**Proposer Details**

| Group Number | *[G-03]* |
| --- | --- |
| Registration Number of Group Members | [2020-CS-110, 2020-CS-104] |

**Proposal Details**

|  |  |
| --- | --- |
| ***Project*** |  |
| Proposed Project Title | *LinkedIn Profile Scrapper* |
| Executive Summary | *LinkedIn is now become one of the most famous platforms which is primarily used for professional networking and career development, and allows job seekers to post their CVs and employers to post jobs. LinkedIn has 774+ million registered members from over 200 countries and territories. LinkedIn allows members (both workers and employers) to create profiles and "connect" with each other in an online social network which may represent real-world professional relationships.*  *This programs “LinkedIn Profile Scrapper” is designed to scrape the details of LinkedIn profile data of different Persons from different countries in the world. The Profile data will consist of their Name, Username, Job designation/ Qualification, Company Name, Number of Connections, Residence, Number of followers, Number of recommendations.*  *This will be a graphical single user desktop program where he/she will be provided with 3 screens. First screen will be available for the user to scrape the data with Start, Stop, Resume, Pause and Save Data with the table of loading data. A loading bar will also be shown indicating the amount of scrapping being done.*  *Second Screen will be shown with all the data scrapped and with the Sorting algorithms available. He/she can choose the algorithm of his choice and can select the column and can sort the column weather in ascending and descending order.*  *Third Screen will be shown will all the scraped data and the filters of the data will also be shown. He/she can choose filter of his/her choice according to the need and can search the persons data from it.* |
| ***Business Case*** |  |
| Outline the business need for the project | *LinkedIn Is the biggest Platform for the companies to find a person of their interest with the best possible skills and experience.*  *From this program, companies can scrape all the existing more than 774+ million profiles and then can filter the data according to their need and can hire or contact particular persons of particular skills and can have communication with them.* |
| End user of the product | *Different Companies Of any type can use this Program to hire their desired faculty/staff.* |
| Motivation for Project | *We just want to provide such a software to the industry which reduces their efforts to find persons of their interest. So that people could be hired sitting at home. In this way, we could serve the society a bit.* |
| State the level of impact expected should the project proceed and implications of not proceeding | *[State whether the implementation would have an impact at an operational level and/or strategic level and state the impact(s) in 2-3 lines]* |
| ***Technical Details*** |  |
| Name of Entity | *Name, Username, Job designation/ Qualification, Company Name, Number of Connections, Residence, Number of followers, Number of recommendations.* |
| Attributes of Entity  (Minimum seven attributes/rows can be increased) | |  |  |  | | --- | --- | --- | | *Name* | *Data Type* | *Description* | | Full Name | String | This will contain the full name of the User. | | Username | String | Every profile has a unique username through which people access their profiles. | | Job Designation | String | Job on which a user is currently working or studying. | | Company Name | String | Company in which user is working right now. | | No. of Connections. | Integers | Total Number of connections he had with other users. | | Address | String | Address of the user from where he is. Such as City, Country. | | No. of Followers | Integer | Number of people who have followed the user to see his achievements, skills and activities. | | No. of Recommendations. | Integer | This will show the total numbers of recommendations the person has from well-known people. | |  |  |  | |
| Sample of Scrapping Source |  |
| GitHub Repository Link | *https://github.com/Nouman0x45/CS261F21PID55* |
| Sorting Algorithms |  |
| |  |  | | --- | --- | | **Algorithm Name** | **Description (Each algorithm in 2-3 lines)** | | Selection Sort | *It sorts an array by repeatedly finding the minimum element from unsorted part and putting it at the beginning.* | | Bubble Sort | *It works by repeatedly swapping the adjacent elements if they are in wrong order.* | | Insertion Sort | *The array is virtually split into a sorted and an unsorted part. Values from the unsorted part are picked and placed at the correct position in the sorted part.* | | Merge Sort | *It divides the input array into two halves, calls itself for the two halves, and then merges the two sorted halves.* | | Quick Sort | *It is a Divide and Conquer algorithm. It picks an element as pivot and partitions the given array around the picked pivot.* | | Counting Sort | *It is based on keys between a specific range. It works by counting the numbers of objects having distinct key values. Then doing some arithmetic to calculate the position of each object in the output sequence.* | | Radix Sort | *It is a linear time sorting algorithm that sort in 0(n+k) time when elements are in the range from 1 to k. It does digit by digit sort starting from least significant digit to most significant digit.* | | Bucket Sort | *It is mainly useful when input is uniformly distributed over a range.* | | Cycle Sort | It is comparison based unstable sorting algorithms. Array to be sorted can be divided into cycle | | Cocktail Sort | It transverse an element to its right position by moving in both direction back and forth. | | Shell Sort | Shell Sort involves sorting elements which are away from each other. We sort a large sub-list of a given list and go on reducing the size of the list until all elements are sorted. | | |
