## **DATA STRUCTURES**

## **PROJECT REPORT**

**BS-SE FALL 2021**

**TEAM MEMBERS:**

* Amna Mubarak (20k-1695)
* Dehya khuraim (20k-0128)
* Tayyab Shahzad (20k-1043)



**Department of Computer Science**

**FAST-National University of Computer & Emerging Sciences, Karachi.**

**E-RECRUITMENT**

**OVERVIEW:**

The recruitment system will help different companies to find potential candidates for the specific position. The systems will allow businesses to collect information, abstract the relevant applications based on experience and skill set.

**PROJECT DETAILS:**

**1- How it works:**

The system will abstract the provided information of the candidates according to the positions. All resumes of the same domain workers will be sorted, then in the next step the system will filter out the relevant applications. The system will also allow the user to search for the desired applications.

**2-Project Goals:**

The goals of the system are :

1- Recruit and place top-tier applicants.

2- To find more qualified applicants with appealing skill sets

3- To Abstracts the relevant and irrelevant applications

4- To screen applicants

5- Allow the user to search for the desired domain worker

6- To shortlist the potential candidates.

7- Allow the user to add new data to the dataset if there is any need

**3- Benefits:**

The E-Recruitment system will help the user to work more efficiently. The user can handle more resumes at a time. It will reduce the cost of time, as well as the cost of employees required to maintain recruitment and selection procedures. Our system will help the user to search for the candidate through his skills which are mapped on the scale of 1 to 10. The system will also allow the user to search the applicants through the year of experience he has and through the desired domain.

**4- DATASET:**

The dataset which we have chosen for this project is approximately 950 mbs consists of information related to the applicants. Basically it consists of the Applicant’s Name, Domain, Experience, Applicant’s ID and the skills. This is basically the pool of applicants,In which we have extracted the applicant’s name, his experience and the work domain,then applied algorithms of sorting and searching on it for further processes.

**5- DATA STRUCTURES:**

* **AVL:**

We have used AVL for saving all the data, because it is a self balancing tree and the main purpose of our project is searching so the tree should be perfectly balanced. AVL helps us to perform all the operations at O(logn) which is pretty faster as compared to other data structures.

* **Why not BST** : If we use BST and insert the values in decreasing or increasing order, BST will be one-sided and the search time will be O(n), but if we use AVL, the tree will self-balance using rotations.
* **Why not Linked List:** If we use Linked List and insert the values in decreasing or increasing order, the Linked List will be skewed right or skewed left and the search time will be O(n), but if we use AVL, the tree will balance itself by using rotations.
* **QUEUES:**

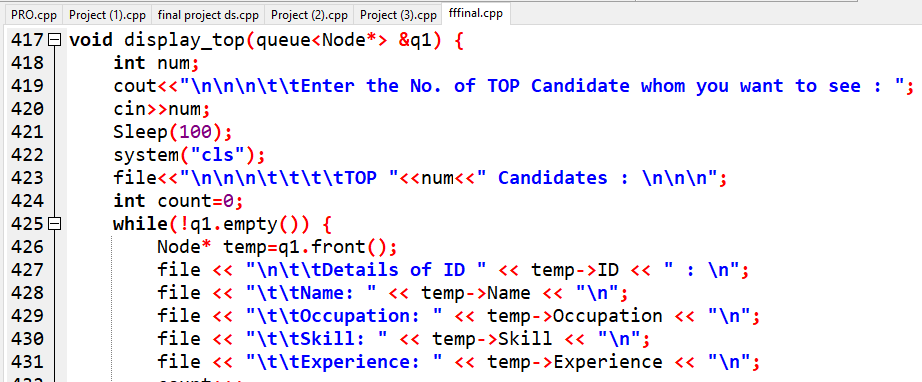
Queues are also used in the project for BFS purpose

**6-Preprocessing Data**

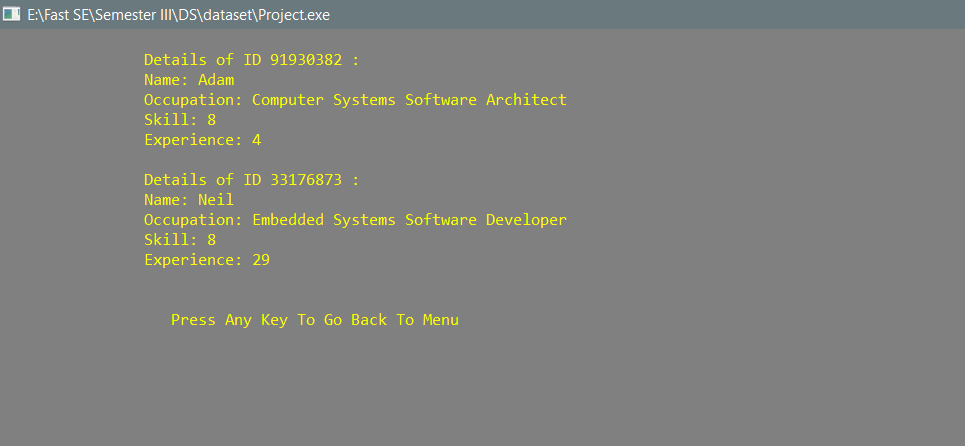
The process of converting raw data into a comprehensible format is known as Preprocessing data. We can't work with raw data, thus this is the key stage when you are working with big data.We have abstracted the relevant data of the applicants from the dataset by manually preprocessing it. The Applicant’s id is used as the primary key, through which we have checked for the duplicate applicants and then eliminated these applications as well. We have erased all the white spaces between the data and also removed the dummy values.

