 CS261-Data Structure and Algorithms

Final Term Project Report (Fall 2021)

Laptop Scraper



Project Supervisor

Mr. Samyan Qayyum Wahla

# Proposer Details:

**Group Number:**

G-06

**Registration Number of Group Members:**

Member 1: Member 2:

2020-CS-129 2020-CS-111

Name: Noor Atif Name: Aqsa Riaz

Department of Computer Science, University of Engineering and Technology, Lahore Pakistan

Table of Contents:

[Proposer Details: 1](#_Toc87026180)

[Proposal Details: 3](#_Toc87026181)

[Project Description: 3](#_Toc87026182)

[Business Case: 3](#_Toc87026183)

[Motivation for Project: 4](#_Toc87026184)

[Benefits: 4](#_Toc87026185)

[Technology Used: 5](#_Toc87026186)

[Technical Details: 5](#_Toc87026187)

[Final GUI Display: 11](#_Toc87026188)

[GUI Description: 11](#_Toc87026189)

[Detailed Description Of Algorithms: 12](#_Toc87026190)

[Selection Sort: 12](#_Toc87026191)

[Merge Sort: 15](#_Toc87026192)

[Insertion Sort: 19](#_Toc87026193)

[Quick Sort: 21](#_Toc87026194)

[Bubble Sort: 25](#_Toc87026195)

[Cube Sort: 28](#_Toc87026196)

[Heap Sort: 29](#_Toc87026197)

[Counting sort: 32](#_Toc87026198)

[Bucket Sort: 35](#_Toc87026199)

[Radix Sort: 38](#_Toc87026200)

[Shell Sort: 42](#_Toc87026201)

[Hybrid Sort: 45](#_Toc87026202)

[Collaboration: 48](#_Toc87026203)

[Task Division: 48](#_Toc87026204)

[Issues Faced: 49](#_Toc87026205)

[Implementation: 50](#_Toc87026206)

[Project Plan 50](#_Toc87026207)

# Proposal Details:

**Proposed Project Title:**

Laptop Scraper

# Project Description:

The Title of our project is Laptop-Scraper. No matter how you use computers, the overarching reason to own a laptop is portability. Unlike smartphones and most tablets, laptop computers run the same programs as their desktop counterparts, so you can take your work and entertainment with you wherever you go without relying on mobile apps. Depending on your needs, you can either own a laptop alongside a desktop computer or replace your desktop entirely. Now a days laptops are essential for us in our day to day life so that is why we are scrapping laptops, all the latest and old models in a single file so that is becomes easier for a person to compare while buying a laptop. We will basically scrap different sort of laptops including gaming ones and other heavy machines. The data will be scrapped from multiple websites including Daraz.pk, Flipkart, Amazon.com, Ali Baba etc and the scrapped data will be then appended in one single Csv/Excel File. It will have all the basic entities including product name, Brand, Generation, General Description, Price, Rating, colour etc. The UI will be designed by using PY-QT. Scrapping will be done by using Beautiful Soup and various other features. It will have the option to pause, start, resume and stop, with the progress bar showing the progress of tasks or the number of entities scrapped. Now talking about sorting of data. We will apply all the sorting algorithms studies in class including Insertion sort, Merge Sort, Selection Sort, Bubble sort etc The user will have the option to sort the data column wise or row wise in many ways in ascending order or descending order based upon various features.

# Business Case:

**Outline Of the business need for the project:**

We are Scraping 1 million laptops, after loading the data in the csv file our python code will present the whole scraped data and then we can sort the data further through different sorting techniques*.*

Different E-commerce companies and Retailers choose the Best laptops to put on their websites, those ones which are good in performance and are meeting all the definite standards. So for this purpose we are scraping the complete description of a laptop so that it can be easier for them to compare and sale them according to their requirements and customer demands.

**End user of the product:**

End user could be any user, any organization particularly retailers or even buyers willing to buy a laptop.

# Motivation for Project:

**Problem:** Suppose the retailer or any person is buying a new laptop. It becomes very difficult to like compare all the specifications of a laptop. One has to check a lot of things including brand, color, generation, Processor and most importantly price range. This whole procedure takes a lot of time. Like opening a particular website then checking different brands etc will require a lot of time and efforts.

