COMP4801 Final Year Project
UniLife
Personalized Career Consulting System

Hung Ka Hing, Jin Shang, Xu Haozhou, Zhao Yiming

Department of Computer Science
Faculty of Engineering
The University of Hong Kong, Pokfulam Road

Index Terms: Personalized Career Consulting System, Natural Language Processing

1. Introduction

1.1. Project Rationale

The transitional phase from higher education to the professional world can be challenging for students, particularly in terms of mapping their career path. More than 90% of university students face two primary issues: a feeling of uncertainty towards their future career and a lack of a structured career path to guide their decision-making process [1]. This situation often results in students missing opportunities and regretting their decisions in retrospect. Therefore, there is a pressing need for a system that can provide personalized career guidance from the earliest stages of their university life, aiding them in shaping a clear vision for their future.

1.2. Project Overview

The Personalized Career Consulting System (PCCS) aims to assist university students in navigating their career paths by providing comprehensive and personalized guidance. The PCCS is structured around three core components: the conversation system, the recommendation system, and the path generation system. The conversation system is designed to elicit information about the student's interests, abilities, resources, and career goals through an interactive dialogue. This information serves as the foundation for the recommendation system, which analyzes the gathered data to propose potential career that align with the user's profile. The path generation system then generates and visualizes the proposed career path(s), providing a detailed timeline that includes key career milestones and deadlines. Together, these components work in synergy to offer a clear, actionable roadmap to career success, enabling students to make informed and strategic decisions about their future.

1.3. Project Objectives

The primary objectives of this project are:

1) To design an interactive conversation system that elicits students' career interests, abilities, resources, and aspirations using a structured question bank and a sophisticated NLP model.
2) To develop a recommendation system that processes the information collected by the conversation system and suggests potential career(s) aligning with the students' profile.

3) To create a path generation system that organizes key career milestones and deadlines into an intuitive and interactive career path(s) timeline for each student.
4) To evaluate the system's effectiveness and user-friendliness through user experience surveys and accuracy assessments of the recommendation system.

1.4. Relevance to Computer Science

This project is highly relevant to the field of Computer Science, specifically in the areas of NLP and machine learning. The project utilizes these advanced technologies to understand user responses, generate interest matrices, and recommend relevant jobs. This is a practical application of theoretical computer science concepts and demonstrates how technology can be used to solve real-world problems.

1.5. Scope of the Project

The scope of this project is limited to students at The University of Hong Kong. The system will be designed to understand the specific needs and requirements of HKU students and provide personalized career guidance accordingly. The project will involve developing the conversation system, recommendation system, and path generation module. The project will also involve conducting user experience surveys and dataset evaluations to assess the performance of the system. The project will not cover job placement or recruitment services.

2. Literature Review

2.1. Current Related Technologies/Studies

Several technologies and studies have made notable strides in areas related to our project. OpenAI's ChatGPT, for instance, is a leading conversational agent that employs machine learning to generate human-like text responses. However, it primarily focuses on answering rather than asking questions, thereby limiting its interactive capability.

On the other hand, recommendation systems like Google Cloud Job Discovery utilize machine learning algorithms to pair job seekers with suitable job listings, based on the individuals' current abilities and objective information. Nevertheless, it fails to take into account the users' future potential and personality traits, thereby possibly restricting their prospects.

When we look at career path construction, the existing solutions predominantly provide piecemeal information from different sources, including posters, leaflets, and emails. These resources, although informative, often overwhelm users with an abundance of paths, and they focus on job titles and hard skills, providing a rather fragmented view of career progression.

2.2. Gap in Existing Solutions

The current solutions exhibit several limitations. Conversational agents, like ChatGPT, might be adept at generating responses but are not designed to initiate context-specific questions. This shortfall hampers their ability to comprehend and capture users' career aspirations and interests fully.

Similarly, existing recommendation systems, such as Google Cloud Job Discovery, base their recommendations on users' present objective information or capabilities, thereby neglecting potential capabilities and personality traits. This approach might prevent users from exploring career paths that they could excel in with further training or development.

Most critically, the information on career path construction that is currently available is fragmented and scattered across various media, making it challenging for users to collate and compare. These resources rarely mention additional experiences that could enhance a candidate's competence, such as internships or research experience, rendering them incomplete.

