# **Progress Report #1**

**Pathfinder: Personalized Career Path Recommender System for University Students**

**Leshan Chathuranga Kuruppuarachchi – 300363077**

**Work Logs**

|  |  |  |
| --- | --- | --- |
| Date | Number of Hours | Description of work done |
| 2025-09-05 | 3 | Initial research on the topic. Refer previous research papers. |
| 2025-09-07 | 4 | Refer previous research papers.  Research for existing web scraping resources |
| 2025-09-08 | 3 | Research for existing web scraping resources |
| 2025-09-09 | 3 | Work on project proposal |
| 2025-09-10 | 2 | Work on project proposal  Set up git repo |
| 2025-09-11 | 3 | Finalizing project proposal |
| 2025-09-15 | 2 | Work on web scraping code |
| 2025-09-16 | 3 | Work on web scraping code |
| 2025-09-20 | 4 | Work on web scraping code |
| 2025-09-23 | 1.5 | Finalizing project proposal with feedback |
| 2025-09-24 | 3 | Work on progress report 1 |

**Work Description**

In the first week our project group was formed and we agreed on a project idea. I began researching existing studies on the topic and reviewed several research papers. I also explored sample web-scraping code to collect data from job sites. In addition, I started drafting the project proposal and set up the Git repository. We held a few team calls to clarify sections of the proposal.

In the second I finalized the project proposal and focused on developing the web-scraping code. I tested several existing scripts, but they failed due to website security measures and frequent CSS changes. I then began investigating methods to bypass these security restrictions.

In the third week I continued improving the web-scraping code and successfully found a way to handle the security restrictions. I am still refining the code to manage CSS changes and capture all required information. Meanwhile, I finalized the proposal resubmission and prepared Progress Report 1. Also, we held a few team calls to discuss the progress of the project.

**Repo Check In**

Created the folder structure and updated the README.md file.

The files/folders I have checked in the repo are as follows:

Documents /

Misc / Research Articles/ (existing studies research papers)

- 293Harsh1.0 (AJCS).pdf

- CareerX.pdf

- JETIR2403A44.pdf

- PCRS\_Personalized\_Career-Path\_Recommender\_System\_for\_Engineering\_Students.pdf

- kainyu-2025-ijca-924808.pdf

Misc / POC Web Scraping / (sample web-scraping code – still refining)

- web\_scraping.ipynb