**Solution:** So our project will help the end user to compare all the information of the laptops at a single click.

**State the level of impact expected should the project proceed and implications of not proceeding:**

Whenever the end user wants to compare laptops, the complete data will be available at a single place with just a single click

# Benefits:

* Much Easier Process
* Convenient
* Very Less Time consumption
* Almost no efforts required
* Everything will be Just A click away

# 

# Technology Used:

|  |  |
| --- | --- |
| Languages | Python |
| Platform | Desktop Application |
| Frontend Technology | Python Qt5, QT Designer |
| IDEs | Jupiter Notebook, VS Code |
| UI Prototype | Pencil Tool |

# Technical Details:

|  |  |
| --- | --- |
| ***Technical Details*** |  |
| Name of Entity | Laptop |
| Attributes of Entity  (Minimum seven attributes/rows can be increased) | |  |  |  | | --- | --- | --- | | *Name* | *Data Type* | *Description* | | Description | String | This attribute will be used to store the brand of the product. Every product has a different brand. So thorough this attribute the user can access any brand easily. | | Discount | Integer | This attribute will show the discount percentage of the relevant product | | Price Before Discount |  | This Attribute will show the Price Before Discount of all the products. | | Review | Float | This attribute will show all the reviews of the product. | | Ratings | Integer | Rating means quality use that how much a user is satisfied with this product. A brand having good rating is considered batter in the industry. So through this attribute the users will access the ratings of the product | | Price | Integer | This attribute will be used to store the price of the product. | | Size | Integer | This attribute we will describe the Size in Inches of each laptop.  Like 13 Inch, 15 Inch etc. | |
| Sample of Scrapping Source |  |
| Git-hub Repository Link | https://github.com/NoorAtif/CS261F21PID54 |
| Sorting Algorithms | Sorting means arrange the data and store it into a list so by using the sorting algorithms we can arrange the data in different orders. Like if we want to arrange the data in descending or in ascending order this will be done with the help of the sorting algorithms.  Following are the sorting algorithms that are used in this project |
| |  |  | | --- | --- | | **Algorithm Names** | **Description(Each algorithm in 2-3 lines)** | | Insertion Sort | In computer science the insertion sort is categorized in the comparison sorting algorithm. Time Complexity of the insertion sort is O (n). Insertion sort is the method that return a completed sorted array by selecting one element per loop | | Selection Sort | In computer science the selection sort is categorized in the comparison sorting algorithm. Time complexity of the election sort is O (n^2).Selection sort is not better than the Insertion sort. It will always return the sorted elements store in the list | | Merge Sort | In computer science the merge sort is categorized in the comparison sorting algorithm. Time complexity of the merge sort is n-log-n. In the merge sort we will get the sorted lit by dividing the unsorted array in to halves and then merging it again to get the sorted array | | Bubble Sort | In computer science the bubble sort is categorized in the comparison sorting algorithm. Time complexity taken by the bubble sort is O (n^2). In the bubble sort algorithm we will get a complete sorted list at the end of the loop by compare each element of the list | | Quick Sort | In computer science the quick sort is categorized in the comparison sorting algorithm. Time complexity by the quick sort is O (n) and it is not better than the merge sort. Merge sort is the best sorting algorithm.Quick sort is a popular sorting algorithm that is often faster in practice compared to other sorting algorithms. It utilizes a divide-and-conquer strategy to quickly sort data items by dividing a large array into two smaller arrays. | | Hybrid Sort | A Hybrid Algorithm is an algorithm that combines two or more other algorithms that solve the same problem, either choosing one (depending on the data) or switching between them throughout the algorithm. | | Time Sort | Time sort is a hybrid stable sorting algorithm, derived from merge sort and insertion sort, designed to perform well on many kinds of real-world data. The algorithm finds subsequences of the data that are already ordered (runs) and uses them to sort the remainder more efficiently | | Heap Sort | Heap sort is a comparison-based sorting technique based on Binary Heap data structure. It is similar to selection sort where we first find the minimum element and place the minimum element at the beginning. We repeat the same process for the remaining elements. | | Shell Sort | Shell sort is an optimization of insertion