**7-Idea Completion/Goals Achieved:**

**1- Recruit and place top-tier applicants.**



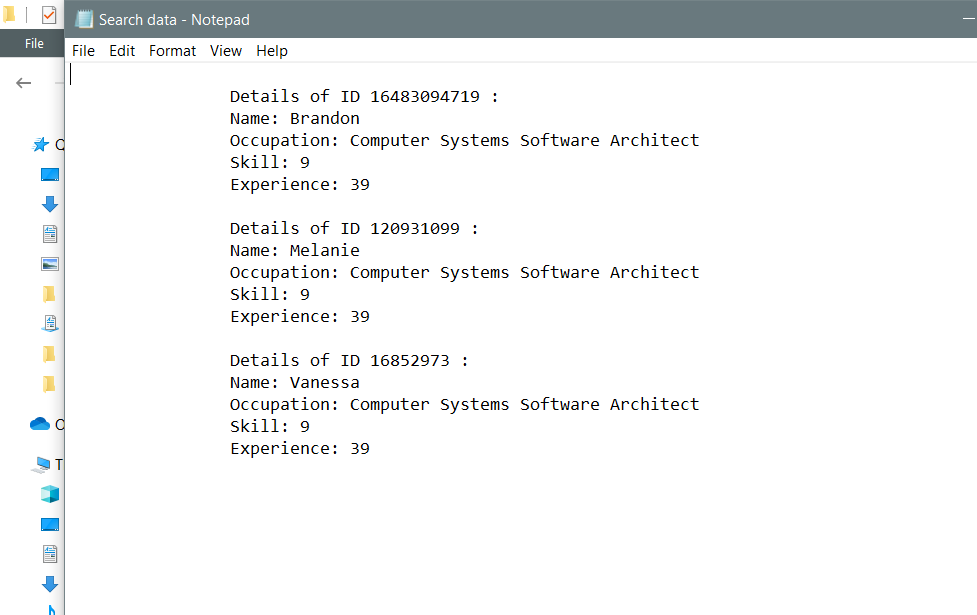
**2- To find more qualified applicants with appealing skill sets**



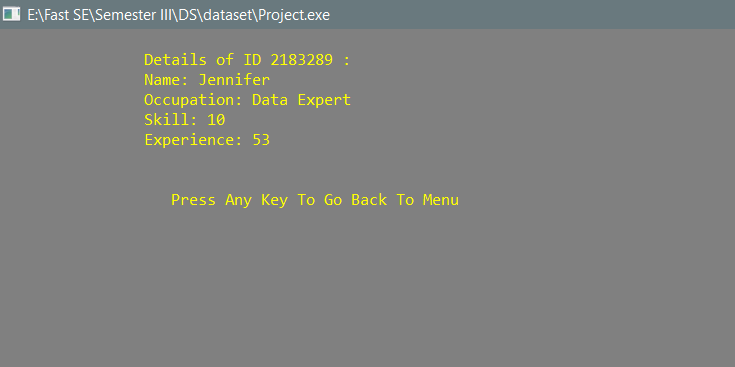
**3- To Abstracts the relevant and irrelevant applications**

[Preprocessing Data](https://docs.google.com/document/d/117OCRuMPxr-WZNjqXEaH39_kKSlv1xXFCfGCFgK-MZk/edit#bookmark=id.70oa9zcos2nx)

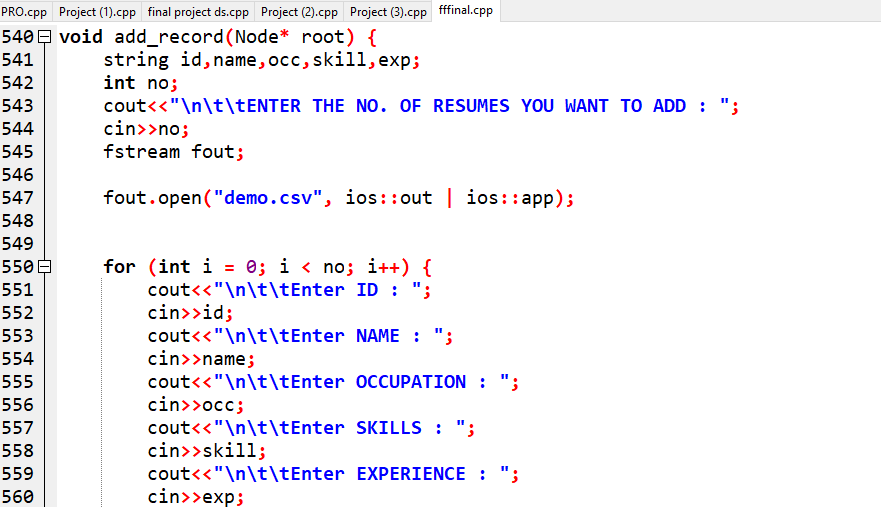
**4- Allow the user to search for the desired domain worker**



**5- To shortlist the potential candidates.**



**6- Allow the user to add new data to the dataset if there is any need**

****

## 