



ISY5001 Intelligent Reasoning Systems

Master of Technology in Intelligent Systems

CareerButterfly

The One Stop Shop for your Tech Career

Project Report

Group Members (IS06):

- Soo Kuan Yong (A0291188U)
- Zachary Chua Feng Kwan (A0180556B)
- Ng Min Teck (A0291213N)
- Ang Han Chuan Wilson (A0291657R)

Table of Contents

1. Executive Summary	5
2. Market Research.....	6
3. Project Description	7
3.1 Project Background	7
3.2 Project Objectives	8
3.3 Report Organisation	8
4. Project Solution.....	9
4.1 Job Recommender	10
4.1.1 Data Collection	10
4.1.1.1 Job Posts Data Collection.....	11
4.1.1.2 User Review of a Company	12
4.1.2 Data Pre-processing	12
4.1.2.1 Resume Parser.....	13
4.1.2.2 User Preferences.....	14
4.1.2.3 Feature Engineering (Postal Code to Location)	14
4.1.3 Cosine Similarity Score	16
4.1.4 Company Reviews Sentiment Analysis.....	16
4.1.4.1 User Review Sentiment and User's Company Rating Score	18
4.1.4.2 User Preference Score	19
4.1.5 Weighted Average Score to rank and recommend job	20
4.2 Resume Generator.....	21
4.2.1 Prompt Engineering Design.....	22
4.3 Cover Letter Generator	22
4.3.1 Prompt Engineering Design.....	22
4.4 Behavioural Interview Chatbot.....	23
4.4.1 Prompt Engineering Design.....	23
4.5 Learning Resource Recommender.....	25
4.5.1 Knowledge-based Reasoning Learning Resource Recommender	25

4.5.1.1 Knowledge Acquisition	25
4.5.1.2 Reasoning System.....	26
4.5.2 Recommended Technical Interview Questions based on company.....	27
4.5.2.1 Data collection	27
4.5.2.2 Recommended Technical Interview Questions.....	28
5. Project Implementation	29
5.1 System Design	29
5.2 Deployment Architecture	29
5.3 Streamlit Application.....	31
5.3.1 Home Page	32
5.3.2 User Preference Page	32
5.3.3 Cover Letter Generator, Resume Generator, Learning Resource Recommender Page.....	34
5.3.4 Behavioural Interview Chatbot Page.....	34
6. Project Performance & Validation.....	36
6.1 Home Page	37
6.2 User Preference Page.....	38
6.3 Cover Letter Generator, Resume Generator, Learning Resource Recommender Page	40
6.4 Behavioural Interview Chatbot Page	44
7. Project Conclusion	46
7.1 Future Implementation	46
Appendices	47
Appendix A: Project Proposal	47
Appendix B: Mapped System Functionalities against knowledge, techniques and skills of modular courses: MR, RS, CGS.....	51
Appendix C: Installation and User Guide.....	52
Appendix D: Individual Reports	61
Appendix E: List of Abbreviations	69
Appendix F: List of Figures.....	70

Appendix G: List of Table	72
Appendix H: References	73

1. Executive Summary

Layoffs are becoming a disturbing norm. Despite frequent layoffs, the median gross salary income for full-time employees in the Information and Communications sector is still on the rise. With the increasing demand for jobs and the increasing gross salary in the technology sector, getting a job in the technology sector becomes increasingly more competitive. While there are many individual tools to help a job seeker in their path to search for a job that fits them, there is not a single platform that combines all the tools together and streamlines the processes to help ease the job-hunting process. This project aims to develop a single platform to streamline this process.

This project develops a single platform application built on Streamlit and uses microservices to recommend jobs fitted to user preferences, generate cover letter, resume and learning resources for our users, and provides an interactive behavioural interview chatbot for mock interview sessions.

This project makes use of recommendation system, knowledge-based reasoning techniques and uses Large Language Model (LLM) and prompt engineering to develop a chatbot and generate documents detailed in the following pointers:

- Job recommender is a recommendation system that uses a weighted average score based on cosine similarity between user resume and job description, user review sentiment and user's company rating score from company reviews and ratings, and user preference score from company reviews.
- Cover letter generator and resume generator both use LLM and well-crafted effective prompts that acts as instruction to guide the LLM to achieve the desired output customised for their application.
- Learning resources recommender uses knowledge-based reasoning techniques to recommend learning resources and generate learning resources document for the user to bridge the knowledge gap.
- Behavioural question interview chatbot is developed with the help of LLM and prompt engineering that identifies possible behavioural questions for user to practice in mock interview sessions with the chatbot and provides timely feedback to user on how to improve their answers in the future.

This project combines multiple job seeking tools together in a single platform and helps to streamline the process of finding a job suitable for our users in a timely manner.

2. Market Research

In the current market, there are many tools that can help job seekers in their job searching process. Some examples of tools such as resume, and cover letter generators shown on Table 1 and Table 2 can provide the candidates with complete documents but often have a paywall behind them.

Table 1. Examples of resume generators in the market and their limitations

Resume Generators	Limitations
Novorésumé ^[1]	Only the first page of the resume is free
Resume.io ^[2]	Must pay to download resume
Zety ^[3]	Must pay to download resume

Table 2. Examples of cover letter generators in the market and their limitations

Cover Letter Generators	Limitations
MyPerfectResume ^[4]	No personalisation and is not free
Grammarly ^[5]	No personalisation

Currently in the market, there is no tool that can identify hard skill gaps and recommend learning resources to supplement the gaps between job description and resume. Another key step in the job seeking process is the interview process. Interview preparation usually comes from experience, talking to career coach and searching for past interview questions from user reviews of a company in websites such as Glassdoor and Nodeflair. This can be very time consuming and might not return great results. This process mostly only helps to uncover behavioural questions.

While there are individual tools out there that can help to ease the process of finding a job but there is not a single platform that combines the abovementioned tools together to provide a seamless process for the job seekers. Hence this project aims to provide all the tools in a single platform for job seekers.

3. Project Description

3.1 Project Background

Although layoffs have constantly been happening over the years as shown in Figure 1, according to the data on recent trends of layoffs in Singapore provided by the Ministry of Manpower, there is a steady increase in retrenchment numbers starting from the second quarter of 2022. The increasing trend will likely continue as companies go through reorganisation and restructuring.



Figure 1. Recent layoff trends ^[6]

With the increasing number of layoffs, more job seekers will be hunting for jobs^[7]. Despite all the layoffs happening, the median gross salary income for full-time employees in the Information and Communications sector is still on the rise. As the demand of job increases and base salary of the jobs increase, getting a job in the technology sector becomes increasingly more competitive.

While there are many job seeking tools such as job recommender, resume and cover letter generators, and the use of LLM applications such as ChatGPT can be helpful in preparing the documents and interview process, they are still individual tools specialised for a single task. However, finding jobs that fit individual preferences, customizing resume, cover letter and preparing for the interviews tuned for each job takes time and effort. There is not a single platform that combines all the tools together and streamlines the processes to help ease the

job-hunting process. In the current market landscape, having a competitive advantage is crucial to land a job quickly.

3.2 Project Objectives

The project aims to develop a single platform that recommend relevant jobs fitted to individual preferences, generate customized resume and cover letter, analyse, and identify technical skill gaps and recommend learning resources including a table detailing the type of technical questions to expect based on company and job title, and provide a behavioural interview chatbot for mock interviews. The project is divided into multiple systems, namely job recommender, resume generator, cover letter generator, learning resource recommender, and behavioural interview chatbot.

The project must achieve the following:

1. **Decision Automation:** Knowledge based reasoning techniques
 - Automatically identify pre-requisite technical skills that are required but not listed in the job description using reasoning techniques.
 - Automatically identify technical skill gaps and recommend learning resources for user to bridge the knowledge gap.
2. **Knowledge discovery & (big) data mining techniques:** recommendation systems
 - Recommend jobs that are most relevant to users using a weighted average score based on cosine similarity between user resume and job description, user review sentiment and user's company rating score from company reviews and ratings, and user preference score from company reviews.
3. **System designed with cognitive techniques or tools:** chatbot
 - Make use of LLM and prompt engineering to develop a chatbot that identifies possible behavioural interview questions for user to practice in mock interviews sessions with the chatbot.
 - The chatbot should also be able to provide feedback to the users on how to improve their interview answers in the future.

The following chapter will detail the how each subsystem works.

3.3 Report Organisation

This report is organised as follows: Chapter 4 covers the project solution. Chapter 5 describes the project implementation. Chapter 6 discusses about the project performance and validation. Chapter 7 covers the conclusion and future implementation.

4. Project Solution

The solution consists of 5 main subsystems that are integrated together into a single web application, namely:

1. Job recommender
2. Resume generator
3. Cover letter generator
4. Learning resource recommender
5. Behavioural interview chatbot

The project takes in user input such as past job experiences, education history, personal information from their resume and information on the job they are looking for to recommend 5 most relevant jobs to the profile. User will have to choose one of these jobs and will generate cover letter, resume, and learning resources. With the cover letter and resume, user will be able to submit their job application and can refer to the learning resources to supplement their skill gaps and practice on relevant technical interview questions. Once user is ready to move on to the last step, user can proceed to practice behavioural interview question with our chatbot. The following sections will cover on how each subsystem works.

4.1 Job Recommender

Job recommender is a recommendation system that takes in user's resume and user's preferences to check against the job posts and user review data that we collected and recommend 5 most relevant jobs to the user profile. For easy reference, we refer to user's resume and user's preferences collectively as candidate information in the following chapters. Figure 2 shows the flow diagram of job recommender system.

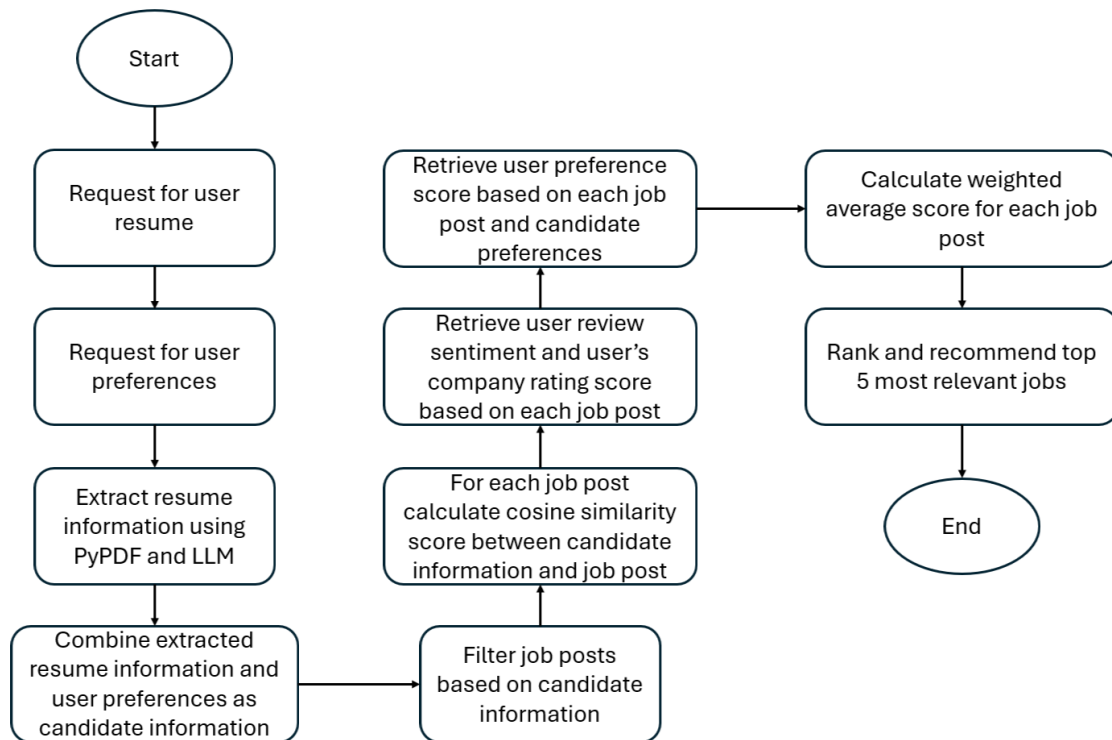


Figure 2. Flow diagram of job recommender system

The following subsections will detail the techniques used to find out 5 most relevant jobs to recommend.

4.1.1 Data Collection

The job recommender uses 2 data sources to make recommendations, namely job posts and user review of the company. The following subsections will describe the data collection process in more detail.

4.1.1.1 Job Posts Data Collection

Before we can make job recommendations, we must first find out what are the jobs available for each role by searching for job posts on job portals. The job posts in this project are scraped from myCareersFuture using a Python library called Selenium. Figure 3 shows the flow diagram of the data collection process from myCareersFuture using Selenium.

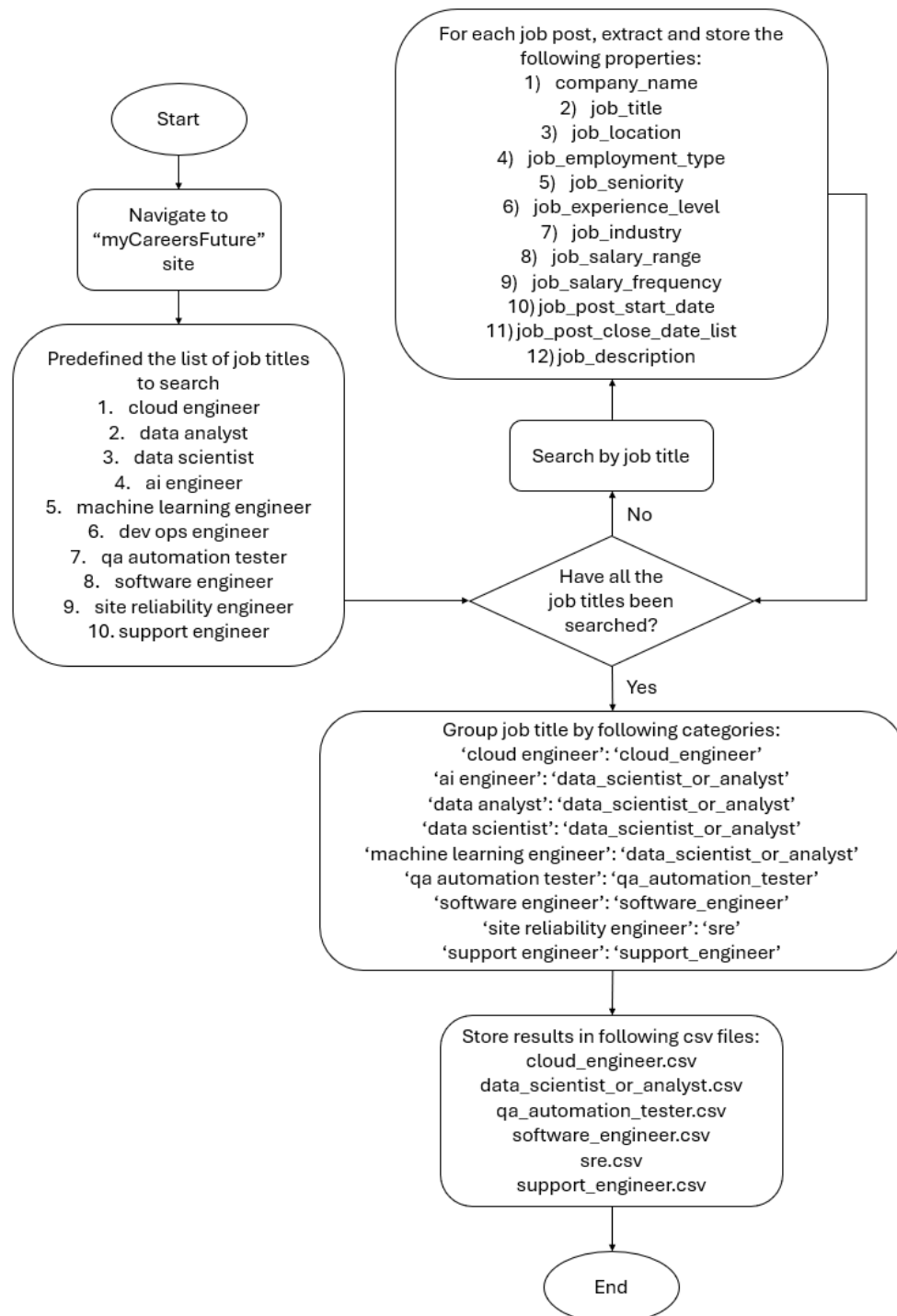


Figure 3. Flow diagram of the data collection process from myCareersFuture using Selenium

The job posts data are stored in multiple csv files such as cloud_engineer.csv, data_scientist_or_analyst.csv, qa_automation_tester.csv, software_engineer.csv, sre.csv, and support_engineer.csv.

4.1.1.2 User Review of a Company

User reviews of a company indicates an employee's perception of their company across various aspects such as work-life balance, compensation, and career growth opportunities. Analysing the review highlights the area of strength and concern. Job seekers can grasp a better understanding if their values and career goals are aligned with that of the company based on the reviews from current and past employees.

Figure 4 shows the activity diagram that outlines the data collection process from Glassdoor using Selenium.