sort that allows the exchange of items that are far apart. The idea is to arrange the list of elements so that, starting anywhere, taking every element produces a sorted list. Such a list is said to be h-sorted. | | Tree Sort | A tree sort is a sort algorithm that builds a binary search tree from the elements to be sorted, and then traverses the tree (in-order) so that the elements come out in sorted order. Its typical use is sorting elements online: after each insertion, the set of elements seen so far is available in sorted order. | | Radix Sort | Radix sort is an integer sorting algorithm that sorts data with integer keys by grouping the keys by individual digits that share the same significant position and value (place value). Radix sort uses counting sort as a subroutine to sort an array of numbers. | | Cube Sort | Cube sort is a parallel sorting algorithm that builds a self-balancing multi-dimensional array from the keys to be sorted. As the axes are of similar length the structure resembles a cube. After each key is  Inserted, the cube can be rapidly converted to an array. | | |
| Searching Algorithms | **Linear Search:**  A linear search is the simplest method of searching a data set. Starting at the beginning of the data set, each item of data is examined until a match is made. Once the item is found, the search ends. If there is no match, the algorithm must deal with this.  **Binary Search:**  Binary search is an efficient algorithm for finding an item from a sorted list of items. It works by repeatedly dividing in half the portion of the list that could contain the item until you have narrowed down the possible locations to just one.  **Jump Search:**  Like Binary Search, Jump Search is a searching algorithm for sorted arrays. The basic idea is to check fewer elements (than linear search) by jumping ahead by fixed steps or skipping some elements in place of searching all elements. |
| Searching Filters for each data type | **For data type of Integer:**  Numbers are compared and a smaller number would be decided.  **For data type of String:**  Strings are compared with each other and a character of smaller ASCII code would be decided. |
| Multi-Level Sorting | Multi-level sorting would be inserted in this project to sort data for particular regions or for particularly filtering out the data. |
| Any other features | **Sorting Label:**  During sorting time of any column, a label that display the progress of the data being sorted.  **URL Option:**  We Will try to take a URL as input and then the data will be scrapped from that particular URL given by the user. |
| ***Interfaces for project*** |  |
| *[Draw layouts in the pencil tool. For each picture of the UI, provide the following table.]*   |  |  |  | | --- | --- | --- | | UI Component Name | Type of UI component | Purpose of UI Component/Other details | | Pause-Button | Button | It will Pause the process of scrapping | | Stop-Button | Button | It will Stop the process of scrapping | | Start-Button | Button | It will again start the process of scrapping once paused or stopped. | | Progress-Bar | Progress-Bar | To show the visual representation of the data being scraped with time in Milliseconds. | | Combo-Selector | Combo Box | To Ask The User various sorting options. I-e ascending, descending etc. And to select the sorting algorithm to sort data by selected column. | | Data-Table | Table | A Table To show all the scrapped data | | Buttons | Button | To manage various tasks and move from one form to another. | | Dash-Board Window | Area | In the dashboard we will have the facility to choose what we want I-e sorting, searching etc. It’s basically a main menu. | | Search and sorting Screens | To-do list | In this part of GUI we will search the scrapped data on the basis of the different attributes like color and brands of the product. On searching the list will provide the selected elements from the scrapped data. So searching the data is the main process of the project. And we can also sort our data | | |
| Final GUI Display:  GUI Description: The GUI Is made By Using Pyqt5 and QT designer tool. First we explored the tools and then started designing our GUI. Different icons and images are used to make the GUI look more attractive. A table is made by using QT Table Widget. Similarly a progress bar is displayed which will show the amount of data scraped. Pause, play and stop button as discussed earlier are designed using the tool. Secondly There is An Option for the user to select the algorithm for sorting. According to the user’s choice data will be sorted. It can be sorted in ascending order, descending order and various other features. Like on the basis of rating, price etc  **Icons Used:**  The icons are selected from Pichon. It Provides many icons and images for GUI.  Following are the icons used in our GUI. | |

