has been selected at the end of the answer

Also rank the resumes based on their scores

Select the highest ranked resume and label them as selected

If two resumes have a matching score, the resume with the highest matching Required Qualifications gets selected and if they have the same amount of Required Qualifications then the person with the highest Preferred Qualifications gets selected

In case the matching number of both skills is same, give the resume with the higher years of experience and a Masters Degree higher priority

Breakdown of the Prompt:

Resumes without matching degrees or required skills are automatically given a lower score and no other factors like years of experience and Preferred Skills are taken into consideration

Resumes with a larger amount of matching Required Skills are given more priority than the resumes which have a larger amount of matching Preferred Skills Bachelor’s Degree and Required Skills have the highest priority with years of experience coming at third

The prompt gives a list of reading for their achieved scores, does a thorough analysis of the resume and lists all the matching skills. It also excels in mentioning the number of matching skills and also includes other criterias like degrees and Experience

The results provided with this prompt were the most consistent. The output provided by the prompt was following:

### Resume 1: Amie Claude

* ***Score:*** *350*
* ***Reasoning:*** *Amie has a Bachelor's degree in Computer Science and 3 years of experience as a Software Engineer Intern. She has proficiency in Python, Java, Unix, and Git, and is familiar with JavaScript, SQL, and HTML/CSS. She has experience with web frameworks (Django) and version control systems (Git). She has strong problem-solving skills and attention to detail.*

### Matching Skills:

* + *Proficiency in Python, Java, and C++*
  + *Experience with web frameworks (Django, Flask, Spring)*
  + *Strong understanding of database technologies (SQL, NoSQL)*
  + *Familiarity with version control systems (Git)*
  + *Excellent problem-solving skills and attention to detail*

### Resume 2: Social Media

* ***Score:*** *200*
* ***Reasoning:*** *This resume does not have a matching Bachelor's degree. The candidate has experience in social media marketing and management, which is not relevant to the job description.*

### Matching Skills: None

***Resume 3: Digital Marketing***

* ***Score:*** *250*
* ***Reasoning:*** *This resume does not have a matching Bachelor's degree. The candidate has experience in digital marketing, which is not relevant to the job description.*

### Matching Skills: None

***Resume 4: Software Engineer***

* ***Score:*** *450*
* ***Reasoning:*** *Charles has a Bachelor's degree in Computer Science and 10 years of experience as a Software Engineer. He has proficiency in Python (Django), JavaScript (NodeJS, ReactJS, jQuery), SQL (MySQL, PostgreSQL, NoSQL), HTML5/CSS, AWS, and Unix/Git. He has experience with web frameworks (Django, Flask, Spring), database technologies (SQL, NoSQL), version control systems (Git), and cloud platforms (AWS). He has strong problem-solving skills and attention to detail.*

### Matching Skills:

* + *Proficiency in Python, Java, and C++*
  + *Experience with web frameworks (Django, Flask, Spring)*
  + *Strong understanding of database technologies (SQL, NoSQL)*
  + *Familiarity with version control systems (Git)*
  + *Excellent problem-solving skills and attention to detail*

### Resume 5: Tina Cook

* ***Score:*** *300*
* ***Reasoning:*** *Tina has a Bachelor's degree in Computer Science and 1 year of experience as a Software Engineer Intern. She has proficiency in C++, Mercurial, Eclipse, MongoDB, and OpenShift. She has experience with version control systems (Mercurial, Git) and database technologies (MongoDB). She has strong problem-solving skills and attention to detail.*

### Matching Skills:

* + *Proficiency in Python, Java, and C++*
  + *Experience with web frameworks (Django, Flask, Spring)*
  + *Strong understanding of database technologies (SQL, NoSQL)*
  + *Familiarity with version control systems (Git)*
  + *Excellent problem-solving skills and attention to detail*

## RANKING AND SELECTION

The resumes are ranked based on their scores and the highest scored resume is selected for the interview process.

Here we might encounter a situation where none of the resumes receive a good score so we set a threshold score for selection. This can also help in cases where companies are looking to hire multiple candidates who are above the threshold score.

Another problem is if the company only wants one candidate but there are two matching scores. To resolve the problem, the prompt was altered to cater to this particular problem. Now whenever the LLM finds matching scores it will look for the candidates with higher amount of Required Qualifications and then the Preferred Qualifications. In a case where, both the candidates have the same skillset, the one with the higher experience and a masters degree will be given higher priority.

The output for the Ranking and Selection was following:

### Ranking:

1. *Software Engineer (450)*
2. *Amie Claude (350)*
3. *Tina Cook (300)*
4. *Digital Marketing (250)*
5. *Social Media (200)*

***Selected:*** *Resume Software Engineer*

## OTHER APPROACHES

Another way to tackle this problem without heavily relying on prompts could have been using the “cosine\_similarity” function.

But in order to use this function the Job Description and Resumes have to be converted to vector embeddings and stored in the vector database.

Problem Faced while using this approach:

Gemini doesn’t offer any functionality where can i can store vector embeddings of different pdfs in different indexes inside the database since it lack the “embed\_text” attribute

Google does offer this functionality using google.cloud which is a paid feature. OpenAI is in a similar boat since it has a subscription based API

Transformers could have been used for vector conversion but the dependencies required for Transformers clashed with the dependencies of the already in-use GenAI tools and the Visual Studio Code Redistributables