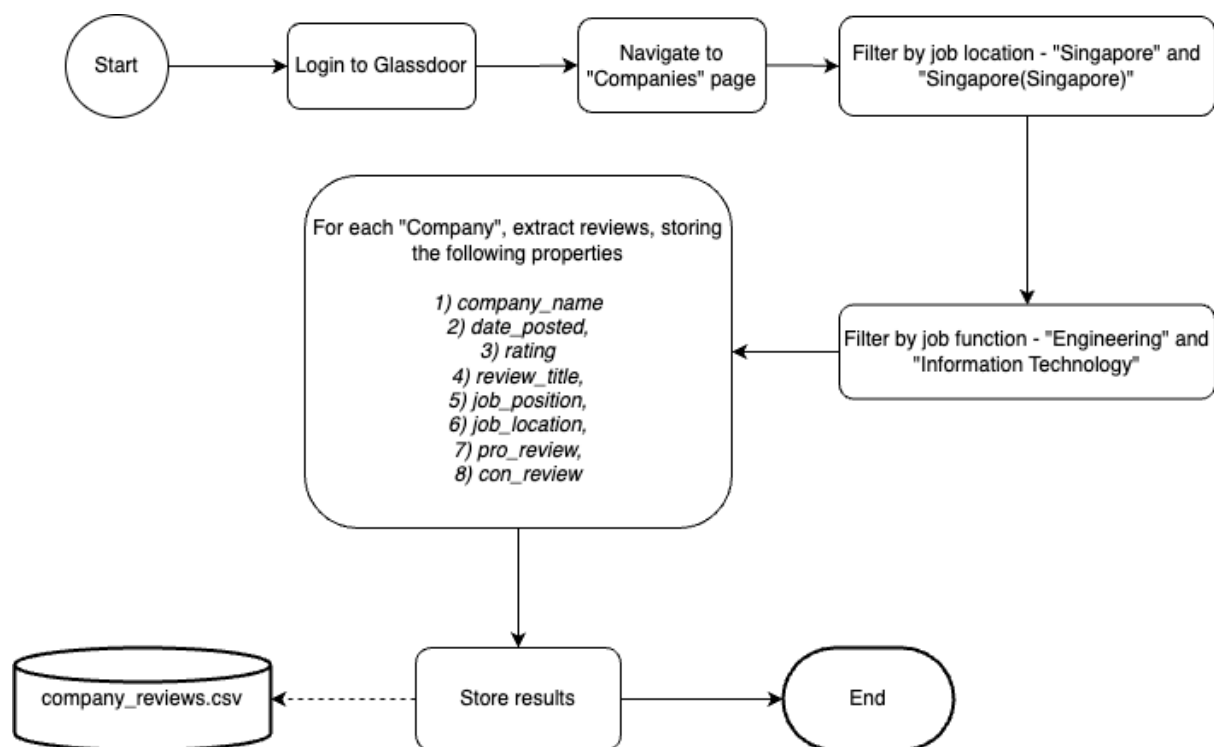


Figure 4. Activity diagram that outlines the data collection process from Glassdoor using Selenium

The user reviews of different companies will be saved in company_reviews.csv.

Once the data is collected, the next step is to clean and prepare the data to score and recommend the jobs.

4.1.2 Data Pre-processing

The following subsections will describe the data pre-processing done in preparation for job recommendation.

4.1.2.1 Resume Parser

Other than job posts and user reviews described above, job recommender also requires user to upload their resume in Portable Document Format (PDF) format. From the resume, we must extract the resume information to recommend jobs for the user. A Python library called PyPDF2 is used to extract the text from in resume. As resumes have different format, to ensure the same key information is extracted, we made use of LLM and prompt engineering to extract user's personal information, past work experiences, education history, and technical skills they have. Figure 5 shows the prompt used to extract key candidate information for the job recommender.

```
input_prompt='''
Act as a skilled Recruiter or very experience ATS(Application Tracking System)
with a deep understanding of tech field, software engineering, data science, data analyst
and big data engineer. Your task is to evaluate the resume and return the following information in this format.
For the skills, take note to add in technical skills gathered throughout the resume.
For Job no., provide it in a numerical form.
For Degree no., provide it in a numerical form.
if the answer is not in provided resume just say, "answer is not available in the context", don't provide the
wrong answer. Scan through the resume multiple times to ensure that no information was missed out:

- Name:
- Work Experiences:
-----
\t- Job no.:
\t- Company Name:
\t- Job title:
\t- Job tenure period:
\t- Job description:
-----
- Education information:
-----
\t- Degree no.:
\t- Degree title:
\t- Degree years:
\t- University:
-----
- Programming Skills:
- Data Analysis Skills:
- Data Science Skills:
- Data Visualization Skills:
- Front End Skills:
- Back End Skills:
- Cloud technology Skills:
resume:{0}'''.format(text)
```

Figure 5. Prompt to extract user's personal information, work experiences, education history, and technical skills

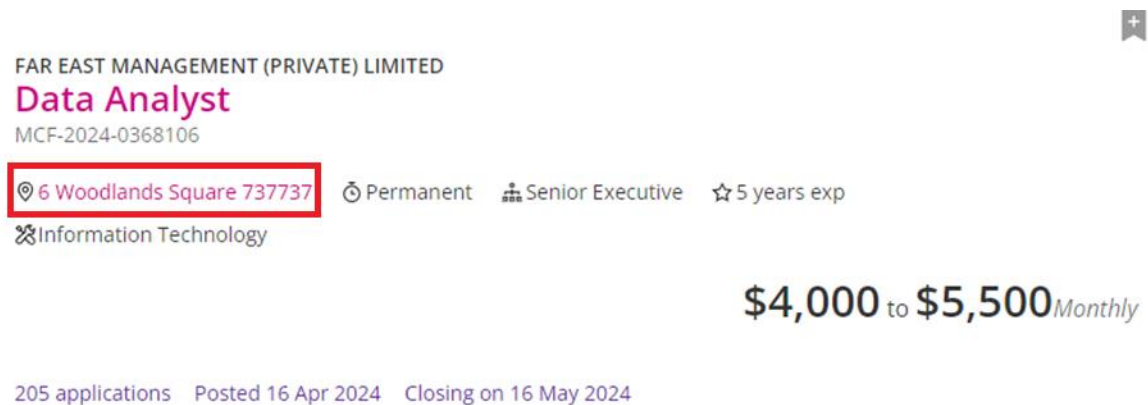
4.1.2.2 User Preferences

In order to have a better understanding of the job that user is looking for, the job recommender also requires user to key in their user preferences. The user preferences that include the job position they are looking for, desired location if any (e.g. Central, North, South), desired employment type (e.g. contract or permanent), desired seniority level (e.g. fresh/entry level, executive, senior executive), required number of years of job experience, desired job aspects (e.g. work life and flexibility or career development and learning), and minimum salary.

The only required fields for user preferences are the job position and the desired job aspects they look for. The rest of the fields could be left blank if the user wishes to.

4.1.2.3 Feature Engineering (Postal Code to Location)

Figure 6 shows an example of a job post that we scraped. As the location provided is given as an address and we will not immediately be able to tell which region of Singapore it belongs to.



FAR EAST MANAGEMENT (PRIVATE) LIMITED
Data Analyst
MCF-2024-0368106

📍 6 Woodlands Square 737737 🕒 Permanent 👤 Senior Executive ⭐ 5 years exp

🔧 Information Technology

\$4,000 to \$5,500 Monthly

205 applications Posted 16 Apr 2024 Closing on 16 May 2024

Figure 6. Location of a job post is given as an address

In order to convert the address into region, we converted the postal code of the address from each job post into postal district by referring to Figure 7 below from Urban Redevelopment Authority (URA)'s webpage ^[8].

List of Postal Districts

Postal District	Postal Sector (1st 2 digits of 6-digit postal codes)	General Location
01	01, 02, 03, 04, 05, 06	Raffles Place, Cecil, Marina, People's Park
02	07, 08	Anson, Tanjong Pagar
03	14, 15, 16	Queenstown, Tiong Bahru
04	09, 10	Telok Blangah, Harbourfront
05	11, 12, 13	Pasir Panjang, Hong Leong Garden, Clementi New Town
06	17	High Street, Beach Road (part)
07	18, 19	Middle Road, Golden Mile
08	20, 21	Little India
09	22, 23	Orchard, Cairnhill, River Valley
10	24, 25, 26, 27	Ardmore, Bukit Timah, Holland Road, Tanglin
11	28, 29, 30	Watten Estate, Novena, Thomson
12	31, 32, 33	Balestier, Toa Payoh, Serangoon
13	34, 35, 36, 37	Macpherson, Braddell
14	38, 39, 40, 41	Geylang, Eunos
15	42, 43, 44, 45	Katong, Joo Chiat, Amber Road
16	46, 47, 48	Bedok, Upper East Coast, Eastwood, Kew Drive
17	49, 50, 81	Loyang, Changi
18	51, 52	Tampines, Pasir Ris
19	53, 54, 55, 82	Serangoon Garden, Hougang, Punggol
20	56, 57	Bishan, Ang Mo Kio
21	58, 59	Upper Bukit Timah, Clementi Park, Ulu Pandan
22	60, 61, 62, 63, 64	Jurong
23	65, 66, 67, 68	Hillview, Dairy Farm, Bukit Panjang, Choa Chu Kang
24	69, 70, 71	Lim Chu Kang, Tengah
25	72, 73	Kranji, Woodgrove
26	77, 78	Upper Thomson, Springleaf
27	75, 76	Yishun, Sembawang
28	79, 80	Seletar

Source: SingPost

Figure 7. List of Postal District and corresponding Postal Sector

Using the postal district, we convert the postal district into individual regions of Singapore (i.e. Central, North, South, East, West) by referring to Figure 8 below from URA source.

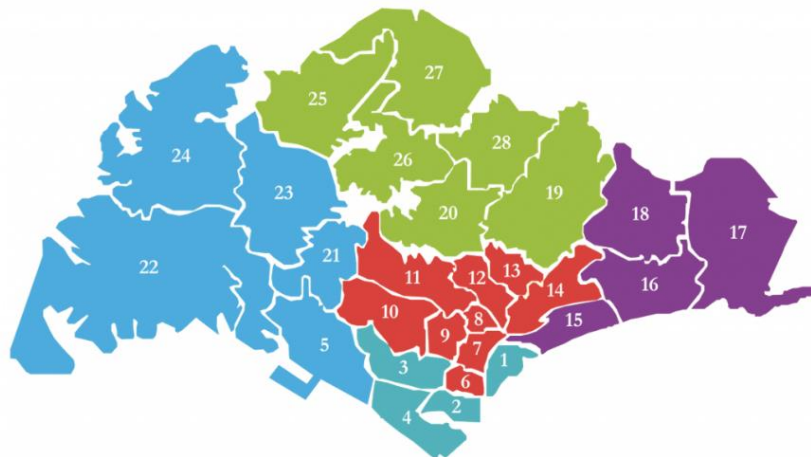


Figure 8. Regions of Singapore based on Postal District

With the region, we can then filter the jobs based on user preference.

4.1.3 Cosine Similarity Score

Once the candidate information and job posts have been pre-processed, we will treat each candidate information and job description as a pair of documents and apply bag of words technique on them. With the count matrix output from bag of words, we will apply cosine similarity to find out how similar the candidate information and job description are. The cosine similarity formula is as follows:

$$\text{cosine similarity}(A, B) = \frac{A \cdot B}{\|A\| * \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2 \sum_{i=1}^n B_i^2}}$$

Cosine similarity score will be a number between 0 to 1. If the value is closer to 0, this means that both candidate information and job description are not similar at all, and if value is closer to 1, this means that both candidate information and job description are very similar.

This score will then be used as part of the weighted average score to rank and recommend jobs that are similar to the user profile.

4.1.4 Company Reviews Sentiment Analysis

Aside from job description, we also look at company reviews and ratings to have a holistic view when determining whether a job position is great to recommend. Therefore, in this section, we will discuss the approach used.

Figure 9 shows the activity diagram that outlines the process of scoring companies based on the implicit score and explicit score. Implicit scores are derived through sentiment analysis of the pro reviews and con reviews, while explicit scores are derived directly from the user's rating.

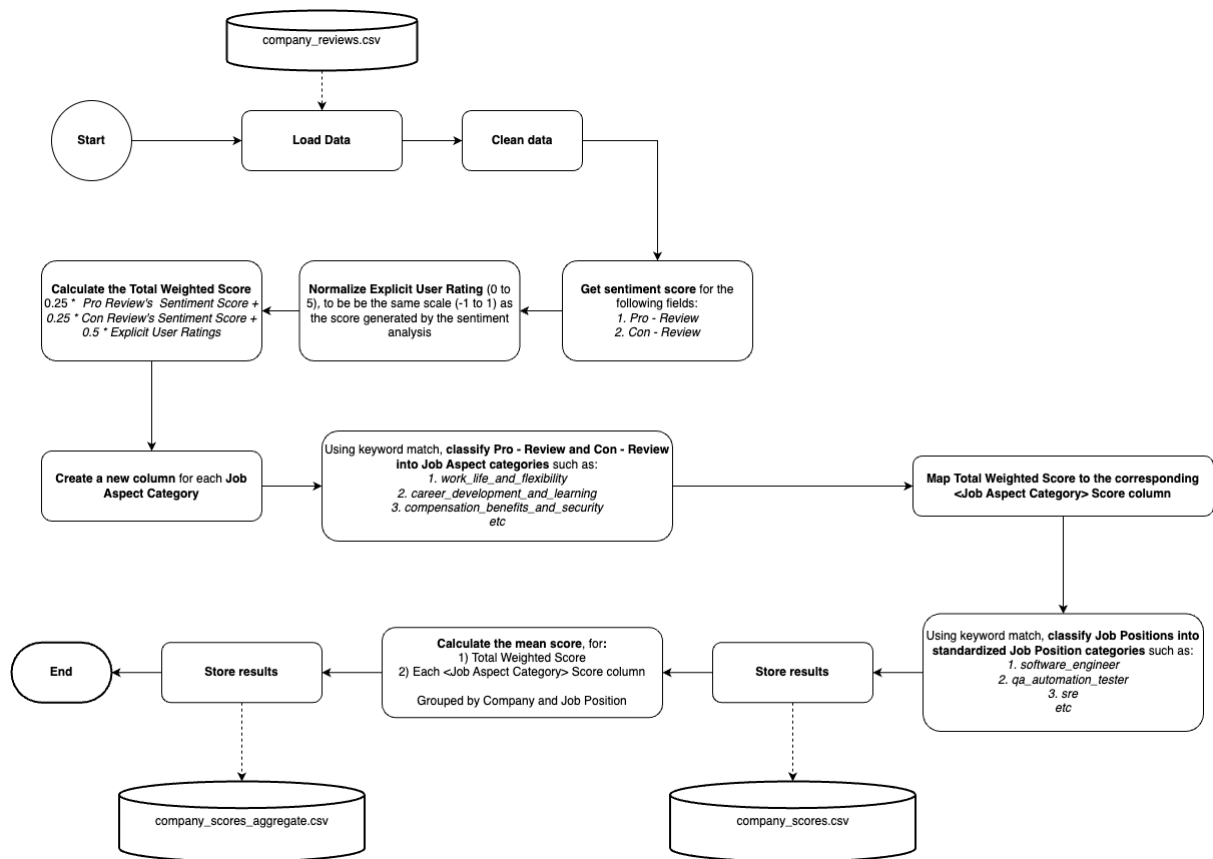


Figure 9. Activity diagram that outlines this

4.1.4.1 User Review Sentiment and User's Company Rating Score

Figure 10 shows an example of a user review. From the user review, we will collect the job title, company name, user's company rating, company location, pros and cons reviews. Sentiment analysis is applied on each of the pros and cons reviews.

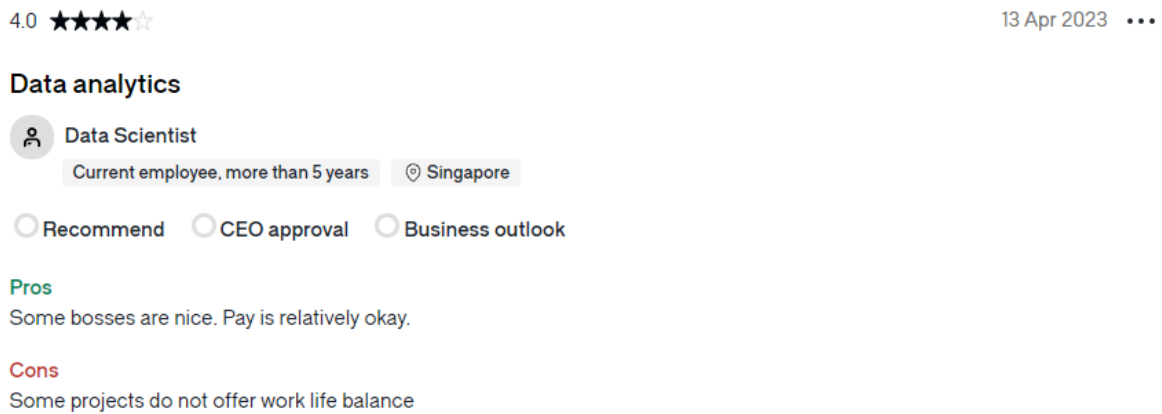


Figure 10. An example of user review

Each of the pros and cons review will receive a score between -1 to 1. If the score is between -1 to -0.05, it will be negative sentiment. If the score is between -0.05 to 0.05, it will be a neutral sentiment. If the score is between 0.05 to 1, it will be a positive sentiment. For easy reference, we will refer to the output as “pro score” and “con score”.

User's company rating score is originally between 1 to 5. We normalised this number to a value between -1 to 1.

We combine the pro score, con score and user's company rating using the following formula:

$$\begin{aligned} & \text{user review sentiment and user's company rating score} \\ &= (0.2 \times P) + (0.25 \times C) + (0.5 \times U) \end{aligned}$$

where:

- P = pro score
- C = con score
- U = user's company rating

This score will then be used in the weighted average score to rank and recommend job.

4.1.4.2 User Preference Score

In order to recommend more relevant jobs to the user, we allow them to express their preferences in the job aspect that they are looking for. Using the pros and cons review, we group them into different job aspects by picking up keywords used. Table 3 shows the list of keywords that we pick up for each job aspect. The job aspect derived from the pro or con review will share the same sentiment score as the pro score and con score.

Table 3. List of keywords to determine each job aspect

Job aspects	Keywords
Work life balance and flexibility	work life balance, no overtime, flexible hours, work-life, work and life balance, remote work, work from home, telecommute, flexible location, work-life integration
Career development and learning	career advancement, promotion opportunities, growth opportunities, career growth, progression, training, courses, skills development, personal growth, learning opportunities, career mobility, lateral moves, role changes, internal opportunities, innovation, creativity, creative thinking, innovative ideas
Compensation, benefits and job security	salary, pay, compensation, benefits, bonus, remuneration, rewards, insurance, retirement, parental leave, job security, employment stability, company stability, layoffs, restructuring
Culture and environment	culture, team spirit, collaboration, work environment, company culture, team dynamics, teamwork, team collaboration, team atmosphere, environmental sustainability, eco-friendly, sustainability efforts, green initiatives, health, safety, well-being, mental health, workplace safety
Management and communication	leadership, management, executive, supervisor, managerial style, communication, transparency, open communication, feedback
Diversity and inclusion	diversity, inclusion, equal opportunity, inclusive environment, cultural diversity, international, multicultural
Employee engagement and satisfaction	satisfaction, engagement, morale, happiness, valued, team bonding, job satisfaction
Operational efficiency and resources	operational efficiency, resources, project management, task management, resource allocation, organizational structure
Innovation and strategic vision	company-wide innovation, strategic direction, future vision, strategic planning, long-term goals
Global impact and social responsibility	global impact, social responsibility, society contribution, positive impact

If the user chooses their preferred job aspects, it will then be used in the weighted average score to rank and recommend job.

