

# Progress Report #3

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

**Leshan Chathuranga Kuruppuarachchi – 300363077**

## Work Logs

| Date       | Number of Hours | Description of work done  |
|------------|-----------------|---|
| 2025-10-11 | 2               | Set up the development folder structure and started work on web scraping indeed.com |
| 2025-10-13 | 3               | Work on web scraping indeed.com to retrieve jobs                                    |
| 2025-10-15 | 3               | Work on web scraping indeed.com to retrieve jobs and save in database               |
| 2025-10-18 | 3               | Implement initial frontend  |
| 2025-10-21 | 2.5             | Work on midterm report  |
| 2025-10-22 | 2               | Work on progress report 3   |

## Work Description

In the sixth week I created initial folder setup for project implementation and continued work on the web-scraping code, also started to work on saving the scrapped data in database. Initially we proposed saving data on MySQL or H2 database but while doing the implementation I faced difficulties related to libraries related to MySQL and H2 databases, so I looked for solutions and found out that SQLite is a better solution for when it comes to handling databases with python. So far other than job descriptions, all the other information is scraped and saved in database currently finding solutions to track job descriptions.

In the seventh week I started implementation of frontend app using React. Currently I have created UI elements and fed with dummy data. Also, I started working on Midterm Report and Progress Report 3.

## **Repo Check In**

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

Documents / (branch: main)

- Project\_LKu077\_Report3.pdf
- Midterm\_Report\_LKu077.docx

Implementation / backend (branch: leshan)

- indeed\_jobs\_backup.csv
- indeed\_web\_scraping.ipynb
- indeed\_web\_scraping.py
- pathfinder\_db.sqlite
- requirements.txt

Implementation / frontend (branch: leshan) all the files inside the frontend (this is the initial checkin with main UI elements)