2.3. Justification for the Project

The identified gaps in the existing technologies and studies underscore the need for our proposed solution, the Personalized Career Consulting System (PCCS). The PCCS amalgamates a conversational system, a recommendation system, and a path generation system to deliver tailor-made career guidance for university students.

Our system surpasses existing conversational agents by designing it to ask meaningful, context-specific questions that capture the user's career interests, abilities, resources, and aspirations. The gathered information forms the foundation of our recommendation system, which considers not only current capabilities but also future possibilities and personality traits.

Most importantly, our path generation system offers a more comprehensive and intuitive view of career progression, incorporating both hard and soft skills. It presents a manageable number of career paths and includes information on additional experiences, such as internships or research experience, that can enhance a candidate's competence. This comprehensive approach enables users to make well-informed decisions about their career path.

In conclusion, the PCCS addresses the limitations of existing solutions by offering a thorough and personalized approach to career guidance for university students.

3. Methodology

The proposed system initiates with a dialogue between the individual and an AI agent to capture the individual's career preferences. Based on these insights, a recommendation system proposes potential careers. Following the initial dialogue and career suggestions, the necessary skills for the chosen careers are extracted and distinct paths for each job are formulated. These paths include required certifications and key activities. The final paths are then presented to the user in a visually engaging format.

As depicted in Fig. 1, the proposed model comprises four distinct sub-systems. This chapter will elaborate on these sub-systems in terms of their architecture, benchmark, input data, implementation steps, assumptions, and limitations.

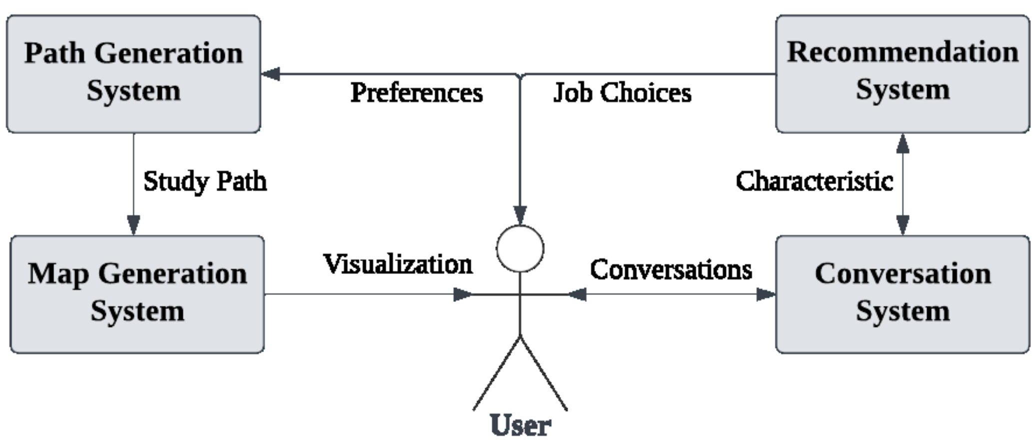

Fig. 1. Overall structure of the Proposed System

3.1. Conversation system

The conversation system serves as the primary data collection module of the model. It facilitates communication with users based on designed question banks, allowing the system to progressively understand and collect user characteristics. The collected and formatted data is then passed on to the subsequent module.

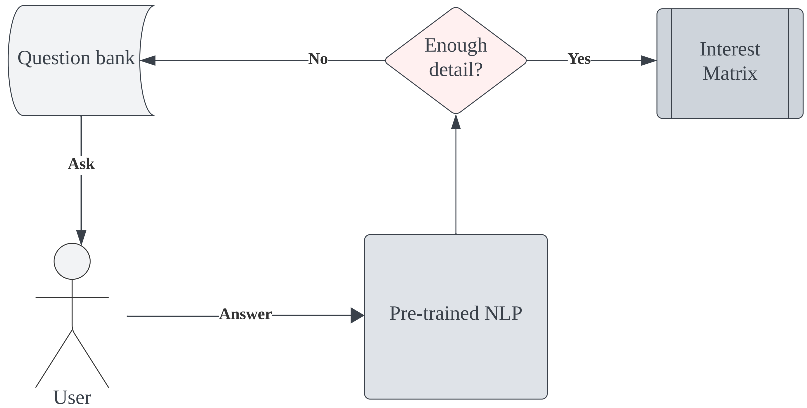

Fig. 2. Architecture of Conversation System

3.1.1. Architecture

The system comprises a question bank and a Natural Language Processing (NLP) model, see figure 2. The question bank contains meticulously crafted questions like MBTI [2]. They are categorized questions based on the nature of relevant jobs and some unbiased questions. A hierarchical structure exists among questions within a category, enabling the system to pose subsequent questions that delve deeper into understanding the user. Unbiased questions help the system define the area of the user's interest; for example, what do you think about yourself?