4.1.5 Weighted Average Score to rank and recommend job

After the cosine similarity score, user review sentiment and user's company rating score, and user preference score is derived, a final weighted average score to rank and recommend the top 5 most relevant job will be calculated using the following formula:

$$\text{Weighted Average Score} = \frac{(W_1 \times S_1) + (W_2 \times S_2) + (W_3 \times S_3)}{(W_1 \times E_1) + (W_2 \times E_2) + (W_3 \times E_3)}$$

where:

- S_1 = Cosine similarity score between candidate information and job description
- S_2 = User review sentiment and user's company rating score
- S_3 = User preference score
- W_1, W_2, W_3 = Weights (all set to 1/3 to hold equal weightage)
- E_1 = When cosine similarity score exists (represents existence of S_1 , 1 if exists, else 0)
- E_2 = When user review sentiment and user's company rating score exist (represents existence of S_2 , 1 if exists, else 0)
- E_3 = When user preference score exists (represents existence of S_3 , 1 if exists, else 0)

All the jobs will be ranked based on this weighted average score and the top 5 jobs will be shown to the user where they will pick one of them to go for.

4.2 Resume Generator

3 of the subsystems make use of LLM, namely cover letter generator, resume generator and the behavioural interview chatbot. The variation of LLM that we use in this project is a Generative Pre-trained Transformer (GPT), specifically the GPT-3.5-Turbo-Instruct model. Figure 11 details the information of available GPT-3.5 Turbo models. In this section, we will be discussing about the resume generator. It leverages on GPT 3.5 model to generate a resume based on candidate information and job description.

GPT-3.5 Turbo

GPT-3.5 Turbo models can understand and generate natural language or code and have been optimized for chat using the [Chat Completions API](#) but work well for non-chat tasks as well.

MODEL	DESCRIPTION	CONTEXT WINDOW	TRAINING DATA
gpt-3.5-turbo-0125	New Updated GPT 3.5 Turbo The latest GPT-3.5 Turbo model with higher accuracy at responding in requested formats and a fix for a bug which caused a text encoding issue for non-English language function calls. Returns a maximum of 4,096 output tokens. Learn more.	16,385 tokens	Up to Sep 2021
gpt-3.5-turbo	Currently points to gpt-3.5-turbo-0125.	16,385 tokens	Up to Sep 2021
gpt-3.5-turbo-1106	GPT-3.5 Turbo model with improved instruction following, JSON mode, reproducible outputs, parallel function calling, and more. Returns a maximum of 4,096 output tokens. Learn more.	16,385 tokens	Up to Sep 2021
gpt-3.5-turbo-instruct	Similar capabilities as GPT-3 era models. Compatible with legacy Completions endpoint and not Chat Completions.	4,096 tokens	Up to Sep 2021
gpt-3.5-turbo-16k	Legacy Currently points to gpt-3.5-turbo-16k-0613.	16,385 tokens	Up to Sep 2021
gpt-3.5-turbo-0613	Legacy Snapshot of gpt-3.5-turbo from June 13th 2023. Will be deprecated on June 13, 2024.	4,096 tokens	Up to Sep 2021
gpt-3.5-turbo-16k-0613	Legacy Snapshot of gpt-3.5-16k-turbo from June 13th 2023. Will be deprecated on June 13, 2024.	16,385 tokens	Up to Sep 2021

Figure 11. Information on GPT-3.5 models

4.2.1 Prompt Engineering Design

In order to guide LLM more effectively to achieve the desired output customised for the resume generator, we have crafted effective prompts that act as instruction to guide the LLM. Figure 12 shows the prompt used to generate the resume.

```
prompt_str=("Write a structured and professional resume with the following information: " + str(candidate_information) + \
            ", extract the information from the json object provided. Do not assume any missing information and \
            instead replace them with [placeholders]. ")
```

Figure 12. Prompt used to generate resume for the user

Figure 13 shows the flow diagram of resume generator.

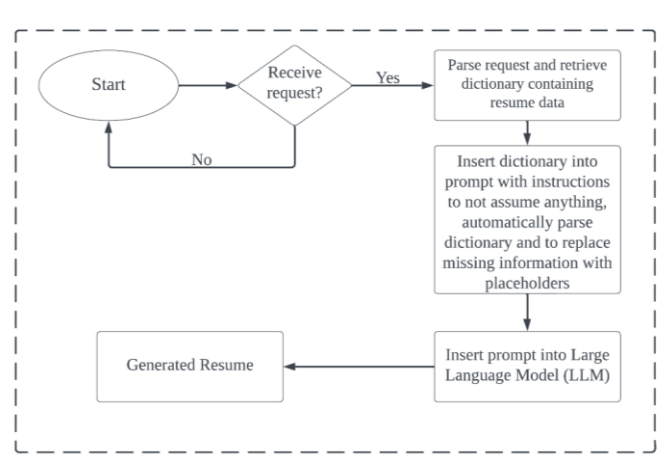


Figure 13. Flow diagram of resume generator

4.3 Cover Letter Generator

In this section, we will be discussing about the cover letter generator. It leverages on GPT 3.5 model to generate a cover letter based on candidate information and job description.

4.3.1 Prompt Engineering Design

Similarly, to guide LLM more effectively to achieve the desired output customised for the cover letter generator, we have crafted effective prompts that act as instruction to guide the LLM. Figure 14 shows the prompt used to generate the cover letter.

```
prompt_str = ("Write a cover letter to " + contact_person + " for a " + role + " job at " + company_name + "." + \
              "My name is " + name + ", my email is " + email + " . My mobile number is " + mobile_number + \
              " . My linkedIn profile is " + linkedin_profile + " . My skills are " + skills + \
              " . I am excited about the job because " + job_description + " I am passionate about "+ passion + \
              " . Do not assume anything and replace any missing information with [placeholders].")
```

Figure 14. Prompt used to generate cover letter for the user

Figure 15 shows the flow diagram of cover letter generator.

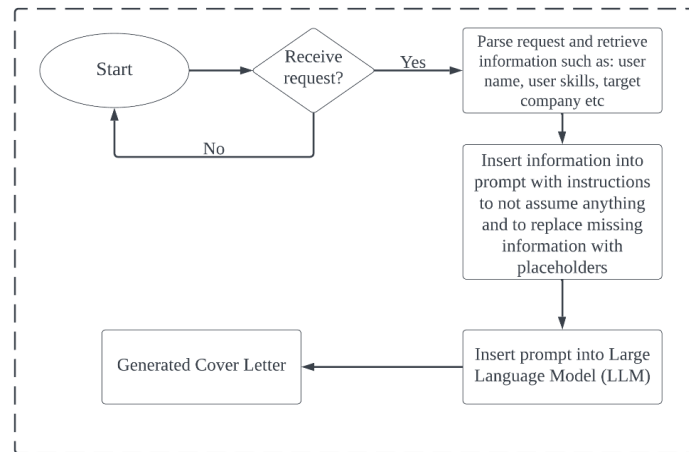


Figure 15. Flow diagram of cover letter generator

4.4 Behavioural Interview Chatbot

In this section, we will be discussing about the behavioural interview chatbot. The behavioural interview chatbot leverages on GPT 3.5 model to provide behavioural questions based on candidate information and job description. Once the user provides their responses to those questions, the chatbot will provide feedback and how the user can improve their answers.

4.4.1 Prompt Engineering Design

Likewise, to guide LLM more effectively to achieve the desired output customised for the behavioural interview chatbot, we have crafted effective prompts that act as instruction to guide the LLM. Figures 16 and 17 show the respective prompts used to generate behavioural interview questions based on candidate information and job description, and to evaluate user response and generate feedback to that response.

```

interview_questions_prompt = '''
Act as a skilled Interviewer with a deep understanding of tech field, software engineering, data science, data analyst
and big data engineer. Your task is to evaluate the resume and given job description to come up with possible interview
questions to help your client interviewee to prepare for their upcoming interview.

resume:{0}
job description:{1}

Add the following at the end of your response:
Please provide me with your response, I will evaluate and provide feedback based on your answers.
'''
.format(candidate_key_information, selected_job_description)

```

Figure 16. Prompt to generate behavioural interview questions based on candidate information and job description

```

evaluate_user_prompt = '''
Act as a skilled Interviewer with a deep understanding of tech field, software engineering, data science, data analyst
and big data engineer. Evaluate user's interview answer, provide feedback on how they can improve their answers
for their upcoming interviews.

Be kind and encouraging in your response.
Provide motivational words at the end of the response to help motivate the interviewee.
interviewee's answer: {0}
'''.format(prompt)

```

Figure 17. Prompt to get GPT model to evaluate user response and generate feedback to that response

Figure 18 shows the flow diagram of behavioural interview chatbot.

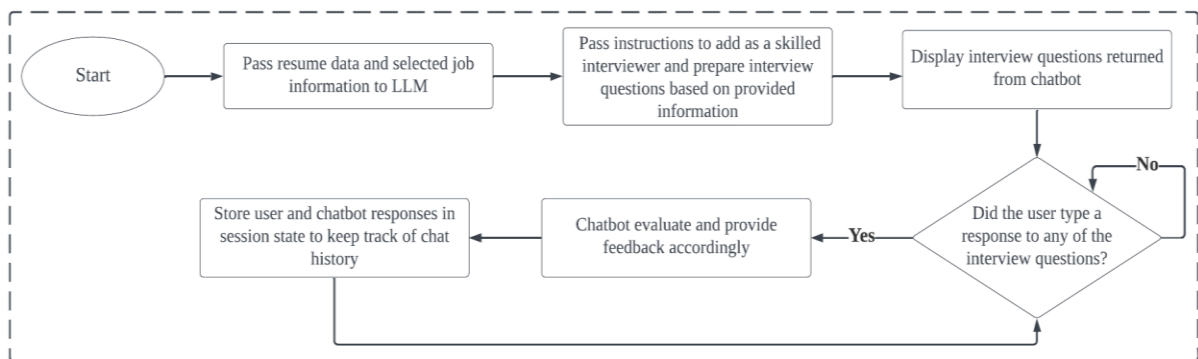


Figure 18. Flow diagram of behavioural interview chatbot

With the above prompts, the individual subsystems are able to achieve a desired output from the LLM model.

4.5 Learning Resource Recommender

Figure 19 shows the flow diagram of learning resource recommender.

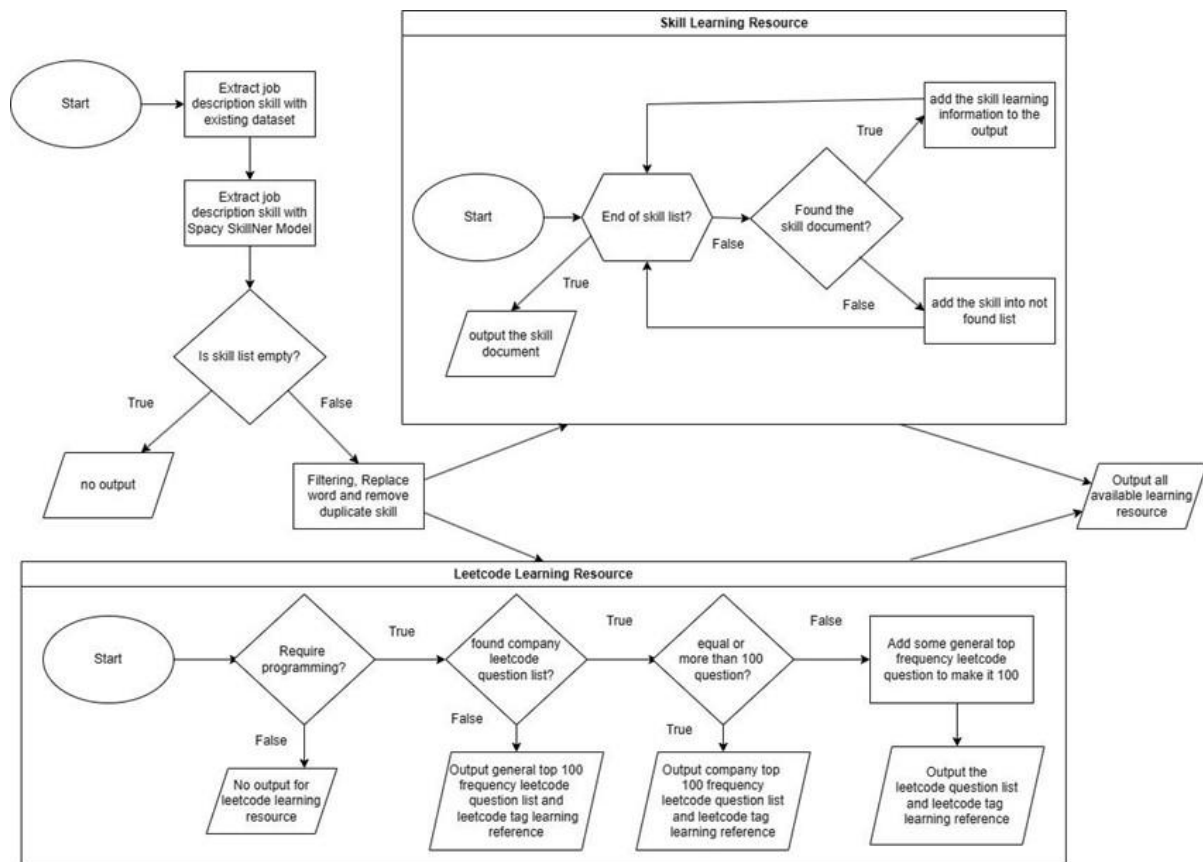


Figure 19. Flow diagram of learning resource recommender

The learning resource recommender makes use of a knowledge-based reasoning approach to identify the technical skills that are required from the selected job. Based on these technical skills, generate a document with the relevant learning resources for each of them using a web-based search. This process can be broken down into knowledge-based reasoning learning resource recommender which talks about knowledge acquisition and reasoning system and recommended technical interview questions based on company.

4.5.1 Knowledge-based Reasoning Learning Resource Recommender

4.5.1.1 Knowledge Acquisition

This project makes use of an existing natural language processing (NLP) library called SkillNER^[9] to automatically extract skills and certifications from a given text. We identified 1500 technical skills commonly used in job descriptions within the technology sector. For each of the skills, a web-based search was performed to find out 5 relevant learning resources and this result is saved for future retrieval. Each of the web-based search results is passed to SkillNER to identify the technical skills that are related to each of the skills. The related skills

will later be used to reason that if skill A is required, skill B which is closely related to skill A will most likely be required. For example, as shown in Figure 20, if a job requires user to have Flask, it is likely that user will need to know Python as well. This helps to build the knowledge base for our reasoning system.

Name	Search Keyword	Resource Path	Groups
flask	flask	skill/flask.html	[python]

Figure 20. An example of Flask being related to Python

4.5.1.2 Reasoning System

After the user selected a job, the learning resource recommender will start. First, SkillNER is applied on the selected job description to extract the technical skills required for this job. The technical skills and its related technical skills identified during the knowledge acquisition process will also be considered as the job's technical requirements. Second, SkillNER is applied on the candidate information and likewise, a list of technical skills and related technical skills will be generated. The 2 lists of technical skills will be compared, and missing skills will be identified. Based on the missing skills, the respective web-based search results with 5 relevant learning resources will be retrieved and populated into a word document.

4.5.1.3 Adding Learning Resources for New Technical Skills

As time passes, new technical skill keywords will be used in job descriptions. When a new technical skill keyword is identified, the system will proceed to generate relevant learning resources for the new skill. On the backend, manual process of verifying the generated content is required to ensure the accuracy, quality, relevancy to the new technical skills. If required, amendments should be made to ensure the contents accurately represent learning resources for the new skills. Once verified, the newly generated resources will need to be stored under the DataCollection/data/skill-learning-resource directory together with the rest of the other learning resources. This ensures the learning resources are made available for the system to use later. The final step is to add in the new technical skill name in the skill.csv file. Figure 21 shows the activity flow diagram when adding learning resources for new technical skills.

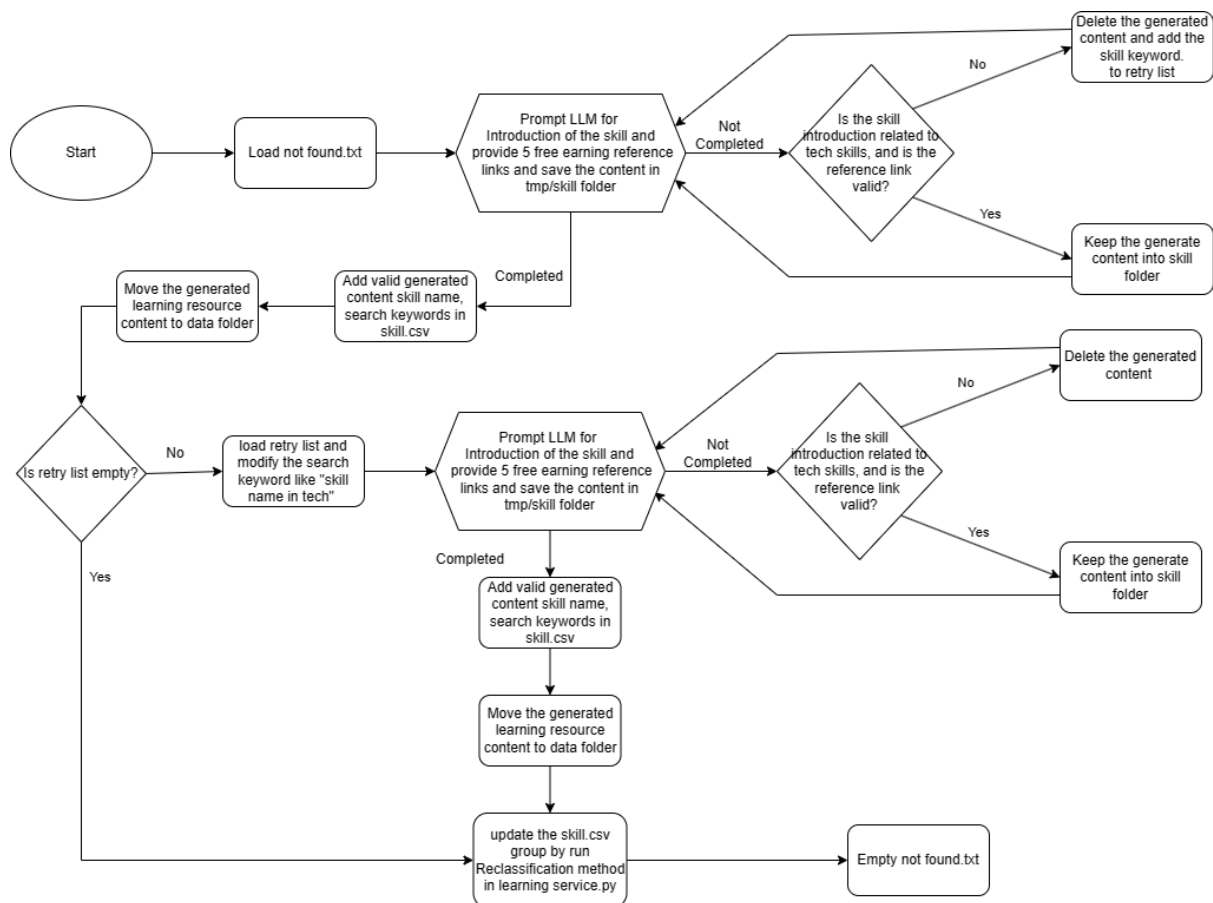


Figure 21. Activity flow diagram when adding learning resources for new technical skills

4.5.2 Recommended Technical Interview Questions based on company

4.5.2.1 Data collection

LeetCode questions are scraped by the question title, link to question, tags (category of questions), acceptance rate (number of accepted submissions out of total number of submissions), difficulty and number of times this question was asked in interviews for roles under the company are scraped by companies and saved in a csv for each company (e.g. AMD.csv). Figure 22 shows an example of technical interview questions scraped. This provides a better insight into the types of technical interview questions that will be asked during the interview.