# Detailed Description Of Algorithms:

# Selection Sort:

**Description of Algorithm:**

Selection sort is the easiest and simplest sorting algorithm. In This Algorithm the entire array is divided into two sections, the arranged part is at the left side and the unsorted part is at the right side. At first, the arranged part is vacant and the unsorted part is the entire array. This calculation will initially track down the smallest component in the exhibit and trade it with the component in the primary position, then, at that point, it will track down the second smallest component and trade it with the component in the subsequent position, and it will continue to do this until the entire array is completely arranged. In computer science the selection sort is categorized in the comparison sorting algorithm. Time complexity of the election sort is O (n^2).Selection sort is not better than the Insertion sort. It will always return the sorted elements store in the list

**Pseudo Code:**

|  |
| --- |
| procedure selection sort  list : array of items  n : size of list  for i = 1 to n - 1  /\* set current element as minimum\*/  min = i  /\* check the element to be minimum \*/  for j = i+1 to n  if list[j] < list[min] then  min = j;  end if  end for  /\* swap the minimum element with the current element\*/  if indexMin != i then  swap list[min] and list[i]  end if  end for  end procedure |

**Code in Python:**

|  |
| --- |
| def selectionSort(array, size):  for step in range(size):  min\_idx = step  for i in range(step + 1, size):  if array[i] < array[min\_idx]:  min\_idx = i  (array[step], array[min\_idx]) = (array[min\_idx], array[step])  data = [-2, 45, 0, 11, -9]  size = len(data)  selectionSort(data, size)  print('Sorted Array in Ascending Order:')  print(data) |

**Time Complexity:**

Selection Sort requires two nested for loops to complete itself, one for loop is in the function selectionSort, and inside the first loop we are making a call to another function indexOfMinimum, which has the second (inner) for loop.

The Time complexity for selection sort algorithm is given below:

* Worst Case Time Complexity [ Big-O ]: **O(n2)**
* Best Case Time Complexity [Big-omega]: **Ω (n2)**
* Average Case Time Complexity[Big-theta]: **Θ(n2)**

**Dry Run:**

|  |
| --- |
| Consider The Following Array:  3  1  4  5  9    We Will set the first element as the least. Contrast least and the subsequent component. If the subsequent component is greater than least, allocate the second component as least. Contrast least and the third component. Once more, in the event that the third component is greater, dole out least to the third component in any case sit idle. The cycle happens until the last component. After every iteration, least is put towards the front of the unsorted list.    For every iteration, ordering begins from the principal unsorted component. Stage 1 to 3 are rehashed until every one of the components are set at their right positions. |

**Three Strengths:**

* It is very efficient
* It works very well for small data sets
* It Is very easy to implement

**Three Weaknesses:**

* It is slower comparative to other sorting algorithms
* Selection sort requires n-squared number of steps for sorting n elements.
* It is not very efficient when we are dealing with a huge list of items

# Merge Sort:

**Description of Algorithm:**

In computer science the merge sort is categorized in the comparison sorting algorithm. Time complexity of the merge sort is n-log-n. In the merge sort we will get the sorted lit by dividing the unsorted array in to halves and then merging it again to get the sorted array Merge sort is one of the most efficient sorting algorithm based on divide and conquer rule. It first divides the array into equal halves and then combines them in a sorted manner. It repeatedly breaks down the array into several sub arrays until each sub array consists of a single element and then it merges those sub arrays in a manner that results into a sorted array list.

**Pseudo Code:**

|  |
| --- |
| MergeSort(arr[], l, r)  If r > l  1. Find the middle point to divide the array into two halves:  middle m = l+ (r-l)/2  2. Call mergeSort for first half:  Call mergeSort(arr, l, m)  3. Call mergeSort for second half:  Call mergeSort(arr, m+1, r)  4. Merge the two halves sorted in step 2 and 3:  Call merge(arr, l, m, r) |

**Code in Python:**

|  |
| --- |
| def mergeSort(array):  if len(array) > 1:  r = len(array)//2  L = array[:r]  M = array[r:]  mergeSort(L)  mergeSort(M)  i = j = k = 0  while i < len(L) and j < len(M):  if L[i] < M[j]:  array[k] = L[i]  i += 1  else:  array[k] = M[j]  j += 1  k += 1  while i < len(L):  array[k] = L[i]  i += 1  k += 1  while j < len(M):  array[k] = M[j]  j += 1  k += 1  def printList(array):  for i in range(len(array)):  print(array[i], end=" ")  print()  if \_\_name\_\_ == '\_\_main\_\_':  array = [4, 1, 3, 9, 7,]  mergeSort(array)  print("Sorted array is: ")  printList(array) |

**Time Complexity:**

Merge Sort is an efficient, stable sorting algorithm with an average, best-case, and worst-case time complexity of O (n log n)*.*

The Time complexity for merge sort algorithm is given below:

* Worst Case Time Complexity [ Big-O ]: **O(n log n)**
* Best Case Time Complexity [Big-omega]: **Ω (n log n)**
* Average Case Time Complexity[Big-theta]: **Θ(n log n)**

**Proof of Correctness:**

**Initialization:**

In the insertion sort the subarray start with the first element of the array and then it is sorted then begin with the other elements

**Maintenance:**

The iteration of loop increase the size of sub array and when new element enters the array it will maintain the sorting property of sub array it is inserted where it is larger than its left one element.