As for the NLP model, it will be trained before use. It processes the user's responses as input and generates a matrix representing their interests. This provisional output aids the system in continuing with the questions from the bank until the matrix is completed.

3.1.2. Algorithm

The core of the NLP model is an interest matrix:

$$\begin{bmatrix} c_0q_1 & c_0q_2 & \cdots & c_0q_n \\ c_1q_1 & c_1q_2 & \cdots & c_1q_n \\ \vdots & \vdots & \ddots & \vdots \\ c_mq_1 & c_mq_2 & \cdots & c_mq_n \end{bmatrix}$$

Where c_i represents cluster i, and q_j represents question j within cluster i (except for c_0, where c_0 represents those unbiased questions). The c_iq_j is a number between 0 and 1, indicating the user's affinity towards the question. A well-formulated matrix should reflect the user's preferences, with high consistency in certain rows. The matrix is like a composed sentimental analysis.

The algorithm is as follows:

Algorithm 1 NLP for filling interest matrix

```
1: InterestMatrix ← [0]_{(m+1)×n}
2: c, q ← 0, 1 //NextQuestion ← Cluster0_Question1
3: while There is a column of interest matrix is all-zero, i.e., q < n do
4:     InterestMatrix[:, q+1] ← NLP(response to QuestionBank[c, q])
5:     q ← q + 1
6:     c ← argmax_c CumulativeInterest(InterestMatrix[:, q])
7: end while
```

The cumulative interest of one row c_i in column $x+1$ is defined as:

$$CI(x+1; c_i) = \sum_{j=1}^{x} (\alpha_{ij} + \beta_{ij} \gamma^j I_{c_i q_j}) \quad (1)$$

where α_{ij} and β_{ij} are constants to be trained, and γ is a discount factor in reducing the influence of further questions, which is common in reinforcement learning [3].

3.1.3. Benchmark

Cronbach's alpha (see equation 2) can be computed for each row c_i (representing the user's responses to a set of questions q_i). A high Cronbach's alpha (close to 1) would indicate that the questions within a cluster are eliciting consistent responses, suggesting that the system understands the user's preferences well [4].

$$\alpha_{c_i} = \frac{m}{m-1} \left(1 - \frac{\sum_{j=1}^{n} \text{Variance}_{q_j}}{\text{Total Variance}}\right) \quad (2)$$

3.1.4. Implementation Steps

3.1.4.a. Build the question bank

1) Collect questions like MBTI [2].
2) Manually categorize the questions and sort them in order according to relevance.

3.1.4.b. Train/Run

1) User Interaction: Begin a conversation with the user, asking questions from the question bank.
2) Response Collection: Collect and store the user's responses.
3) NLP Processing.
4) Propose questions of the most interested area.
5) Iteration: Continue the conversation with the user, updating the matrix as more data is collected.

3.1.5. Assumptions

The primary assumptions for this system are:

- The user responses are truthful and accurately reflect their interests and preferences.
- The NLP model can accurately interpret and quantify the user's responses.
- The questions in the question bank are well-designed and can effectively elicit user preferences related to different job clusters.
- The user will not end the conversation too early.

3.1.6. Limitations

The main limitations of the system are:

- Dependence on User Input: The accuracy of the matrix relies heavily on the user's willingness to provide truthful and comprehensive responses.
- NLP Interpretation: While NLP has made significant strides, it's not perfect. There may be nuances or sentiments in the user's responses that the NLP model fails to capture.
- Question Bank Coverage: The question bank may not cover all possible job clusters or may not include questions that can accurately determine a user's interest in specific areas. It's also possible that the user's interests do not align neatly with the predefined clusters.
- Cronbach's Alpha Limitations: While Cronbach's alpha is a useful measure of internal consistency, it assumes that all variables are equally reliable, which may not always be the case. Also, a high Cronbach's alpha doesn't guarantee that the matrix accurately reflects the user's