	No	Title	Link	Tags	Acceptance	Difficulty	SYNAPSE PTE. LTD. Company Frequency	Overall Frequency
0	1	Two Sum	https://leetcode.com/problems/two-sum/description/	[Array][Hash Table]	52.20%	Easy	0	100%
1	2667	Create Hello World Function	https://leetcode.com/problems/create-hello-world-function/description/	NaN	81.10%	Easy	0	99.46%
2	88	Merge Sorted Array	https://leetcode.com/problems/merge-sorted-array/description/	[Array][Two Pointers][Sorting]	49.40%	Easy	0	98.92%

Figure 22. An example of technical interview questions scraped

Figure 23 shows the activity flow diagram for LeetCode data collection.

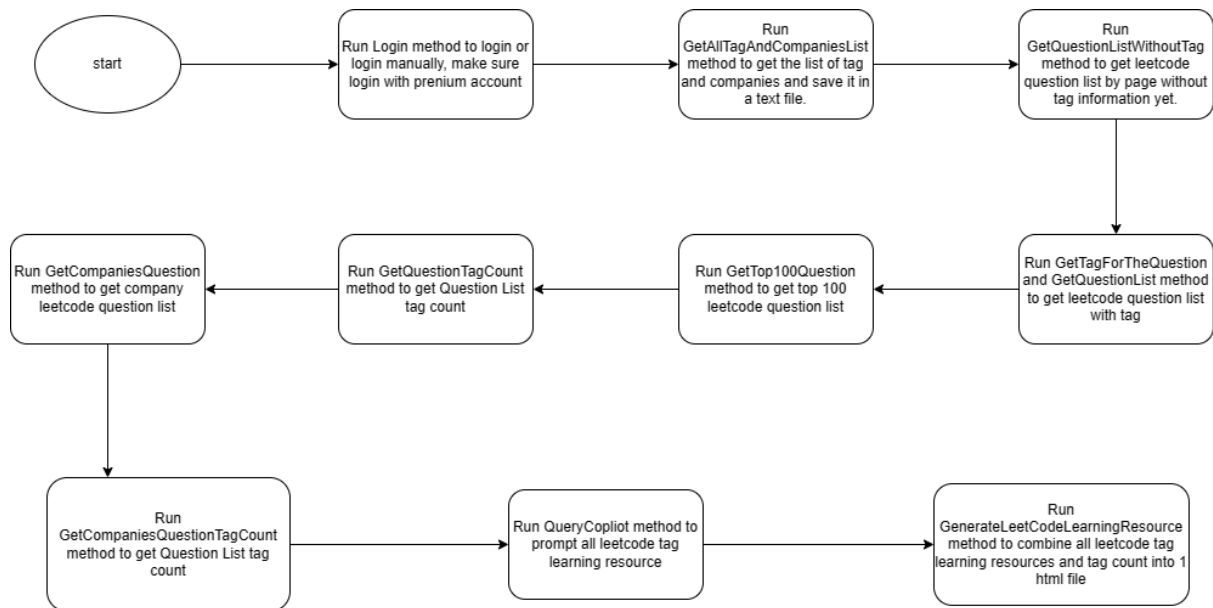


Figure 23. Activity flow diagram for LeetCode data collection

4.5.2.2 Recommended Technical Interview Questions

From the LeetCode questions that were scraped, we can tell the top 100 questions based on frequency that are asked during technical rounds. This information is presented in a table format and appended as part of the relevant learning resources document. Below the table, a list of learning resources for the technical questions will also be provided. This will help users better prepare for their upcoming interviews.

This chapter covered the project solutions for each subsystem used in this project. The following chapter will discuss about the project implementation and how the subsystems come together in a Streamlit application.

5. Project Implementation

This chapter covers the project implementation and how each subsystem is pieced together in a Streamlit application as a final product. It will cover the system design, followed by deployment architecture, continuous integration and continuous delivery (CI/CD) process and the Streamlit application.

5.1 System Design

Figure 24 shows the system design of the project and how each subsystem is connected to the Streamlit user interface (UI).

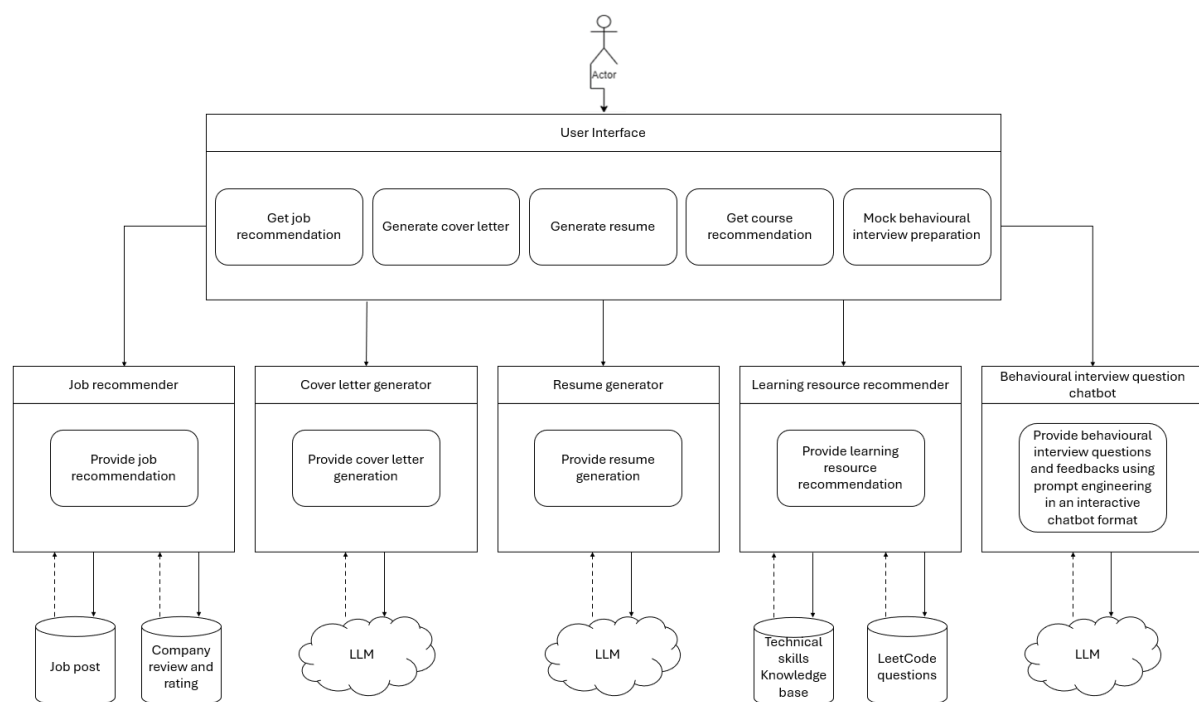


Figure 24. System design of the project and how each subsystem is connected to the Streamlit UI

The system design provides a brief overview of how the Streamlit UI is linked to each subsystem.

5.2 Deployment Architecture

The application is made up of various subsystems and each subsystem provides information for the Streamlit user interface. Each subsystem is viewed as a service. As each service is developed separately, individual service is packaged into its own container, allowing for independent scaling, and updating.

Kubernetes was selected as the orchestration tool to manage these containers efficiently across a cluster of machines.

Using a synchronous microservice architecture, the team has effectively segregated the responsibilities between services, hence improving continuous integration while reducing complexity.

1. **Containerization:** Each microservice was containerized using Docker, creating isolated environments for each service that include all necessary dependencies.
2. **Kubernetes Setup:** A Kubernetes cluster was set up using DigitalOcean (DO) Kubernetes to leverage DO capabilities like scalability and security.
3. **Service Deployment:** Each microservice is deployed as a separate pod within the Kubernetes cluster. Services are managed through deployments which allow the team to scale and update the pods without downtime.
4. **Service Discovery:** Kubernetes services were defined for each microservice to enable service discovery and easy communication between services. This is crucial for operations like connecting the Streamlit Backend application with the other services.
5. **Load Balancing:** Implemented load balancers to distribute traffic across pods for each service, enhancing performance and reliability.
6. **Configuration Management:** Used Kubernetes ConfigMaps and Secrets to manage configuration across different environments securely.

Figure 25 shows the deployment architecture of the project.

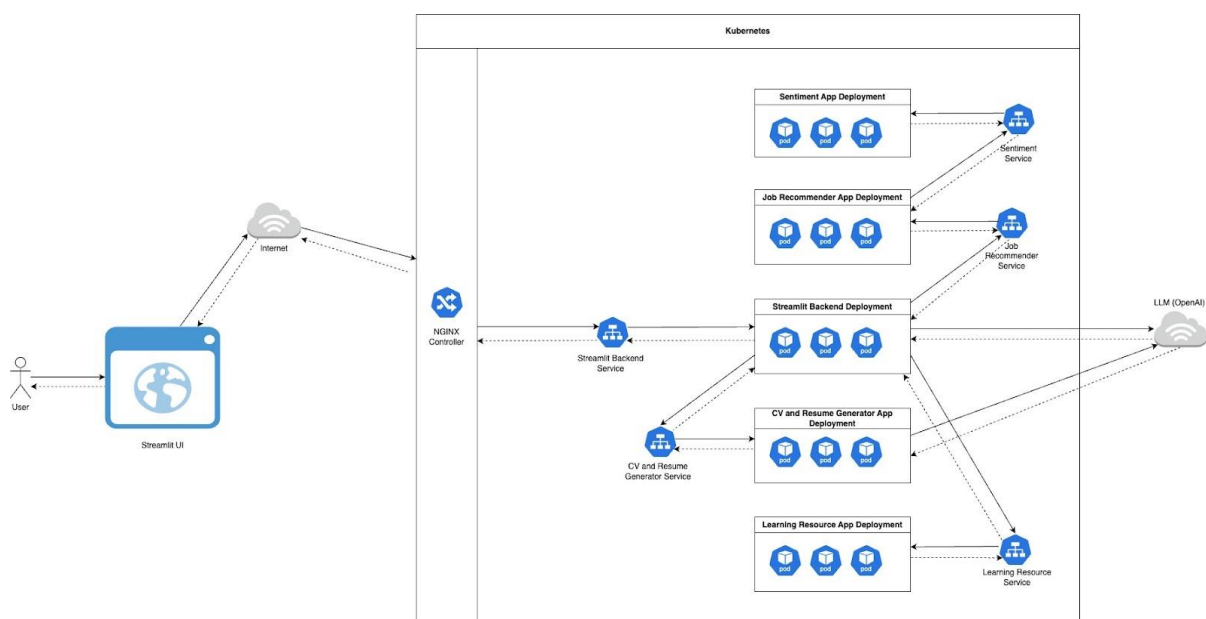


Figure 25. Deployment architecture of the project

5.3 Streamlit Application

Streamlit is an open-source Python framework that allows the project to be quickly turned into a web application ^[10]. The project makes use of Streamlit to build the following pages that will communicate with the services deployed.

Figure 26 shows a flow chart of the Streamlit application.

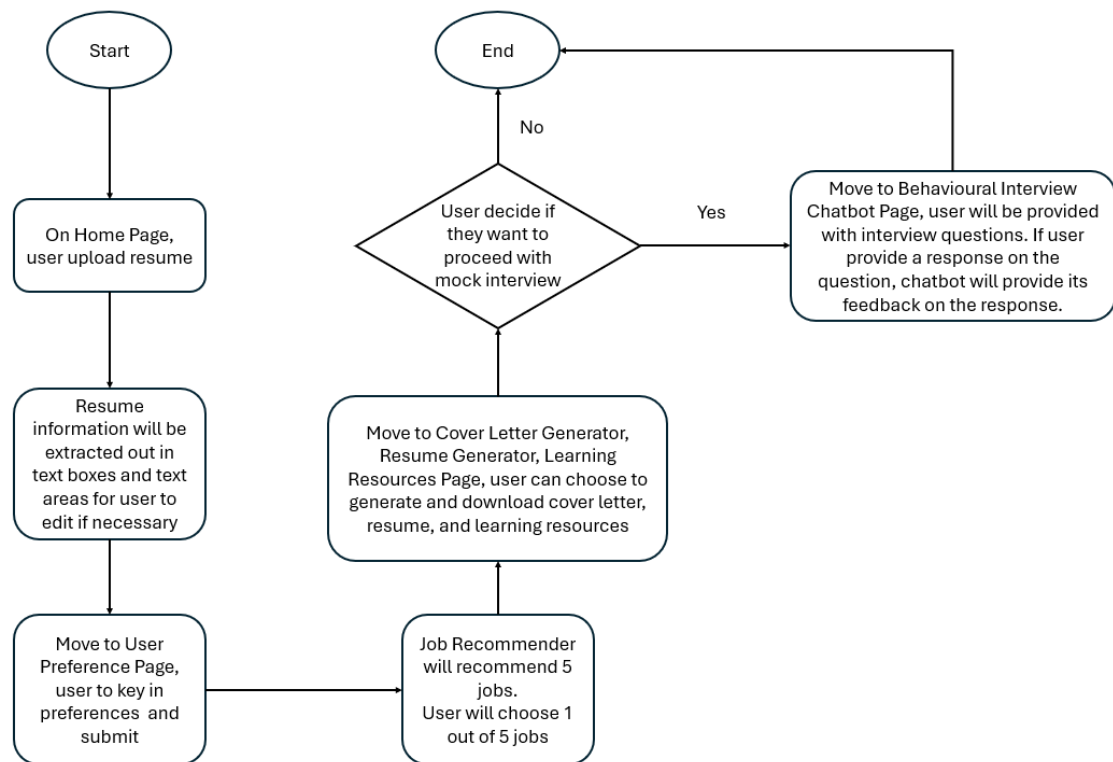


Figure 26. Flow chart of the Streamlit application

5.3.1 Home Page

At the start of the application, on the home page, user will be asked to upload their resume to proceed further. Figure 27 shows the home page of the application before the resume is uploaded in our UI storyboard.

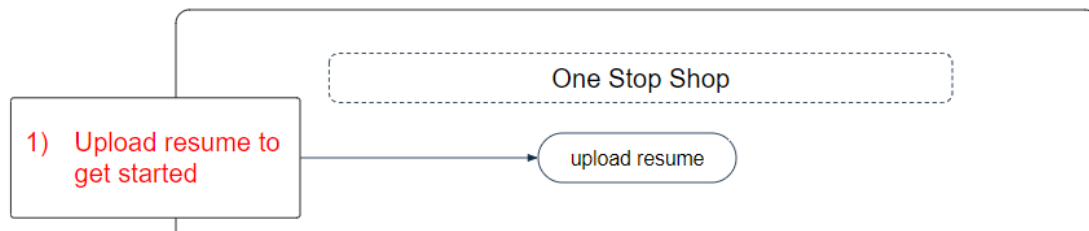


Figure 27. Home page of the application before the resume is uploaded in our UI storyboard

Once the user uploads their resume, the resume parser will extract the user's personal information, past work experiences, education history, and technical skills as explained in the earlier chapter. Figure 28 shows the extended home page where the user information is extracted and displayed in the text boxes in our UI storyboard.

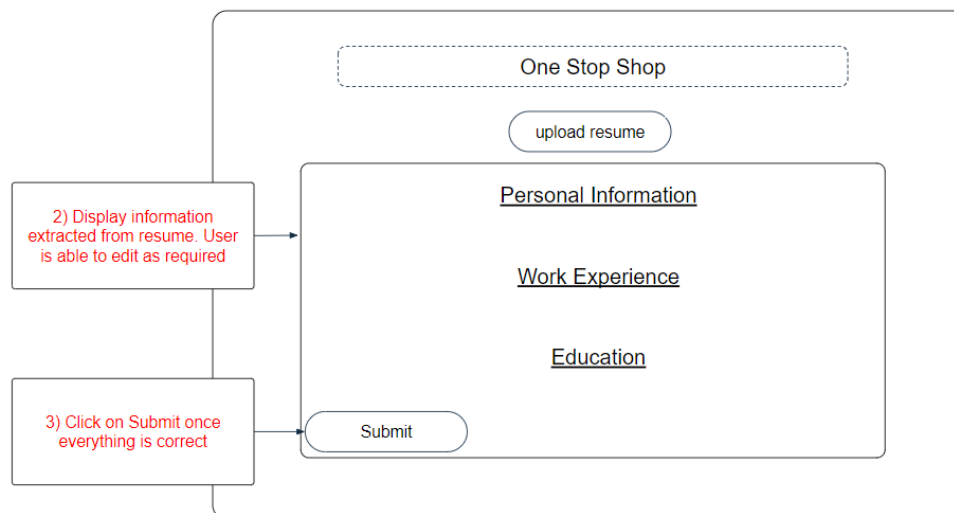


Figure 28. Extended home page where the user information is extracted and displayed in the text boxes in our UI storyboard

User will verify if the information is correct and make necessary edits before clicking on the submit button to continue to the next page.

5.3.2 User Preference Page

On the user preference page, user will fill in their job preferences. Figure 29 shows the user preference page before clicking on submit button in our UI storyboard.