**Terminology:**

The loop will be terminated when the sub array which is sorted to the size of actual array so I indicated the size of original array*.*

**Three Strengths:**

* It is very efficient
* It can be applied to files of any size.
* It works very well for larger inputs.

**Three Weaknesses:**

* It is slower comparative than other sorting algorithms for smaller inputs.
* Memory sort requires additional space.
* It goes through the whole process even if the array is sorted.

**Dry Run:**

|  |
| --- |
|  |

# Insertion Sort:

**Description of Algorithm:**

In computer science the insertion sort is categorized in the comparison sorting algorithm. Time Complexity of the insertion sort is O (n). Insertion sort is the method that return a completed sorted array by selecting one element per loopInsertion Sort Algorithm is an in-place comparison based algorithm In Insertion sort, The array is searched sequentially and unsorted items are moved and inserted into the sorted sub-list (in the same array).It Is the most simplest sorting algorithm and works efficiently.

**Pseudo Code:**

|  |
| --- |
| INSERTION-SORT(A)  for i = 1 to n  key ← A [i]  j ← i – 1  while j > = 0 and A[j] > key  A[j+1] ← A[j]  j ← j – 1  End while  A[j+1] ← key  End for |

**Code in Python:**

|  |
| --- |
| def insertionSort(array):  for step in range(1, len(array)):  key = array[step]  j = step - 1  while j >= 0 and key < array[j]:  array[j + 1] = array[j]  j = j - 1  array[j + 1] = key  data = [ 5, 7, -8, 9, 10, 4, -7, 0,-12, 1, 6, 2, 3, -4, -15, 12 ]  insertionSort(data)  print('Sorted Array in Ascending Order:')  print(data) |

**Time Complexity:**

Insertion Sort is an easy-to-implement, stable sorting algorithm with time complexity of O (n²) in the average and worst case*,* and O (n)in thebest case*.*

The Time complexity for insertion sort algorithm is given below:

* Worst Case Time Complexity [ Big-O ]: **O(n2)**
* Best Case Time Complexity [Big-omega]: **Ω (n)**
* Average Case Time Complexity[Big-theta]: **Θ(n2)**

**Three Strengths:**

* It is very easy to implement. One of the simplest sorting algorithm.
* It is an in-place sorting algorithm so the memory requirement is minimum.
* It works very well for smaller inputs.

**Three Weaknesses:**

* Many Sorting algorithms are better in performance in comparison to insertion sort.
* The insertion sort is particularly useful only when sorting a list of few items.
* Insertion sort does not deal well with a huge input.

**Dry Run:**

|  |
| --- |
| Consider The Following Array:  3  1  4  5  9    The 1st array element is thought to be sorted. Take the subsequent component and store it independently in key. Contrast key and the 1st array element. In the event that the 1st array element is greater than key, key is put before the 1st element.  Presently, the initial two components are arranged.  Take the third component and contrast it and the components on its left. Set it simply behind the component smaller than it. Assuming there is no component smaller than it, place it toward the start of the array. Similarly place the unsorted elements in their respective positions I-e the correct positions. |

# Quick Sort:

**Description of Algorithm:**

In computer science the quick sort is categorized in the comparison sorting algorithm. Time complexity by the quick sort is O (n) and it is not better than the merge sort. Merge sort is the best sorting algorithm.Quick sort is a popular sorting algorithm that is often faster in practice compared to other sorting algorithms. It utilizes a divide-and-conquer strategy to quickly sort data items by dividing a large array into two smaller arrays Quick sort is a highly efficient sorting algorithm. It is based on partitioning of array of data into smaller arrays. A large array is partitioned into two arrays one of which holds values smaller than the specified value, say pivot, based on which the partition is made and another array holds values greater than the pivot value. It partitions an array and then calls itself recursively twice to sort the two resulting subarrays.