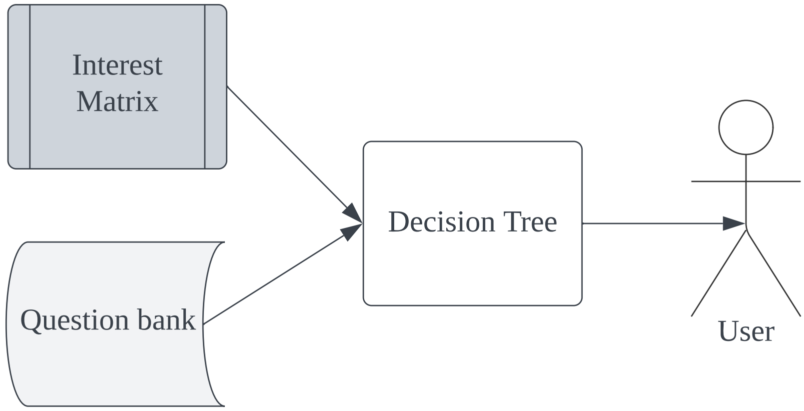

Fig. 3. Architecture of Recommendation System

preferences; it only indicates that the responses to the questions within each cluster are consistent [5].

3.2. Recommendation System

The recommendation system follows the conversation system, using the interest matrix and prior knowledge of the question bank to recommend jobs to users.

3.2.1. Architecture

The system is composed of an input layer and a decision tree. The input layer passes the user's interest on different questions of the bank to the decision tree. The decision tree will generate a list of jobs for the user.

The decision tree model is used because it is easy to understand and interpret. The decision tree mimics the human decision-making process and can handle both categorical and numerical data. The tree is constructed based on the question bank categories and job types, with each question acting as a decision node, leading to different job recommendations based on the user's interests.

3.2.2. Benchmark

This system will be evaluated on the satisfactory rate of users given recommended jobs. In addition to the satisfaction rate of users, the accuracy of the system can also be evaluated by how many users accept the job recommendations. A high acceptance rate would indicate that the system is accurately capturing user preferences and providing relevant job suggestions. User feedback can be collected to further evaluate the performance of the system and refine it.

3.2.3. Implementation Steps

The input is the interest matrix, while the question bank is used to construct the decision tree. After the decision tree, the system will output a list of jobs.

3.2.3.a. Construct the Decision tree

1) Construct the tree based on the categorized question banks manually.
2) Add the job types as the tree leaves.

3.2.3.b. Train/Run

1) Set/Update the threshold of each question's interest.
2) Pass in the interest matrix.

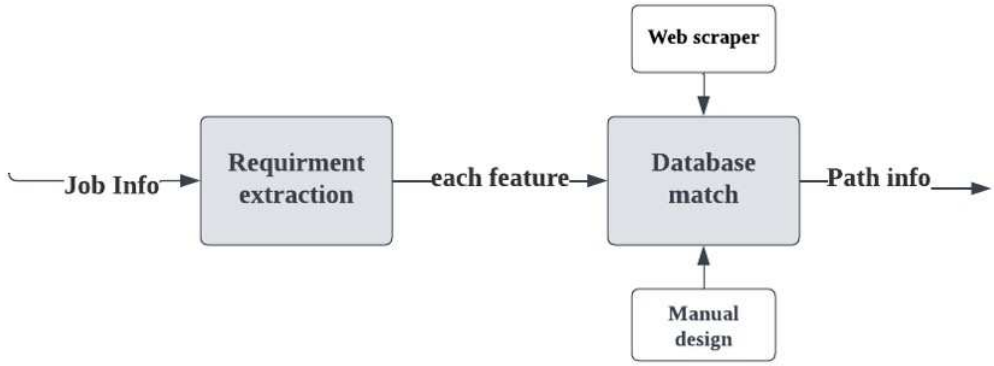

Fig. 4. Architecture of Path Generation System

3.2.3.c. Propose the job list to users

3.2.4. Assumptions

The recommendation system makes a few assumptions:

- The user's responses to questions accurately reflect their interests.
- The interest matrix accurately captures the user's preferences.
- The job types associated with the questions in the question bank are an accurate representation of the job market.