Figure 29. User preference page before clicking on submit button in our UI storyboard

After user clicks on submit button on user preference page, job recommender service will start the process to find out the top 5 most relevant jobs. The 5 most relevant jobs will be shown at the bottom of the user preference page. User will have to choose one of the jobs to proceed to the next page. Figure 30 shows the user preference page with 5 jobs recommended in our UI storyboard.



Figure 30. User preference page with 5 jobs recommended in our UI storyboard

User will have to decide which job to proceed with and move to the next page.

5.3.3 Cover Letter Generator, Resume Generator, Learning Resource Recommender Page

On this page, user can click on cover letter generator to use cover letter generator service to generate the cover letter. It can be downloaded once it is ready. User can click on resume generator and the resume generator service will generate the resume. It can be downloaded once it is ready. User can also click on learning resource recommender and the learning resource recommender service will generate the learning resource document. It can be downloaded once it is ready.

Figure 31 shows the cover letter generator, resume generator, learning resource recommender page in our UI storyboard.

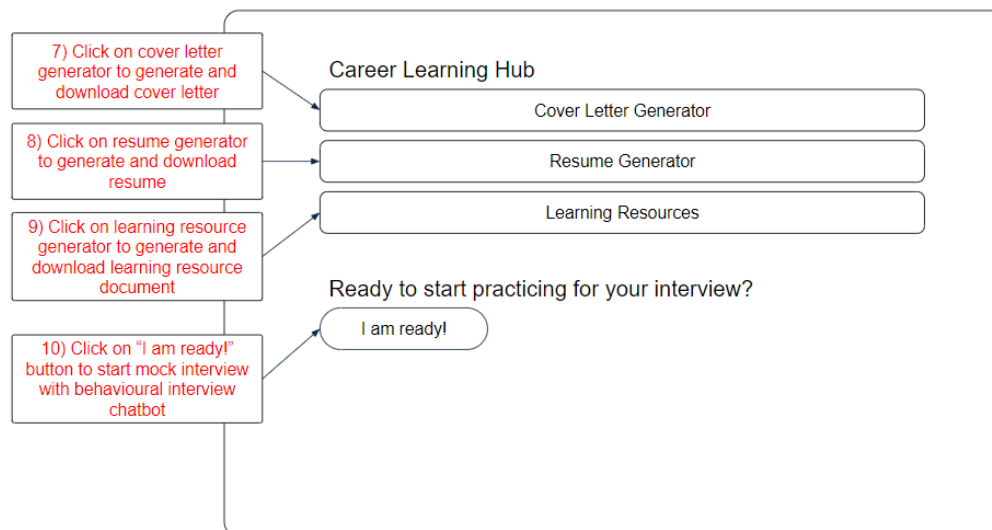


Figure 31. Cover letter generator, resume generator, learning resource recommender page in our UI storyboard

Lastly, user can click on the “I am ready!” button when the user is ready for the mock interview with the behavioural interview chatbot. This will load the behavioural interview chatbot page.

5.3.4 Behavioural Interview Chatbot Page

When the behavioural interview chatbot page loads, the behavioural interview question chatbot service will generate and provide user with a list of possible behavioural interview questions to practice and receive feedback from the chatbot. User can write down responses to the questions and click on the submit button to receive feedback from the chatbot. When the submit button is clicked, the behavioural interview question chatbot service will evaluate the user’s response and generate feedback for the user to improve their answer in the future.

Figure 32 shows the behavioural interview chatbot page in our UI storyboard.

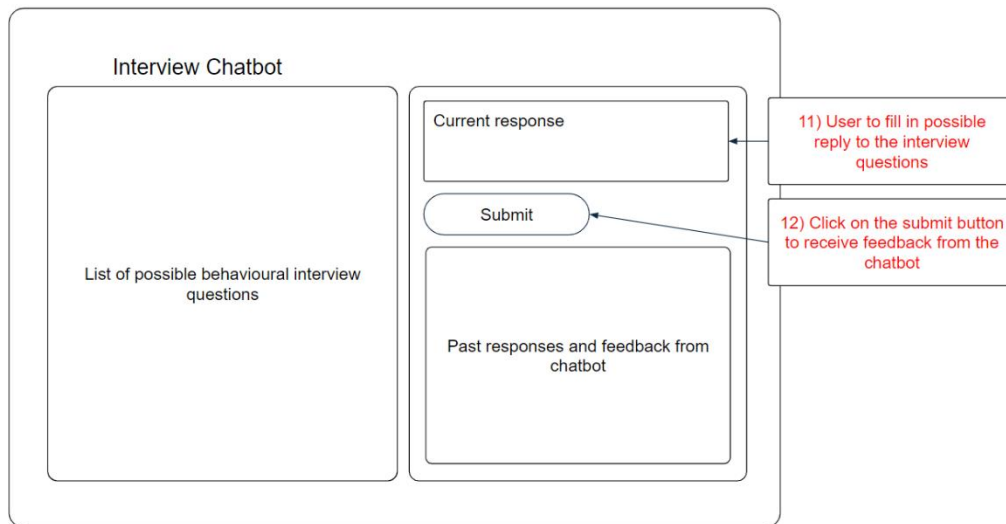


Figure 32. Behavioural interview chatbot page in our UI storyboard

This chapter covered the project implementation, detailing system design, followed by deployment architecture, CICD process and the Streamlit application. The following chapter will discuss about the project performance and validation.

6. Project Performance & Validation

The project has met the original aim to develop a single platform that recommend relevant jobs fitted to individual preferences, generate customized resume and cover letter, analyse, and identify technical skill gaps and recommend learning resources including a table detailing the type of technical questions to expect based on company and job title, and provide a behavioural interview chatbot for mock interviews.

This project makes use of knowledge-based reasoning techniques to identify pre-requisite technical skills that are required but not listed in the job description, and to recommend learning resources and generate learning resources document for the user to bridge the knowledge gap.

This project also uses a recommendation system to recommend top 5 most relevant jobs to the users using a weighted average score based on cosine similarity between user resume and job description, user review sentiment and user's company rating score from company reviews and ratings, and user preference score from company reviews.

Lastly, this project uses LLM and prompt engineering to develop a chatbot that identifies possible behavioural questions for user to practice in mock interview sessions with the chatbot and provides timely feedback to user on how to improve their answers in the future.

Therefore, the project has achieved the project objectives.

6.1 Home Page

On the home page, user will first be asked to upload their resume. Figure 33 shows the home page of the Streamlit application before resume is uploaded.

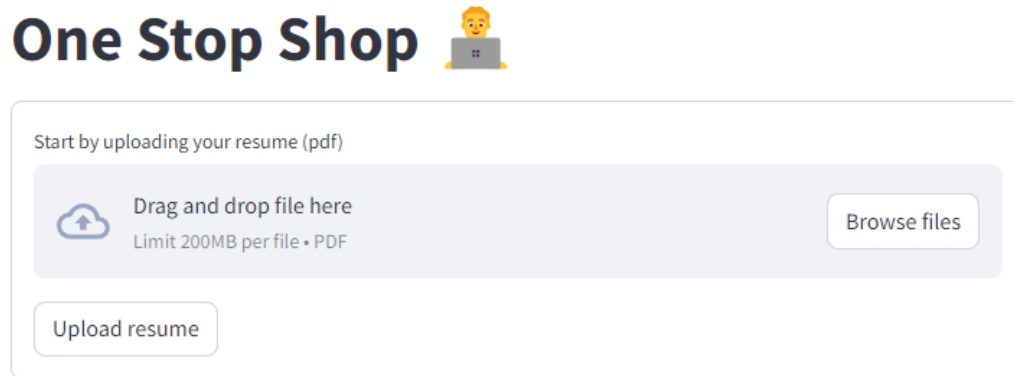


Figure 33. Home page of the Streamlit application before resume is uploaded

After the resume is uploaded, user will need to verify the information and make necessary edits before clicking on the submit button to continue to the user preference page.

Figure 34 shows the home page of the Streamlit application after resume is uploaded.

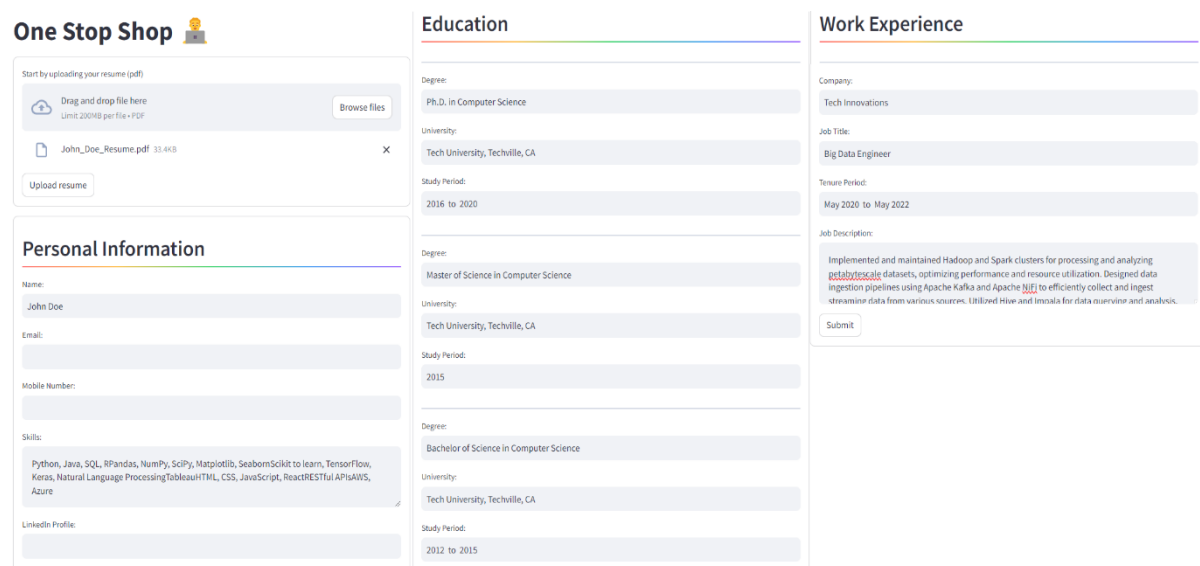


Figure 34. Home page of the Streamlit application after resume is uploaded

6.2 User Preference Page

On this page, user will fill in their preferences. Figure 35 shows the user preference page of the Streamlit application before clicking on submit button.

Previous

Job Preference

Job Position:

Cloud Engineer

Desired Location

Desired Employment Type

Desired Seniority Level

Required Job Experience (no. of years)

0

Desired Aspect(s):

work life and flexibility job aspect

Minimum Salary:

0

Submit

Figure 35. User preference page of the Streamlit application before clicking on submit button

After submit button is clicked, job recommender will recommend 5 most relevant jobs. Figure 36 shows the user preference page of the Streamlit application with the 5 recommended jobs.

Desired Seniority Level

Required Job Experience (no. of years)

Desired Aspect(s):

Minimum Salary:

Submit

These are the recommended jobs:

- AVENSYS CONSULTING PTE. LTD.
- ELLIOTT MOSS CONSULTING PTE. LTD.
- IT CONSULTANCY & SERVICES PTE LTD
- KS CONSULTING PTE. LTD.
- PERSOLKELLY SINGAPORE PTE. LTD.

Figure 36. User preference page of the Streamlit application with the 5 recommended jobs

The job details are shown when the job is clicked to expand as shown in Figure 37.

These are the recommended jobs:

AVENSYS CONSULTING PTE. LTD.

Job Title

cloud engineer

Job Location

VERTEX, 33 UBI AVENUE 3 408868

Job Employment Type

Contract

Job Seniority

Executive

Job Experience Level

5 years exp

Job Industry

Information Technology

Salary Range

\$5,000 to \$10,000

Number of Applications

9 applications

Application Closing Date

Closing on 28 Mar 2024

Job Description

Avensys is a reputed global IT professional services company he We are currently looking to hire Cloud Data Engineer (AWS, Data This is an exciting opportunity to expand your skill set, achie Responsibilities:

A Cloud Data Engineer specializing in AWS, Databricks, and Info

- Design and architect data storage solutions, including databa
- Create, manage, and optimize data pipelines for ingesting, pr
- Integrate data from various sources, both internal and extern
- Develop ETL (Extract, Transform, Load) processes to cleanse,
- Monitor and optimize data processing and query performance in
- Implement security best practices and data encryption methods
- Implement automation for routine tasks, such as data ingestio
- Maintain clear and comprehensive documentation of data infras
- Collaborate with cross-functional teams, including data scien
- Identify and resolve data-related issues and provide support
- Optimize AWS, Databricks, and Informatica resource usage to c
- Stay up-to-date with AWS, Databricks, Informatica IDMC servic

Requirements / Qualifications

- Bachelor's or master's degree in computer science, data engin
- Minimum 5 years of experience in data engineering, with exper
- Proficiency in programming languages such as Python, Java, or
- Evaluate potential technical solutions and make recommendatio
- Strong knowledge of SQL and NoSQL databases.

- Bachelor's or master's degree in computer science, data engin
- Minimum 5 years of experience in data engineering, with exper
- Proficiency in programming languages such as Python, Java, or
- Evaluate potential technical solutions and make recommendatio
- Strong knowledge of SQL and NoSQL databases.
- Familiarity with data modeling and schema design.
- Excellent problem-solving and analytical skills.
- Strong communication and collaboration skills.
- AWS certifications (e.g., AWS Certified Data Analytics - Spec

Preferred Skills:

- Experience with big data technologies like Apache Spark and H
- Knowledge of containerization and orchestration tools like Do
- Familiarity with data visualization tools like Tableau or Pow
- Understanding of DevOps principles for managing and deploying
- Experience with version control systems (e.g., Git) and CI/CD
- Knowledge of data governance and data cataloging tools, espe

WHAT'S ON OFFER

You will be remunerated with an excellent base salary and entit To submit your application, please apply online or email your U

CONSULTANT DETAILS

Consultant Name: Preethi Kanthappan

Reg No: R1765546

Avensys Consulting Pte Ltd

EA Licence: 12C5759

Privacy Statement: Data collected will be used for recruitment

Select This Job

ELLIOTT MOSS CONSULTING PTE. LTD.

IT CONSULTANCY & SERVICES PTE LTD

KS CONSULTING PTE. LTD.

PERSOLKELLY SINGAPORE PTE. LTD.

Figure 37. User preference page of the Streamlit application with job details expanded

If the user is interested in this role, user can click on the “Select This Job” to proceed to the next page.

6.3 Cover Letter Generator, Resume Generator, Learning Resource Recommender Page

On this page, the user will be shown cover letter generator, resume generator, learning resources, and a “I am ready!” button. Figure 38 shows the cover letter generator, resume generator, learning resource recommender page of the Streamlit application.

Career Learning Hub

Previous

Cover Letter Generator

Resume Generator

Learning Resources

Ready to start practicing for your interview?

I am ready!

Figure 38. Cover letter generator, resume generator, learning resource recommender page of the Streamlit application

Clicking on cover letter generator will reveal a list of information used to generate the cover letter. User can make necessary edits on the information before generating the cover letter with more accurate information. Figure 39 shows the example of cover letter generator with the list of information to generate the cover letter. Clicking on the “Generate Cover Letter” button will generate the cover letter. It will be downloadable once generated.

Cover Letter Generator

Company Name:
AVENSYS CONSULTING PTE. LTD.

What role are you applying for?
cloud engineer

Who are you emailing?

Name:
John Doe

Email:

Mobile Number:

Skills:
Python, Java, SQL, RPandas, NumPy, SciPy, Matplotlib, SeabornScikit to le

LinkedIn Profile:

I have experience in...

I am excited about the job because...

I am passionate about...

Generate Cover Letter

Figure 39. Example of cover letter generator with the list of information to generate the cover letter

When resume generator is clicked, it will show a button to generate resume as shown in Figure 40. Clicking on the “Generate Resume” will trigger the resume generator service to generate a resume.

The screenshot shows the 'Career Learning Hub' interface. At the top, there's a navigation bar with a 'Previous' button and a dropdown menu currently set to 'Cover Letter Generator'. Below this, the 'Resume Generator' section is highlighted with a red border. It contains the text 'Generate a resume template based on the information you have filled in:' and a 'Generate Resume' button. Below the highlighted section is a 'Learning Resources' dropdown menu. At the bottom, there's a prompt 'Ready to start practicing for your interview?' with an 'I am ready!' button.

Figure 40. Example of resume generator to generate the resume.

A list of the resume information will be shown on screen before user download the resume. If user is satisfied with the resume content, user can click on the “Download Resume” button to download the resume. Figure 41 shows the list of resume information before user downloads the resume.

The screenshot shows the 'Resume Generator' interface with the 'Generate Resume' button highlighted. Below the button, the 'Generated Resume' section is displayed. It includes contact information: 'John Doe Email: - Mobile Number: - LinkedIn: - Skills:'. A list of skills is shown: Python, Java, SQL, Pandas, NumPy, SciPy, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras, Natural Language Processing, Tableau, HTML, CSS, JavaScript, React, RESTful APIs, AWS, and Azure. To the right, the 'Education' section lists three degrees: Ph.D. in Computer Science (2016 to 2020), Master of Science in Computer Science (2015), and Bachelor of Science in Computer Science (2012 to 2015). The 'Work Experiences' section lists 'Big Data Engineer' at 'Tech Innovations' from 'May 2020 to May 2022'. A detailed description of the role is provided, mentioning the implementation and maintenance of Hadoop and Spark clusters, data ingestion pipelines using Apache Kafka and Apache NiFi, and data querying and analysis using Hive and Impala. Below the work experience, there is a 'Download Resume' button. At the bottom, there's a 'Learning Resources' dropdown menu and a prompt 'Ready to start practicing for your interview?' with an 'I am ready!' button.

Figure 41. List down the resume information before user downloads the resume

Clicking on the learning resources reveals a button “generate learning resources” as shown in Figure 42.

Career Learning Hub

Previous

Cover Letter Generator

Resume Generator

Learning Resources

generate learning resources

I am ready!

Figure 42. Clicking on the learning resources reveals a button “generate learning resources”

When the “generate learning resources” button is clicked, the learning resource recommender service will start to generate the learning resource document. Figure 43 shows the generated learning resource documents on Streamlit application.

Career Learning Hub

Previous

Cover Letter Generator

Resume Generator

Learning Resources

generate learning resources

Cloud

Certainly! In a nutshell, cloud computing refers to accessing servers, software, and databases over the Internet, eliminating the need for managing physical servers or running applications on local machines^[214].