| Searching Algorithms | *Searching algorithms are designed to check for an element or retrieve an element from any data structure where it is stored. There are two categories of the algorithm:*   1. *Sequence Search: In this, the list or array is traversed sequentially, and every element is checked.*. 2. *Interval Search: These are specifically designed for searching in sorted data-structure.* 3. *Binary Search: Searches a sorted list on the basis of comparison..* |
| Searching Filters for each data type | *1: A search bar will be Available if he/she will start simple*  *writing name than it will be consider as Name and matching*  *result will be shown from “****Name column****”.*  *2: A search bar will be Available if he/she will start writing with*  *“@” symbol then it will be considered as Username and*  *matching result will be shown from “****Username****” column.*  *3: A search bar will be Available if he/she will start writing with*  *“#” symbol then it will be considered as Company name and*  *matching result will be shown from “****Company Name****” column.*  *4: A dropdown list will be available with* ***Countries*** *name present*  *in scrapped data. He/she can choose the countries.*  *5: A slider with integers number will be available for “****Number of***  ***connections****” and result will contain profiles with the given*  *range of connections.*  *6: A slider with integers number will be available for “****Number of***  ***Followers****” and result will contain profiles with the given*  *range of Followers.* |
| Multi-Level Sorting | Multi-Level sorting will be done in a way, if the user select the column “No. of Connections” and we will sort the data based on selected column with given algorithm and then the user selects another column “Name” then the previous data will stay sorted according to the previous selected column, then within the sorted data, we will again sort the data according to given column with the range of previously sorted data. |
| Any other features | Before Starting the scraping, user can also enter the country name. Then the data scraped will be from that countries Profile. |
| ***Interfaces for your project*** |  |
| ***UI # 01:***     |  |  |  | | --- | --- | --- | | UI Component Name | Type of UI component | Purpose of UI Component/Other details | | Logo | Pic | Its just a logo of Program will be in top left corner. | | Navigation Bar | Buttons | It is the navigation bar which will direct pages of the program from one to another. | | Scraping bar | Loading bar | This bar will indicate that how much profiles have been scrapped. | | Profile’s table | Table | This table will show the scrapped profiles from the LinkedIn. | | Select Country | Drop-Down menu | This will contain all the countries | | Profiles Amount | Text | This will ask user to enter the amount of profiles data to scrape. | | 4 Buttons | Button | These buttons will stop, start resume or load the scraped data into the PC. |   ***UI # 02:***     |  |  |  | | --- | --- | --- | | UI Component Name | Type of UI component | Purpose of UI Component/Other details | | Logo | Pic | It’s just a logo of Program will be in top left corner. | | Navigation Bar | Buttons | It is the navigation bar which will direct pages of the program from one to another. | | Profile’s table | Table | This table will show the scrapped profiles from the LinkedIn. | | Select Algorithm | Drop-Down | User will select available algorithm from here. | | Select Entity | Drop-Down | User will select which entity he want to sort accordingly. | | Select Order | Drop-Down | User will select how to sort it e.g ascending or descending. |   ***UI # 03:***     |  |  |  | | --- | --- | --- | | UI Component Name | Type of UI component | Purpose of UI Component/Other details | | Logo | Pic | Its just a logo of Program will be in top left corner. | | Navigation Bar | Buttons | It is the navigation bar which will direct pages of the program from one to another. | | Profile’s table | Table | This table will show the scrapped profiles from the LinkedIn. | | Type Name or Username | Text | This will take name or username or Company name and will search according to it. | | Select Country | Drop-Down | Countries in scrapped data will show here and user can select it here. | | 3 Amount Selectors | Scale | This will be used by user to select the amount of entity to search. | |  |  |  | | |