3.2.5. Limitations

The recommendation system also has a few limitations:

- The system assumes that user interests are static and do not change over time. However, in reality, user interests can evolve, and the system may need to adapt accordingly.
- The decision tree model assumes that the questions in the question bank and their associated job types cover all possibilities. However, there may be job types that are not captured in the system, leading to incomplete recommendations.
- Finally, the system's performance largely depends on the quality of the question bank and the accurate categorization of these questions.

3.3. Path Generation System

The Path Generation System aims to provide users with a feasible development plan for their desired job positions. It uses NLP techniques to extract job requirements such as skills and education from the job descriptions. The system then links these requirements to our database to determine the necessary steps for the user to meet them. This could include obtaining certifications, completing relevant coursework, or gaining specific experience. The system's goal is to help users create a practical plan to qualify for their target jobs.

3.3.1. Architecture

The architecture of the Path Generation System is illustrated in the Fig. 3.3 above and consists of several key components:

- Extraction Component: This component uses Natural Language Processing (NLP) techniques to extract job requirements from the job descriptions. Tools like spaCy and MITIE can be used for this purpose [6], [7]
- Database Component: The extracted job requirements are matched with our existing database

of job skills and qualifications.
- Path Information Input Component: This component allows us to manually input path information into our system, such as details about specific courses, certifications, or experiences required for a particular job skill.
- Web Scraping Component: This component uses web scraping techniques to automatically gather information related to the job requirements. Tools like Scrapy or Octoparse can be used to gather information from online resources [8], [9]. Methods such as manual verification will be adopted to select information for entry into the database.

3.3.2. Benchmark

The benchmarking process for the Path Generation System evaluates its accuracy, efficiency, and usability. Accuracy is assessed by comparing the system's output with paths suggested by professionals in the relevant fields. Also, the Skill2Vec dataset which provides a large collection of standardized job descriptions with required skills can be used as a benchmark for evaluating the performance of our system [10]. Efficiency is gauged by the time taken to generate a development path. Usability is determined through user surveys, focusing on aspects such as ease of use and overall user satisfaction.

3.3.3. Implementation Steps

The system takes job descriptions from recruitment websites as input and generates a to-do list for the user with suggested timeline and importance rates.
The implementation process of the Path Generation System involves:

3.3.3.a. Framework build-up

1) Setting up the necessary tools and frameworks, including NLP tools for data extraction and a database for storing job skills and qualifications.

3.3.3.b. Train/Run

1) Feature Extraction: Using NLP to extract job requirements from the job descriptions.
2) Data Matching: Matching the extracted requirements with our database.
3) Path Generation: Creating a development path for each requirement.

3.3.3.c. Update data in database

3.3.4. Assumptions

The Path Generation System operates under several assumptions:

- The job descriptions provided by the user are accurate and comprehensive.
- The system's database contains up-to-date and relevant information about job skills and qualifications.
- The run-time of the system for each job should be controlled within a few seconds to tens of seconds.

3.3.5. Limitations

The Path Generation System has certain limitations:

- Validity of Online Information: The system relies on online information, which may not always be reliable or accurate.
- Ranking Importance: The system does not rank the importance of different job requirements.
- Dynamic Nature of Job Market: The system may not always reflect the most current trends or requirements in the job market.

These will be factors to be considered in future implementation.

3.4. Map Generation System

The Map Generation System, a sub system under path generation system, visualizes the career path suggestions from the Path Generation System in a tree map. It features a main path for general skills and branches for specialized roles, diverging at specific points. The map can be exported in various formats like JPEG, PDF, or HTML for user convenience.

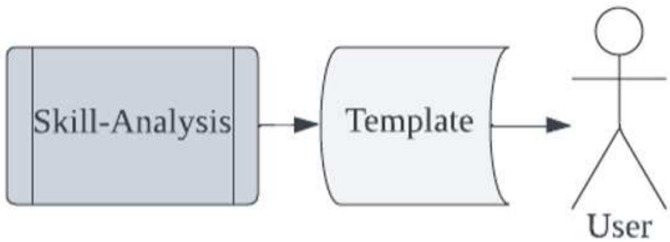

Fig. 5. Architecture of Map Generation System

3.4.1. Architecture

The architecture of the Map Generation System comprises two main parts: Skills Analysis and the Template System.

- Skills Analysis identifies the main skills required for the chosen career path and any specialized skills needed for specific roles within that path.
- The Template System uses this information to generate a visual map, creating a main path for general skills and branches for specialized skills, with divergence points indicating when to acquire each skill.