Here are five free resources where you can learn more about cloud computing:

1. [Cloudflare Learning Center](#): Understand cloud basics, explore different cloud models, and learn how the cloud works.
2. [Google Cloud](#): Dive into on-demand availability of computing resources and grasp the essentials of cloud computing.
3. [IBM Cloud Computing](#): Explore cloud resources, virtual servers, data storage, and more with pay-per-use pricing.
4. [PhoenixNAP](#): Learn about cloud virtualization, scalability, hybrid cloud, DevOps, and automation.
5. [Coursera Cloud Computing Courses](#): Discover essential concepts like IaaS, PaaS, and SaaS, and explore different layers of cloud computing.

Feel free to explore these resources and enhance your understanding of cloud technology!

Career Learning Hub

Previous

Cover Letter Generator

Resume Generator

Learning Resources

generate learning resources

Happy learning! 🎉

AVENSYS CONSULTING PTE. LTD. 100 Leetcode QuestionList

No	Title	Link	
0	1 Two Sum	https://leetcode.com/problems/two-sum/description/	[Array][Table]
1	2667 Create Hello World Function	https://leetcode.com/problems/create-hello-world-function/description/	NaN
2	88 Merge Sorted Array	https://leetcode.com/problems/merge-sorted-array/description/	[Array][Pointer][Sort]
3	20 Valid Parentheses	https://leetcode.com/problems/valid-parentheses/description/	[String]
4	146 LRU Cache	https://leetcode.com/problems/lru-cache/description/	[Hash][Link][Deque][Double-Link]
5	9 Palindrome Number	https://leetcode.com/problems/palindrome-number/description/	[Math]

5	9	Palindrome Number	https://leetcode.com/problems/palindrome-number/description/	[Math]
6	5	Longest Palindromic Substring	https://leetcode.com/problems/longest-palindromic-substring/description/	[String][Dynamic Programming]
7	121	Best Time to Buy and Sell Stock	https://leetcode.com/problems/best-time-to-buy-and-sell-stock/description/	[Array][Dynamic Programming]
8	3	Longest Substring Without Repeating Characters	https://leetcode.com/problems/longest-substring-without-repeating-characters/description/	[Hash][String][Sliding Window]
9	42	Trapping Rain Water	https://leetcode.com/problems/trapping-rain-water/description/	[Array][Pointer][Dynamic Programming][Stack][Monotonic Stack]
10	56	Merge Intervals	https://leetcode.com/problems/merge-intervals/description/	[Array][Sort]
11	314	Binary Tree Vertical Order	https://leetcode.com/problems/binary-tree-vertical-order-traversal/description/	[Hash][Tree][First S]

Download Learning Resource

Ready to start practicing for your interview?

I am ready!

Figure 43. Generated learning resource documents on Streamlit application

The user can download this information as a document for future references by clicking the “Download Learning Resource” button.

Clicking on “I am ready” button will proceed to the next page where user can have mock interview with the behavioural interview chatbot.


6.4 Behavioural Interview Chatbot Page

On this page, user will be provided with a list of possible behavioural interview questions based on the job description and user resume information. This will allow the user to practice interview questions with the chatbot.

Figure 44 shows the list of questions the interview chatbot provides.

Interviewer ChatBot

Previous



Possible Interview Questions:

1. Can you walk us through your experience in implementing and maintaining Hadoop and Spark clusters for processing and analyzing petabyte-scale datasets?
2. How have you optimized performance and resource utilization in your previous projects involving big data technologies like Hadoop and Spark?
3. Could you explain your approach to designing data ingestion pipelines using Apache Kafka and Apache NiFi to efficiently collect and ingest streaming data from various sources?
4. What tools and techniques have you utilized to enable fast and interactive data exploration for business users using Hive and Impala?
5. How have you collaborated with data scientists, data analysts, and other cross-functional teams in your previous roles to identify data requirements and deliver solutions that meet business and analytical needs?
6. Can you provide an example of a data quality and integrity issue you encountered and how you resolved it in your previous projects?
7. How do you approach optimizing and tuning database queries and data processing workflows to achieve high performance and efficiency?
8. What experience do you have with cloud platforms such as AWS or Azure in the context of data engineering and analytics?
9. How have you applied your knowledge of machine learning algorithms and AI technologies in your previous projects as a Data Engineer?
10. Can you share an example of a challenging data-related issue you faced and how you troubleshooted and resolved it in a timely and accurate manner?

Please provide me with your response, I will evaluate and provide feedback based on your answers.

write here: e.g. regarding question 1, <my response> ...


Submit

Figure 44. List of questions the interview chatbot provides

Providing a response of an interview question to the interview chatbot will allow it to evaluate the answer and provide feedback on how the user should improve their answers. Figure 45 shows the interview chatbot evaluate a user response and provide feedback on how to improve their answer.

Interviewer ChatBot

Previous




Possible Interview Questions:

1. Can you walk us through your experience in implementing and maintaining Hadoop and Spark clusters for processing and analyzing petabyte-scale datasets?
2. How have you optimized performance and resource utilization in your previous projects involving big data technologies like Hadoop and Spark?
3. Could you explain your approach to designing data ingestion pipelines using Apache Kafka and Apache NiFi to efficiently collect and ingest streaming data from various sources?
4. What tools and techniques have you utilized to enable fast and interactive data exploration for business users using Hive and Impala?
5. How have you collaborated with data scientists, data analysts, and other cross-functional teams in your previous roles to identify data requirements and deliver solutions that meet business and analytical needs?
6. Can you provide an example of a data quality and integrity issue you encountered and how you resolved it in your previous projects?
7. How do you approach optimizing and tuning database queries and data processing workflows to achieve high performance and efficiency?
8. What experience do you have with cloud platforms such as AWS or Azure in the context of data engineering and analytics?
9. How have you applied your knowledge of machine learning algorithms and AI technologies in your previous projects as a Data Engineer?
10. Can you share an example of a challenging data-related issue you faced and how you troubleshooted and resolved it in a timely and accurate manner?

Please provide me with your response, I will evaluate and provide feedback based on your answers.

write here: e.g. regarding question 1, <my response> ...

Submit



It's great that you are seeking guidance on how to handle a question that you may not have direct experience with. In this situation, it's important to be honest with the interviewer about your current level of familiarity with Hive and Impala, but also demonstrate your willingness and ability to quickly learn and adapt to new technologies.

You could respond by acknowledging that while you may not have direct experience with Hive and Impala, you have a strong foundation in data exploration techniques and tools in general. You could mention your experience with similar tools or technologies, and highlight your ability to quickly grasp new concepts and technologies.

For example, you could say something like:

"While I haven't had the opportunity to work extensively with Hive and Impala, I have a solid understanding of data exploration techniques and tools. I have experience with other data processing and analysis tools such as SQL, Python, and Tableau, which have allowed me to effectively explore and analyze data for business insights. I am confident in my ability to quickly learn and adapt to new technologies like Hive and Impala to enable fast and interactive data exploration for business users."

Remember, interviewers appreciate honesty and a willingness to learn. By showcasing your adaptability and eagerness to expand your knowledge, you can leave a positive impression even if you don't have direct experience with a particular technology.

Keep up the positive attitude and continue to showcase your enthusiasm for learning and growth in your interviews. You're on the right track, and with your determination, you can overcome any challenge that comes your way! Good luck!

With regards to the question "What tools and techniques have you utilized to enable fast and interactive data exploration for business users using Hive and Impala?" I have not touched on such technology, will you be able to help provide me with a plausible answer to respond back to the interviewer if I do not actually have such knowledge?




Figure 45. Interview chatbot evaluate a user response and provide feedback on how to improve their answer

The project has delivered a Streamlit application that uses job recommender, resume generator, cover letter generator, learning resource recommender and an interactive behavioural interview chatbot. This chapter covered the project validation and showed that the project has met its original project objectives. The following chapter will cover on the project conclusion.

7. Project Conclusion

CareerButterfly is a single platform application built on Streamlit and uses microservices to recommend jobs fitted to user preferences, generate cover letter, resume and learning resources for our users, and provides an interactive behavioural interview chatbot for mock interview sessions. It combines multiple job seeking tools together in a single platform and helps to streamline the process of finding a job suitable for our users.

This project makes use of recommendation system, knowledge-based reasoning techniques and uses LLM and prompt engineering to develop a chatbot and generate documents.

Job recommender is a recommendation system used to recommend top 5 most relevant jobs to the users using a weighted average score based on cosine similarity between user resume and job description, user review sentiment and user's company rating score from company reviews and ratings, and user preference score from company reviews.

Cover letter generator and resume generator both use LLM and well-crafted effective prompts that act as instruction to guide the LLM to achieve the desired output customised for their application.

Learning resources recommender uses knowledge-based reasoning techniques to recommend learning resources and generate learning resources document for the user to bridge the knowledge gap.

Behavioural question interview chatbot is developed with the help of LLM and prompt engineering that identifies possible behavioural questions for user to practice in mock interview sessions with the chatbot and provides timely feedback to user on how to improve their answers in the future.

7.1 Future Implementation

In the future, the system can undergo the following enhancements to improve the overall experience.

1. As our data collection is done manually, the next step is to automate the data collection process to refresh every 2 weeks.
2. Currently the application mainly focuses on the technology sector, the next step is to broaden the scope of the fields by searching for more job titles to have a more comprehensive application.
3. Automate process of adding in new technical skills and learning resources. Currently this system requires manual intervention to keep the list of skills up to date.

Appendices

Appendix A: Project Proposal

GRADUATE CERTIFICATE: Intelligent Reasoning Systems (IRS) PRACTICE MODULE: Project Proposal

Date of proposal:
25 February 2024
Project Title:
Career Butterfly - The One Stop Shop for Your Tech Career
Sponsor/Client: <i>(Name, Address, Telephone No. and Contact Name)</i>
Institute of Systems Science (ISS) at 25 Heng Mui Keng Terrace, Singapore NATIONAL UNIVERSITY OF SINGAPORE (NUS) Contact: Mr. GU ZHAN / Lecturer & Consultant Telephone No.: 65-6516 8021 Email: zhan.gu@nus.edu.sg
Background/Aims/Objectives:
<p>Due to the recent tech layoffs affecting thousands of employees, our team came up with the plan to build a One-Stop-Shop for tech employees and people who are breaking into tech to minimize their down time and optimize their career seeking process and experience.</p> <p>All users need to do is to attach their resume in our platform, select options to filter the search to cater to their preference. We will</p> <ol style="list-style-type: none">1. Screen through the resume and suggest recommendations to help improve the resume.2. Based on user selected options for their next role (e.g. next job title, current experience level, next industry, next role's minimum salary)<ol style="list-style-type: none">a. Identify skill sets that are missing from user's resume,b. We will recommend jobs similar to user profile.3. Provide a fully written cover letter and resume recommended for users to submit when applying for the suggested jobs.4. Identify technical skill gaps and provide a document with learning resources for users to bridge their knowledge gaps.5. Provide mock interview questions for users to practice and retrieve suggested answers to these questions in an interactive chatbot format.
Requirements Overview:
<ul style="list-style-type: none">• Research ability• Programming ability• System integration ability

Resource Requirements (please list Hardware, Software and any other resources)

Hardware proposed for consideration:

- laptop

Software proposed for consideration:

- SciPy for Cosine similarity
- Selenium, beautiful soup, scrapy
- Pandas, NumPy, Matplotlib, scikit-learn, nltk, spacy, TensorFlow, Streamlit
- LLM, Keras Chatbot
- Docker, Kubernetes
- GitHub, BitBucket

Number of Learner Interns required: (Please specify their tasks if possible)

Zachary Chua, Clarence Ng, Wilson Ang, Soo Kuan Yong

Methods and Standards:

Procedures	Objective	Key Activities
Requirement Gathering and Analysis	The team should meet with ISS to scope the details of project and ensure the achievement of business objectives.	<ol style="list-style-type: none">1. Gather & Analyse Requirements2. Define internal and External Design3. Prioritize & Consolidate Requirements4. Establish Functional Baseline
Technical Construction	<ul style="list-style-type: none">- Develop code to data scrape from job portal- Develop code for data analysis and processing- Develop sentiment analysis on company reviews to provide a score which will be fed in as an input to the automatic job application model- Develop a recommendation system that recommends fully written cover letter to send along with resume- Develop a recommendation system that recommends short courses/mini projects for user to pick up required skill sets to improve chances- Chatbot for Mock interview question for specific roles based on your resume. Gets instant feedback on how to improve your answer- Recommend job application	<ol style="list-style-type: none">1. Setup Development Environment2. Understand the System Context, Design3. Perform Coding4. Conduct Unit Testing
Integration Testing and acceptance testing	<ul style="list-style-type: none">- End-to-end testing- Integration test	<ol style="list-style-type: none">1. End-to-end testing2. Integration test
Acceptance Testing	To obtain ISS user acceptance that the system meets the requirements.	<ol style="list-style-type: none">1. Plan for Acceptance Testing2. Conduct Training for Acceptance Testing3. Prepare for Acceptance Test Execution4. ISS Evaluate Testing5. Obtain Customer Acceptance Sign-off

Delivery	To deploy the system into production (ISS standalone server) environment.	<ol style="list-style-type: none"> 1. Software must be packed by following ISS's standard 2. Deployment guideline must be provided in ISS production (ISS standalone server) format 3. Production (ISS standalone server) support and troubleshooting process must be defined.
-----------------	---	---

Team Formation & Registration

Team Name: Career Butterfly
Project Title (repeated): Career Butterfly - The One Stop Shop for Your Tech Career
System Name (if decided):
Team Member 1 Name: Ang Han Chuan Wilson
Team Member 1 Matriculation Number: A0291657R
Team Member 1 Contact (Mobile/Email): 96905452 / e1333153@u.nus.edu
Team Member 2 Name: Ng Min Teck
Team Member 2 Matriculation Number: A0291213N
Team Member 2 Contact (Mobile/Email): 8845 4484 / ng_min_teck@u.nus.edu (e1330352@u.nus.edu)
Team Member 3 Name: Zachary Chua Feng Kwan
Team Member 3 Matriculation Number: A0180556B
Team Member 3 Contact (Mobile/Email): 92224279 / zacharychua@u.nus.edu (e0284042@u.nus.edu)
Team Member 4 Name: Soo Kuan Yong
Team Member 4 Matriculation Number: A0291188U
Team Member 4 Contact (Mobile/Email): 93850368 / e1330327@u.nus.edu

For ISS Use Only		
Programme Name:	Project No:	Learner Batch:
Accepted/Rejected/KIV:		
Learners Assigned:		
<p>Advisor Assigned:</p> <p>Contact: Mr. GU ZHAN / Lecturer & Consultant Telephone No.: 65-6516 8021 Email: zhan.gu@nus.edu.sg</p>		

Appendix B: Mapped System Functionalities against knowledge, techniques and skills of modular courses: MR, RS, CGS

The proposed project has developed, integrated, and demonstrated multiple aspects from the following 3 technique groups:

1. **Knowledge discovery & (big) data mining techniques:** recommendation systems (RS)
Job recommender is developed as a recommendation system used to recommend top 5 most relevant jobs to the users using a weighted average score based on cosine similarity between user resume and job description, user review sentiment and user's company rating score from company reviews and ratings, and user preference score from company reviews.

2. **Decision Automation:** Knowledge based reasoning techniques (MR and CGS)

Knowledge-based reasoning technique is used to identify pre-requisite technical skills that are required but not listed in the job description and recommend learning resources and generate learning resources document for the user to bridge the knowledge gap.

3. **System designed with cognitive techniques or tools:** chatbot (CGS)

Chatbot is developed using LLM and prompt engineering that identifies possible behavioural questions for user to practice in mock interview sessions with the chatbot and provides timely feedback to user on how to improve their answers in the future.

Appendix C: Installation and User Guide

CareerButterfly System: User Guide

[ProjectReport](#)

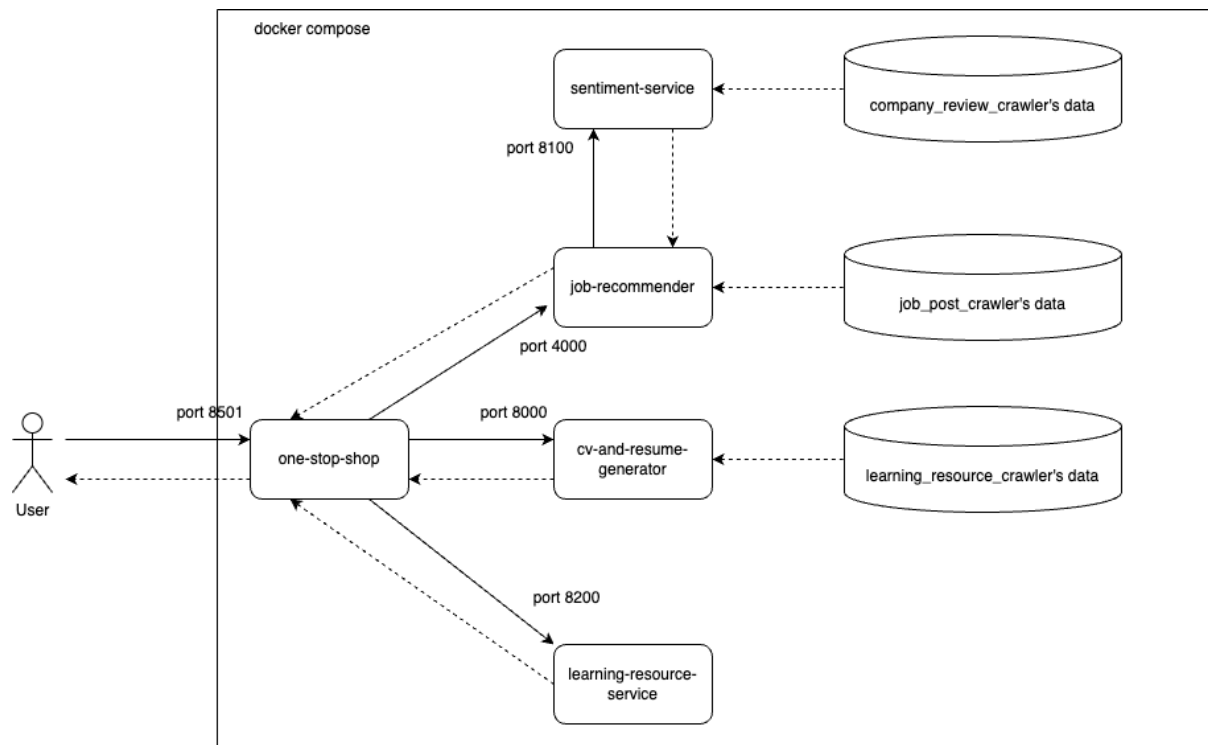
[1] [CareerButterfly System](#)