3.4.2. Benchmark

The benchmark of the Map Generation System is evaluated mainly through user feedback, assessing how effectively the system visualizes their career path.

3.4.3. Implementations Steps

The input data are suggested events generated by the Path Generation System, and the output is a visualized data map of these events.
The implementation involves:

3.4.3.a. Train/Run

1) Input Processing
2) Main Path Generation: The events are sorted using a formula like:

$$\text{Score} = \mathbf{w_1} \times \text{Frequency} + \mathbf{w_2} \times \text{Importance Rate} \tag{3}$$

3.4.3.b. Create a template for visualization

3.4.3.c. Deliver the result to the user

3.4.4. Assumptions

The system assumes that:

- It can accurately distinguish between mainline and branch information.
- It can provide a visual result that satisfies the user’s needs.

3.4.5. Limitations

The system does present a couple of limitations:

- The system may not have enough information to accurately define the mainline content, particularly in less common or emerging fields. This is largely due to the need for extensive manually marked importance data.
- For users with a wide range of interests, the single mainline approach may not fully cater to their needs as it could oversimplify their diverse career paths. More sophisticated methods are required to address this issue.

4. Evaluation

Considering that our entire agent system operates on the basis of human-computer interaction, the user experience is a crucial metric for evaluating our system, including their propensity to engage with our system. As outlined in Section 3.2.2 and Section 2.4.2, a user experience survey can be implemented to assess our system.

4.1. Survey Procedure

Our proposed survey will encompass two components: system trial and questionnaire completion. However, specific details of the survey may be subject to modification in line with the progression of our system development. Any alterations and more comprehensive information will be detailed in subsequent reports.

4.1.1. System Trial

In the initial segment, users will be invited to trial our system by accomplishing a complete flow, which includes interacting with the agent, reviewing the generated paths and nodes, selecting preferred paths, and engaging with the final map.

4.1.2. Questionnaire Completion

In the latter segment, users will be required to fill out a questionnaire. Our proposed question list for the questionnaire is as follows:

1) The questions posed by the agent effectively elicited your requirements.
2) The agent accurately interpreted your responses.
3) The paths generated (prior to your selection) included career plans that piqued your interest.
4) The map generated (post your selection) clearly delineated how you can achieve your career goals during your university tenure.
5) If this system were made available online, you would be highly likely to use it.

Each question is formatted as an agree-disagree statement, and users should score each statement on a scale of 1 (strongly disagree) to 5 (strongly agree).

Among the five questions above, question 1 and question 2 evaluate the quality of the question bank as well as the natural language model (Section 3.1). Question 3 assesses our recommendation system (Section 3.2). Question 4 evaluates the quality of nodes and paths (Section 3.3 and Section 3.4). Question 5 gauges the users' valuation of our system.

4.2. Survey Improvement

4.2.1. Incentive Promise

Despite the survey-based evaluation potentially limiting the quality of the assessment of our system, as users' behaviors are less predictable, we plan to provide incentives to users who complete the survey, such as a Starbucks gift card worth twenty Hong Kong dollars. This approach may result in a higher response rate and improved survey quality [11].

4.2.2. In-person Survey

In addition to the provision of incentives, the quality of the survey can be further enhanced by inviting students on the University of Hong Kong campus for in-person surveys. This approach is deemed effective based on the findings of [12], which suggest that personal interviews are prioritized over mail surveys.

4.2.3. Five-category Response

Another limitation of this approach may be the difficulty in quantifying user experience and satisfaction. To address this issue, adjustments will be made. As mentioned in Section 4.1.2, questions will be set in a scoring format from 1 to 5 instead of other numbers of categories or answering in natural language. This approach minimizes bias caused by differences in expressive capabilities and language customs among individuals. The design is inspired by the research conducted by [13], which argues that for agree-disagree scales, five answer categories yield higher quality data than seven and eleven.

4.3. Assumptions

In light of the evaluation strategy and improvements made, we make the following assumptions to facilitate the user experience survey for our proposal:

1) Users intend to provide authentic feedback, rather than completing the survey carelessly or providing dishonest responses.
2) Users can fully comprehend the objectives of our system through trials during the survey.

4.4. Limitations

Despite the improvements made, including promised incentives, in-person surveys, and five-category responses, there remains one limitation that is challenging to address. As a user-based survey, the evaluation results are inevitably influenced by the variance of values and preferences among individuals, which increases the subjectivity of the evaluation.

5. Project Timeline

The proposed timeline will be:

Phase 1: Project Planning and Design (October 2023)

- Week 1: Research on job recommendation systems and natural language processing models
- Week 2: Finalization of detailed project plan and strategy
- Week 3: Design of detailed system architecture and components
- Week 4: Collection of Necessary Data

Phase 2: System Development and Implementation (November 2023 - Feburary 2024)

- Week 1-4: Development of the conversation system
- Week 5-8: Development of the recommendation system
- Week 9-12: Development of the path generation system
- Week 13 - 16: Development of web interface

Phase 3: Testing, Evaluation and Refinement (March 2024)

- Week 1: Integration testing of all system components
- Week 2: Conducting user experience surveys and system evaluation
- Week 3: Refinement of system based on user feedback and evaluation results
- Week 4: Final system testing and preparation for project exhibition

Phase 4: Documentation and Project Exhibition (April 2024)

- Week 1: Completion of final report and system documentation

- Week 2: Preparation for project exhibition
- Week 3: Project exhibition

6. Project Management

6.1. Roles and Responsibilities

- **Henry Hung**: Hung will design the overall system architecture and develop the front-end of the system. He will integrate the Natural Language Processing model into the conversation system and develop question asking function for the elicitation of user profile.
- **Jin Shang**: Jin will collect and process data, and develop the data models required for the recommendation system. He will manage the data flow between different subsystems, and handle data pre-processing tasks such as cleaning and normalization to ensure it is ready for use in the system.
- **Xu Haozhou**: Xu will develop the recommendation system by applying machine learning algorithms. He will fine-tune the machine learning models, analyze the data outputs, and develop the Path Generation System, which involves creating algorithms to generate career paths.
- **Zhao Yiming**: Zhao will implement the reinforcement learning algorithms for the recommendation system, and design and develop the data extraction and interpretation modules for conversation system.

Each team member will participate in regular team meetings, contribute to problem-solving discussions, assist in testing and debugging the system, and collaborate on the preparation and presentation of project reports and documentation.

6.2. Communication Plan

The success of our project heavily relies on effective communication among team members. We have established the following communication plan to ensure smooth and efficient collaboration:

- **Regular Meetings:** We will hold weekly meetings to discuss the project's progress, address any issues or challenges, and plan for the upcoming week. These meetings will also serve as a platform for brainstorming and exchanging ideas.
- **Instant Messaging:** We will use an instant messaging platform for day-to-day communication. This will allow for quick exchanges of information, problem-solving discussions, and timely updates.
- **Collaboration Tools:** We will use tools like Google Docs and GitHub for real-time collaboration on documents and code respectively. These platforms allow us to work together simultaneously, track changes, and maintain version control.
- **Individual Responsibilities:** Each team member will be responsible for clearly communicating their progress, challenges, and needs. If a team member is unable to meet a deadline or complete a task, they should communicate this to the team as early as possible.
- **Conflict Resolution:** In the event of a disagreement or conflict, we will endeavor to resolve it through open discussion and compromise. If necessary, we may resort to a majority vote.

6.3. Quality Assurance Strategy

- **Documentation:** We will maintain thorough documentation of our code, system architecture, and key decisions. This will facilitate understanding, troubleshooting, and future modifications.

7. Expected Challenges and Solutions

7.1. Competing with Established Systems

One of the primary challenges our project is expected to face is the competition from established job recommendation systems. Notably, Google has developed a sophisticated job recommendation system that our system needs to outperform. To address this, our system has been designed to consider not just objective parameters such as skills, location, and language, but also subjective factors and future possibilities. For instance, if a user is a student with four years left in their degree, our system considers the skills and qualifications they could potentially acquire over this period. We also factor in personal preferences and aspirations, which are crucial to job satisfaction but often overlooked in traditional job recommendation systems. This multidimensional approach provides a more comprehensive and personalized career recommendation, distinguishing our system from existing solutions.

7.2. Ensuring Validity of Online Information

Another challenge is the validity of the job descriptions and prerequisites obtained online. The accuracy of this information is critical to the effectiveness of our path generation system. To mitigate this, we plan to source our data primarily from the recruitment websites of large, reputable companies. These sources tend to provide more professional and accurate job descriptions and expectations. Moreover, we will implement data verification measures, such as cross-referencing job descriptions from multiple sources, to ensure the accuracy of our data.