Pre-requisite

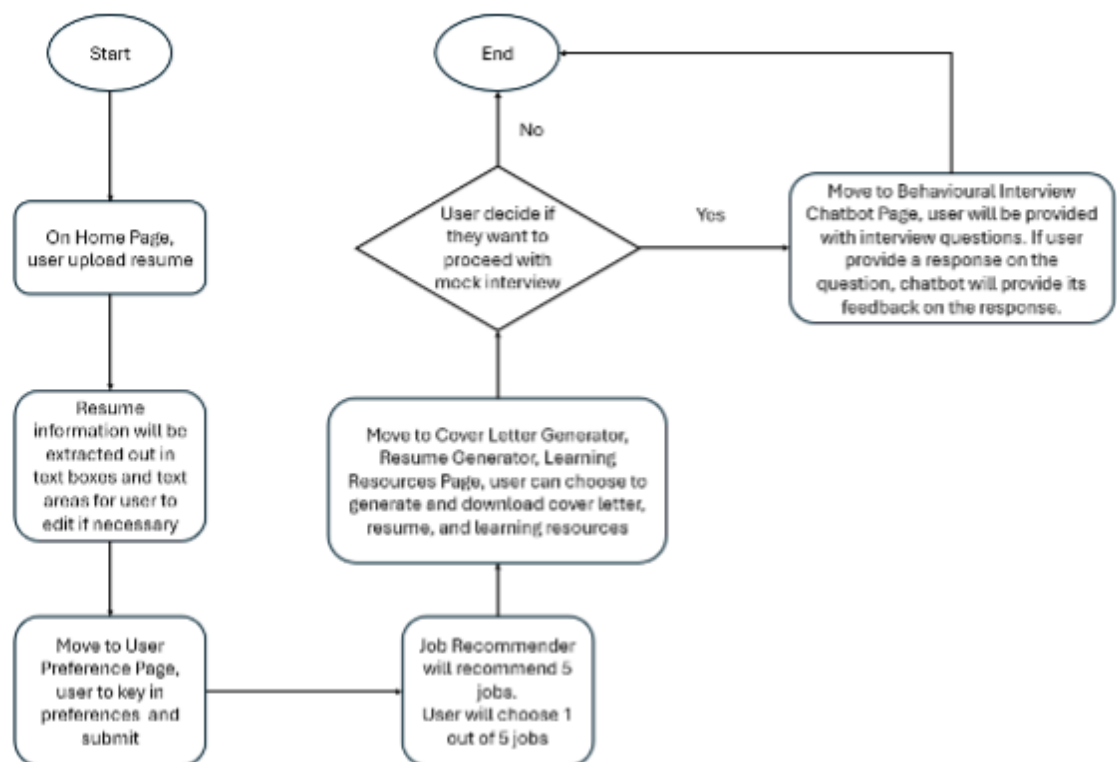
- Docker desktop - <https://www.docker.com/products/docker-desktop/>
-

Setup

- Navigate to project directory
cd SourceCode/CareerButterflySystem
- Build docker images with docker compose
docker compose up
- The image below illustrates the connected containers and volumes in the
docker compose



- Navigate to url, <http://localhost:8501>, with browser
- The following diagram outlines the process flow chart of using the application



Environment variables

Each individual project may have one or more environment variables. The environment variables are hardcoded into their respective Dockerfile. The following are the environment variables that you may choose to change.

- `cv_and_resume_generator`
 - `OPEN_AI_KEY`
The api key from OpenAI. Change this if you wish to use your own api key.
 - `one_stop_shop`
 - `OPEN_AI_KEY`
The api key from OpenAI. Change this if you wish to use your own api key.
-

Volume mounts

Each container, created by docker-compose, may have a volume mount. The volume mount binds the data, that was created by the data collectors.

- `sentiment_service`
 - `DataCollection/company_review_crawler/data` is mounted to `/data`
 - `job-recommender`
 - `DataCollection/job_post_crawler/data` is mounted to `/data`
 - `learning-resource-service`
 - `DataCollection/learning_resource_crawler/data` is mounted to `/data`
-

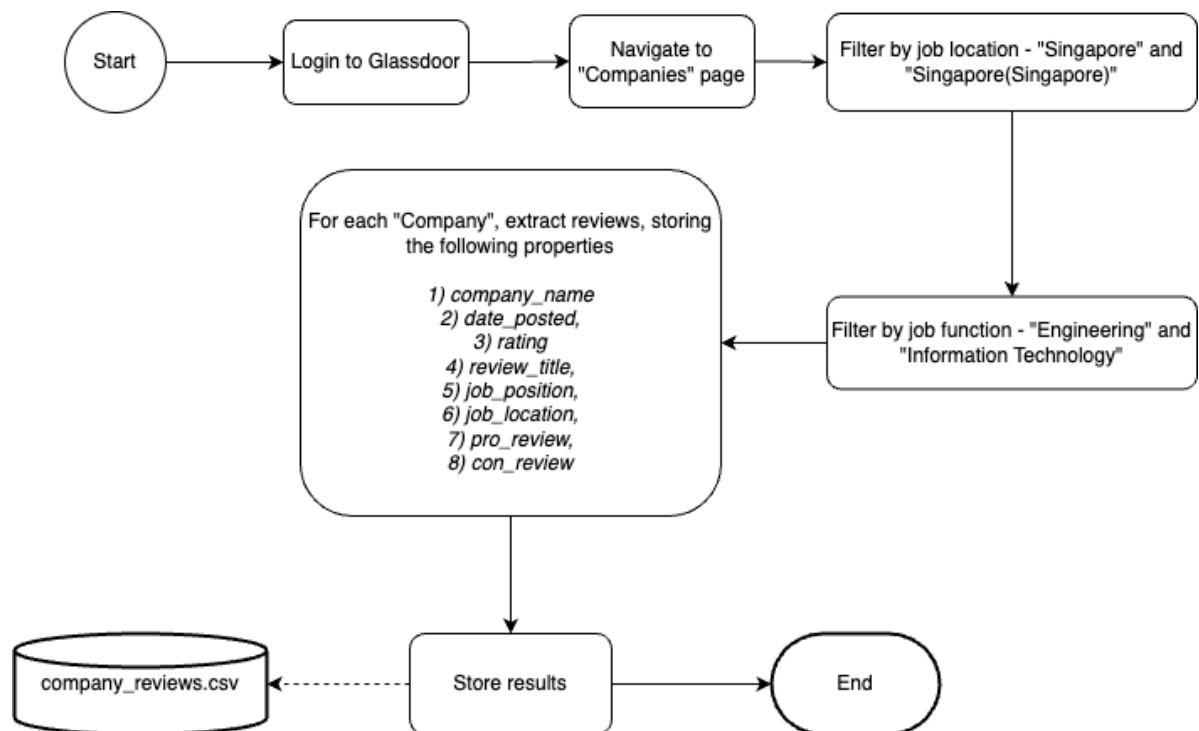
Pre-requisite

- Conda 23 - <https://conda.io/projects/conda/en/latest/user-guide/getting-started.html>
 - Python 3 - <https://www.python.org/downloads/>
 - Jupyter - <https://jupyter.org/install>
-

Setup

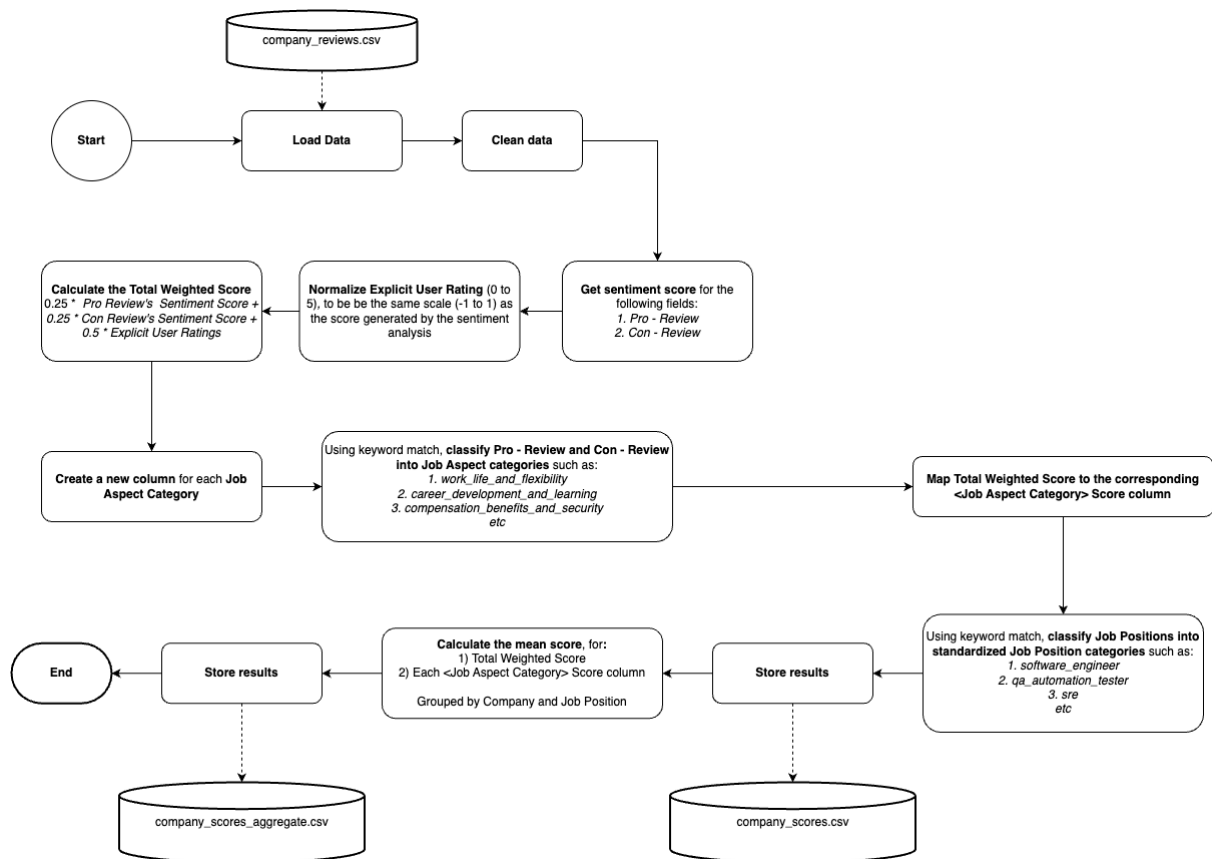
1. Scraping company reviews

- Navigate to project directory
`cd SourceCode/DataCollection/company_review_crawler`
- Create conda environment
`conda env create -f environment.yml`
- Activate conda environment
`conda activate company_review_crawler`
- Run all cells in `crawl_reviews.ipynb`
 - Note: Results will be saved in the data folder which will later be used in rating company review
- This diagram outlines the process of collecting company reviews



1.1. Rating company reviews

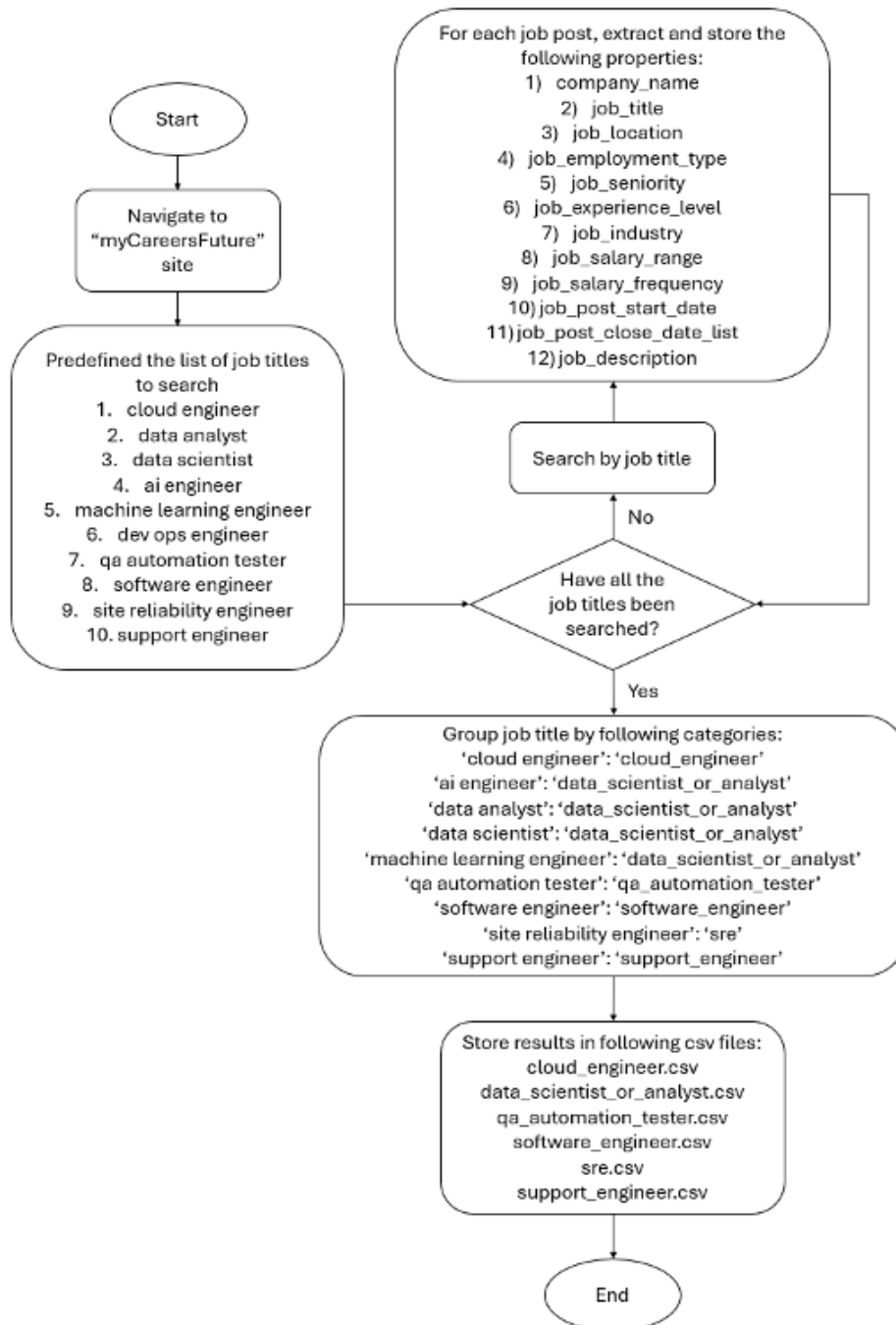
- Run all cells in company_reviews_sentiment.ipynb
 - Note: Results will be saved in the data folder which will be used by SourceCode/CareerButterflySystem/sentiment_service
- This diagram outlines the process of rating company reviews



2. Scraping job posts

- Navigate to project directory
`cd SourceCode/DataCollection/job_post_crawler`
- Create conda environment
`conda create --name job_post_crawler python=3.10`
- Activate conda environment
`conda activate job_post_crawler`
- Install python dependencies
`pip install -r requirements.txt`
- Create 2 folders and rename them as "new_data" and "data" respectively
 - Note: Ignore this step if folders already exist
- Run data_collection_job_information.py
`python data_collection_job_information.py`
 - Results will be saved in the new_data folder which will be used by later in
- Run rename_files.py
`python rename_files.py`

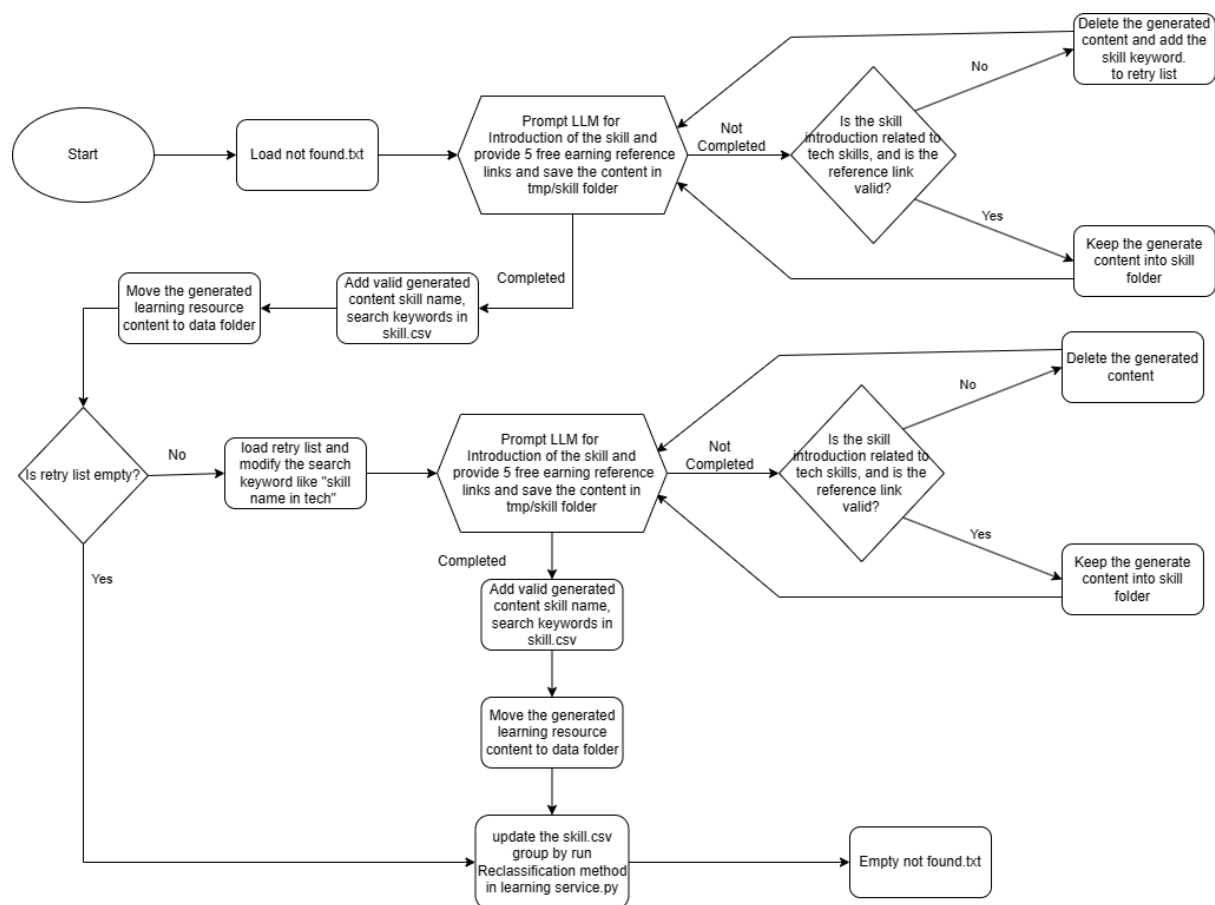
- Results will be saved in the data folder which will be used by
SourceCode/CareerButterflySystem/job_recommender
- This diagram outlines the process of collecting job posts