7.3. Ranking Career Importance

Determining the importance of a career to a specific person is a complex task that our system must accomplish. To do this, we consider a variety of factors, including both objective parameters (such as skill match and career prospects) and subjective factors (personal interest and satisfaction). By balancing these diverse factors, our system can provide a personalized ranking that truly reflects the user's career preferences and goals.

7.4. Preventing Interaction Divergence

Finally, our system must prevent interaction divergence, where the agent continues to ask questions without converging on a career recommendation. To address this, we have proposed a structured conversation system based on a meticulously designed question bank and an interest matrix. The interest matrix ensures that the questions posed are progressively more targeted, based on the user's responses, allowing the conversation to converge on a set of suitable career recommendations. This approach ensures a focused and efficient interaction, leading to timely and relevant career guidance for the user.

8. Conclusions

8.1. Summary of the Project Plan

The primary goal of this project is to develop a Personalized Career Consulting System (PCCS) that aids university students in navigating their career paths. The PCCS consists of three main components: a conversation system, a recommendation system, and a path generation system. The conversation system interacts with the student to gather information about their interests, abilities, resources, and career goals. The recommendation system then uses this data to suggest potential career paths. Finally, the path generation system creates a visual timeline detailing the steps required to achieve these career paths. The project will be evaluated through user experience surveys and accuracy assessments of the recommendation system.

8.2. Expected Impact

The PCCS is expected to significantly improve the career planning process for university students. By providing personalized guidance early in their university life, students will be better equipped to make informed decisions about their future, reducing uncertainty, and increasing career satisfaction. The system also aims to help students identify and seize opportunities that align with their career goals, reducing the chance of missed opportunities and regrets. Furthermore, by visualizing the career path, students can gain a clear overview of their career progression, helping them to stay motivated and focused. In conclusion, the PCCS is expected to have a significant positive impact on the career planning and decision-making process for university students.

Acknowledgements

The authors would like to express their deepest gratitude to Professor Kao, Benjamin C.M., for his invaluable input and guidance throughout the planning stages of this project. His expertise and insights greatly contributed to the development of this project plan.

References

[1] T. Vadera, "Students are confused about their careers: A survey," https://www.edtechreview.in/research/student-career-jobs-employability-survey/, 2017, accessed: 30-Sep-2023.
[2] I. Myers and M. McCaulley, *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Consulting Psychologists Press, 1985.
[3] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 2018.
[4] M. Tavakol and R. Dennick, "Making sense of cronbach's alpha," *International journal of medical education*, vol. 2, p. 53, 2011.
[5] K. Sijtsma, "On the use, the misuse, and the very limited usefulness of cronbach's alpha," *Psychometrika*, vol. 74, no. 1, p. 107, 2009.
[6] "spacy: Industrial-strength natural language processing in python," https://spacy.io/, accessed: 24-Sep-2023.
[7] "Mitie: Mit information extraction," https://github.com/mit-nlp/MITIE, accessed: 24-Sep-2023.
[8] "Scrapy — a fast and powerful scraping and web crawling framework," https://scrapy.org/, accessed: 24-Sep-2023.
[9] "Octoparse — free automatic web scraper, web data extractor," https://www.octoparse.com/, accessed: 24-Sep-2023.
[10] "Github - duyet/skill2vec-dataset: Dataset and pre-trained model for . . . ," https://github.com/duyet/skill2vec-dataset, accessed: 24-Sep-2023.
[11] M. L. Berk, N. A. Mathiowetz, and A. A. White, "The effect of prepaid and promised incentives: Results of a controlled experiment," *Journal of Official Statistics*, vol. 3, no. 4, p. 449, 1987. [Online]. Available: http://eproxy.lib.hku.hk/login?url=https://www.proquest.com/scholarly-journals/effect-prepaid-promised-incentives-results/docview/1266807435/se-2
[12] J. Yu and H. Cooper, "A quantitative review of research design effects on response rates to questionnaires," *Journal of Marketing Research*, vol. 20, no. 1, pp. 36–44, Feb 1983.
[13] M. A. Revilla, W. E. Saris, and J. A. Krosnick, "Choosing the number of categories in agree–disagree scales," *Sociological Methods Research*, vol. 43, no. 1, pp. 73–97, Dec 2013.