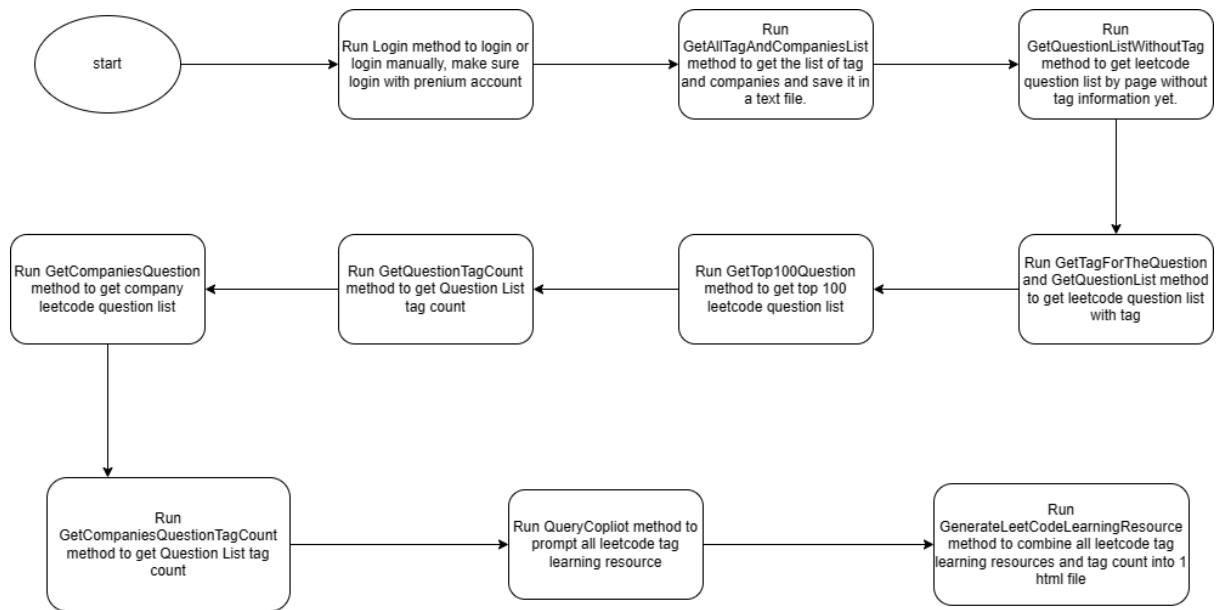
3. Scraping learning resources

- Navigate to project directory
`cd SourceCode/DataCollection/learning_resource_crawler`
- Create conda environment
`conda env create -f environment.yml`
- Activate conda environment
`conda activate learning_resource_crawler`
- Run all cells in leetcode.ipynb for scraping leetcode content
 - Note: Results will be saved in the data folder which will later be used in `SourceCode/CareerButterflySystem/learning_resource_service`
- Run all cells in copilot query1.ipynb for adding new skill learning resource
 - Note: Results will be saved in the data folder which will later be used in `SourceCode/CareerButterflySystem/learning_resource_service`

This diagram outlines the process of collecting of the skill learning resources



This diagram outlines the process of collecting of the leetcode learning resources



Appendix D: Individual Reports

Individual Report

Soo Kuan Yong (A0291188U)

Personal Contribution:

I have contributed to multiple aspects throughout the project. I helped to chip in during project ideation, prepared project proposal document, prepared project report, managed the team making sure we are on track, developed Python script to scrape job post data using Selenium, developed job recommender from scratch in Python using cosine similarity and created a custom weighted average score that makes use of company reviews sentiment analysis service that Zachery developed and cosine similarity score to recommend relevant jobs.

Developed the resume parser using PyPDF2, GPT 3.5 Turbo model with well-crafted prompt to extract key candidate information. Developed the base interactive behavioural interview chatbot using GPT 3.5 Turbo model and prompt engineering to identify a list of possible interview questions, be able to interact and provide feedback on user response on how to improve their answers in a timely manner.

Most useful skills learned:

I find the most useful soft skills from this project experience is the project team management skills that I have picked up along the way. I was given the opportunity by the team to lead this project. It provided me with a sense of responsibility to ensure this project is managed and done on time.

Throughout the project, I was able to pick up new technical skills along the way. I learned how to extract information from PDF documents using PyPDF2 library. I also learned how to make use of GPT 3.5 Turbo model and craft effective prompts that acts as instruction to guide the LLM to achieve the desired output. I learned how to develop a web scrapper using Selenium to scrap job post information online.

How can I apply the skills I have learnt:

With the technical skills, I can explore ways to develop an interactive chatbot that takes in multiple documents and respond to user query at my workplace. This will help significantly reduce the time and effort required to go through older documentations when we have questions.

With the soft skills, I will be able to better understand my project manager's perspective, be able to help the team break down larger projects into smaller portions and foster better communication with the team.

Individual Report

Ang Han Chuan Wilson (A0291657R)

Personal Contributions:

Planning and development of user interface for one stop shop application. User Interface is created using Streamlit; a free and open source Python library for building web applications. User Interface consists of multiple pages with interactive elements such as text boxes, buttons etc. I integrated the user interface with multiple services developed by other team members and carried out debugging to fine tune the user interface and services to improve user experience.

Development of the CV and Resume Generator which is powered by OpenAI GPT 3.5 model. The generator accepts key information such as target company, target job role etc and it will insert this information into the prompts which are fed to the model. The model is instructed to retrieve the key information to generate either a CV or Resume and if there is missing information, to replace them with placeholders. In addition, I integrated and refined the Resume Parser developed by Kuan Yong as part of the UI utilities. Lastly, I integrated and refactor the Interviewer Chatbot developed by Kuan Yong as one of the pages of the UI and to have it utilize the same GPT 3.5 model.

Most useful skills learned:

From this project, I have learnt several skills. Firstly, I learnt more about the functionality and characteristics of Large Language Models. I learnt how to carry out prompt engineering to manipulate the behaviour of the Large Language Models and improve the accuracy of the generated output. I also learnt various frameworks and methodology such as Q&A, prompt engineering etc when developing or refining the various LLM based services (CV and Resume Generator, Interviewer Chatbot etc).

Secondly, I get to experience how to go about creating a User Interface from scratch, starting from creating a storyboard for each of the pages, laying out the steps or options that users can undertake for each page (e.g. pressing the back button will return them back to the previous page), integrating the User interface with the various services via HTTP requests etc. Since Streamlit is based on Python, it makes developing and debugging easier for a beginner like me.

How can I apply the skills I have learnt:

Since my work experience so far have mostly been with the backend's development and maintenance, with my newfound experience in creating a User Interface, I can try to contribute to the creation of UIs for my various projects.

As my work involves the use of knowledge graphs and ontologies, I can look into using LLMs to maintain and extend my knowledge graphs and ontologies dynamically based on any new information that is added to the digital twins without the need for any manual intervention. This is because the addition of concepts to ontologies have always been a manual task that is tedious and non-efficient.

Individual Report

Ng Min Teck (A0291213N)

Personal Contributions:

My contributions to the project were significant and diverse, playing a crucial role in its success. Below is a detailed account of my contributions:

- **Implemented Learning Resource Service:** I successfully implemented a learning resource service that provides users with valuable educational materials, enhancing their learning experience.
- **Created Technical Skill Dataset:** I created a comprehensive dataset of technical skills keywords, which suit the needs of the project
- **Compiled all Technical SkillFree Learning Resource Reference Link:** Compiled a list of free learning resources reference links for all technical skills in the dataset, by prompting Copilot.
- **Compiled LeetCode Information:** I gathered the most frequently asked company-specific LeetCode questions, along with all LeetCode tags learning guide for prepare Leetcode technical interview

These contributions demonstrate my dedication to enhancing our platform and providing users with the tools they need to succeed.

Most useful skills learned:

- **Web Scrapping:** This is my first time doing web scraping. I used Selenium with SeleniumBase to extract data from LeetCode. Additionally, I employed web scraping techniques to prompt Copilot and save its responses in HyperText Markup Language (HTML) format for later use.
- **Technical Skill Keyword Research:** Technical skill keyword research can be an arduous process, it is a backward-chain process. As a backend software engineer, I am familiar with some backend development skills. However, there is a vast landscape to explore: different programming languages, frameworks, libraries, databases, infrastructure components, and system architecture elements. Considerations range from message queues and distributed computing to load balancers and monitoring tools. Each tech role requires a specific skill set, making this exploration both challenging and rewarding.
- **Information Extraction:** In addition to utilizing my custom-built keyword dataset, I leverage SkillNER—one of the popular named entity recognition tools—to extract both

hard skills and soft skills from job descriptions. This approach helps me identify any technical skill keywords that I might have overlooked.

- Flask: While I seldom delve into RESTful, this project has significantly boosted my confidence in my understanding of RESTful principles and implementation of RESTful.

How can I apply the skills I have learnt:

I am considering a transition into AI, which differs significantly from my previous and current roles. One area that stands out to me is information extraction—it has broad applicability and can be used effectively in various contexts. Additionally, I find the process of prompting LLM (large language model) particularly intriguing. Also, I plan to research a procedural generation for an interior environment for twin cities, so information extraction of the keyword to generate the 3D model will be very useful.

Individual Report

Zachary Chua Feng Kwan (A0180556B)

Personal Contributions:

During the initial phase of the project, I played a pivotal role in brainstorming and shaping the idea of this project. This ideation stage set the foundation for the targeted functionalities that would later define our project's uniqueness in the competitive landscape.

I am responsible for the crucial task of collecting company reviews and rating the reviews. This process involves using a python library, Selenium, to scrape user reviews from Glassdoor. The process involved writing robust scraping scripts that not only fetched the data efficiently but also handled exceptions and errors to ensure data integrity.

I led the rating of the scraped company reviews which involves scoring companies based on the implicit score and explicit score. Implicit scores are derived through sentiment analysis of the pro reviews and con reviews, while explicit scores are derived directly from the user's rating. This process provided nuanced insights into employee satisfaction across companies, significantly enriching the data points for our recommender system. As job seekers are usually interested in certain aspects of a job, such as "work-life balance", "diversity", "compensation", I took a step further by classifying the reviews using a pre-defined dictionary of job categories and keywords match.

Finally, I implemented a service that enables Kuan Yong's job recommender service to fetch the aggregated company scores based on certain job aspect and job role.

My research on cloud solutions was instrumental in choosing DigitalOcean as our deployment platform. I evaluated various cloud services based on cost, scalability, and reliability. DigitalOcean's Kubernetes-based offerings were chosen due to their robustness and ease of scaling, making them ideal for our application needs.

In the deployment phase, I assisted in configuring and managing Kubernetes on DigitalOcean. This involved setting up Docker containers, creating Kubernetes clusters, and managing service deployments. My efforts ensured that the application was scalable, with high availability and minimal downtime, traits crucial for handling the expected traffic and data load.

To help the team get up to speed in deploying their application, I took charge in the design and implementation of the CI/CD architecture. I designed and implemented a CI/CD pipeline using GitHub actions which facilitated a streamlined workflow for integration and deployment. This implementation significantly enhanced our team's productivity by allowing for real-time deployment of code changes, reducing downtime in our integration and development process.

Most useful skills learned:

Conducting sentiment analysis of company reviews allowed me to deepen my understanding and practical skills in NLP. I became proficient in using Python's NLTK library to analyze text data, categorize sentiments, and interpret complex datasets. This skill is vital for roles involving data analysis, machine learning, and AI-driven technologies.

Throughout the project, I developed my problem-solving skills by addressing various technical challenges, from data collection to deployment. Additionally, my role in project ideation and design refined my project management skills, including planning, team collaboration, and execution, preparing me for leadership roles in technology projects.

How can I apply the skills I have learnt:

Scoring reviews based on implicit and/or explicit ratings can be a potential use case for many areas. Sentiment analysis holds immense potential across various industries by enabling the extraction of insights from vast amounts of unstructured text data, such as social media posts, customer reviews, or news articles. It provides a powerful means to leverage unstructured data for strategic advantage and responsive decision-making.

Being in the cryptocurrency industry, one such use case is by effectively analyzing public sentiment through tweets, I can gain insights into market trends and potential shifts in cryptocurrency prices before they are fully reflected in the market, potentially leading to more informed investment decisions.

Appendix E: List of Abbreviations

Abbreviations	Definition
LLM	Large Language Model
URA	Urban Redevelopment Authority
GPT	Generative Pre-trained Transformer
NLP	Natural Language Processing
PDF	Portable Document Format
CICD	Continuous Integration and Continuous Delivery
UI	User Interface
DO	DigitalOcean
HTML	HyperText Markup Language
Q&A	Question and Answer
CV	Curriculum Vitae

Appendix F: List of Figures

Figure 1. Recent layoff trends ^[6]	7
Figure 2. Flow diagram of job recommender system	10
Figure 3. Flow diagram of the data collection process from myCareersFuture using Selenium	11
Figure 4. Activity diagram that outlines the data collection process from Glassdoor using Selenium	12
Figure 5. Prompt to extract user's personal information, work experiences, education history, and technical skills	13
Figure 6. Location of a job post is given as an address	14
Figure 7. List of Postal District and corresponding Postal Sector	15
Figure 8. Regions of Singapore based on Postal District	15
Figure 9. Activity diagram that outlines this	17
Figure 10. An example of user review	18
Figure 11. Information on GPT-3.5 models	21
Figure 12. Prompt used to generate resume for the user	22
Figure 13. Flow diagram of resume generator	22
Figure 14. Prompt used to generate cover letter for the user	22
Figure 15. Flow diagram of cover letter generator	23
Figure 16. Prompt to generate behavioural interview questions based on candidate information and job description	23
Figure 17. Prompt to get GPT model to evaluate user response and generate feedback to that response	24
Figure 18. Flow diagram of behavioural interview chatbot	24
Figure 19. Flow diagram of learning resource recommender	25
Figure 20. An example of Flask being related to Python	26
Figure 21. Activity flow diagram when adding learning resources for new technical skills	27
Figure 22. An example of technical interview questions scraped	27
Figure 23. Activity flow diagram for LeetCode data collection	28
Figure 24. System design of the project and how each subsystem is connected to the Streamlit UI	29
Figure 25. Deployment architecture of the project	30
Figure 26. Flow chart of the Streamlit application	31
Figure 27. Home page of the application before the resume is uploaded in our UI storyboard	32

Figure 28. Extended home page where the user information is extracted and displayed in the text boxes in our UI storyboard	32
Figure 29. User preference page before clicking on submit button in our UI storyboard	33
Figure 30. User preference page with 5 jobs recommended in our UI storyboard	33
Figure 31. Cover letter generator, resume generator, learning resource recommender page in our UI storyboard	34
Figure 32. Behavioural interview chatbot page in our UI storyboard.....	35
Figure 33. Home page of the Streamlit application before resume is uploaded	37
Figure 34. Home page of the Streamlit application after resume is uploaded	37
Figure 35. User preference page of the Streamlit application before clicking on submit button	38
Figure 36. User preference page of the Streamlit application with the 5 recommended jobs.....	39
Figure 37. User preference page of the Streamlit application with job details expanded	40
Figure 38. Cover letter generator, resume generator, learning resource recommender page of the Streamlit application	40
Figure 39. Example of cover letter generator with the list of information to generate the cover letter.....	41
Figure 40. Example of resume generator to generate the resume.....	42
Figure 41. List down the resume information before user downloads the resume	42
Figure 42. Clicking on the learning resources reveals a button “generate learning resources”	43
Figure 43. Generated learning resource documents on Streamlit application	43
Figure 44. List of questions the interview chatbot provides	44
Figure 45. Interview chatbot evaluate a user response and provide feedback on how to improve their answer.....	45

Appendix G: List of Table

Table 1. Examples of resume generators in the market and their limitations	6
Table 2. Examples of cover letter generators in the market and their limitations	6
Table 3. List of keywords to determine each job aspect.....	19

Appendix H: References

- [1] "Resume builder for 2024 | Free Resume Builder | Novorésumé." Novorésumé. Accessed: Apr. 21, 2024. [Online]. Available: <https://novoresume.com/>
- [2] "Online Resume builder to help you land the job (try for free)." Resume.io. Accessed: Apr. 21, 2024. [Online]. Available: <https://resume.io/>
- [3] "Zety – Professional Resume & Cover Letter Tools For Any Job." Zety. Accessed: Apr. 21, 2024. [Online]. Available: <https://zety.com/>
- [4] "Top Resume and Cover Letter Tools for 2024." myperfectresume. Accessed: Apr. 21, 2024. [Online]. Available: <https://www.myperfectresume.com/>
- [5] "Grammarly: Free AI Writing Assistance." grammarly. Accessed: Apr. 21, 2024. [Online]. Available: <https://www.grammarly.com/>
- [6] "Outlook For The Tech Sector In Singapore." Ministry of Manpower. Accessed: Apr. 21, 2024. [Online]. Available: <https://stats.mom.gov.sg/Pages/Outlook-For-The-Tech-Sector-In-Singapore.aspx>
- [7] R. Ang. "More tech layoffs in 2024 likely, due to rising costs, greater automation, say experts" The Straits Times. Accessed: Apr. 21, 2024. [Online]. Available: <https://www.straitstimes.com/business/more-tech-layoffs-in-2024-likely-due-to-rising-costs-and-greater-automation-say-experts>
- [8] "List of Postal District." Urban Redevelopment Authority. Accessed: Apr. 21, 2024. [Online]. Available: https://www.ura.gov.sg/Corporate/-/media/Corporate/Property/PMI-Online/List_Of_Postal_Districts.pdf
- [9] "AnasAito/SkillNER: A (smart) rule based NLP module to extract job skills from text." GitHub. Accessed: Apr. 21, 2024. [Online]. Available: <https://github.com/AnasAito/SkillNER?tab=readme-ov-file>
- [10] "Streamlit documentation." Streamlit. Accessed: Apr. 21, 2024. [Online]. Available: <https://docs.streamlit.io/>