**Pseudo Code:**

|  |
| --- |
| quickSort(arr[], low, high)  {  if (low < high)  {  /\* pi is partitioning index, arr[pi] is now  at right place \*/  pi = partition(arr, low, high);  quickSort(arr, low, pi - 1); // Before pi  quickSort(arr, pi + 1, high); // After pi  }  }  partition (arr[], low, high)  {  // pivot (Element to be placed at right position)  pivot = arr[high];  i = (low - 1) // Index of smaller element and indicates the  // right position of pivot found so far  for (j = low; j <= high- 1; j++)  {  // If current element is smaller than the pivot  if (arr[j] < pivot)  {  i++; // increment index of smaller element  swap arr[i] and arr[j]  }  }  swap arr[i + 1] and arr[high])  return (i + 1)  } |

**Code in Python:**

|  |
| --- |
| def partition(arr, low, high):  i = (low-1)  pivot = arr[high]  for j in range(low, high):  if arr[j] <= pivot:  i = i+1  arr[i], arr[j] = arr[j], arr[i]  arr[i+1], arr[high] = arr[high], arr[i+1]  return (i+1)  def quickSort(arr, low, high):  if len(arr) == 1:  return arr  if low < high:  pi = partition(arr, low, high)  quickSort(arr, low, pi-1)  quickSort(arr, pi+1, high)  return arr  arr = [4, 1, 3, 9, 7]  n = len(arr)  quickSort(arr, 0, n-1)  print(arr) |

**Dry Run:**

|  |
| --- |
| Consider The Following Array:  3  1  4  5  9    There are various varieties of quicksort where the pivot component is chosen from various positions. Here, we will choose the furthest right component of the array as the pivot component. Presently the components of the array are revamped so components that are smaller than the pivot are put on the left and the components greater than the pivot are put on the right. Presently, pivot is contrasted with different components. In the event that a component smaller than the turn component is reached, the smaller element is swapped with the larger element viewed as before.The interaction happens until the second last component is reached. At last, the pivot component is swapped with the subsequent pointer. |

**Time Complexity:**

The Time complexity for quick sort algorithm is given below:

* Worst Case Time Complexity [ Big-O ]: **O(n2)**
* Best Case Time Complexity [Big-omega]: **Ω (n log(n))**
* Average Case Time Complexity[Big-theta]: **Θ(n log(n))**

**Three Strengths:**

* The quick sort is regarded as the best sorting algorithm.
* It can deal well with a huge list of items.
* As it sorts in place, so no additional storage is required.

**Three Weaknesses:**

* Its worst-case performance is similar to average performances of the bubble, insertion or selections sorts.
* If the list is already sorted than quick sort is not much efficient.
* It is fragile, i.e. a simple mistake in the implementation can go unnoticed and cause it to perform badly.

# Bubble Sort:

**Description of Algorithm:**

In computer science the bubble sort is categorized in the comparison sorting algorithm. Time complexity taken by the bubble sort is O (n^2). In the bubble sort algorithm we will get a complete sorted list at the end of the loop by compare each element of the listBubble Sort is the simplest sorting algorithm that works by repeatedly swapping the adjacent elements if they are in wrong order.

**Pseudo Code:**

|  |
| --- |
| procedure bubbleSort( list : array of items )  loop = list.count;  for i = 0 to loop-1 do:  swapped = false  for j = 0 to loop-1 do:  /\* compare the adjacent elements \*/  if list[j] > list[j+1] then  /\* swap them \*/  swap( list[j], list[j+1] )  swapped = true  end if  end for  /\*if no number was swapped that means  array is sorted now, break the loop.\*/  if(not swapped) then  break  end if  end for  end procedure return list |

**Code in Python:**

|  |
| --- |
| def bubbleSort(arr):  n = len(arr)  for i in range(n-1):  for j in range(0, n-i-1):  if arr[j] > arr[j + 1] :  arr[j], arr[j + 1] = arr[j + 1], arr[j]  arr = [64, 34, 25, 12, 22, 11, 90]  bubbleSort(arr)  print ("Sorted array is:")  for i in range(len(arr)):  print ("% d" % arr[i]), |

**Proof of Correctness:**

**Initialization:**

In the insertion sort the subarray start with the first element of the array and then it is sorted then begin with other elements.

**Maintenance:**

The iteration of loop increase the size of sub array and when new element enters the array it will maintain the sorting property of sub array it is inserted where it is larger than its left one element.

**Terminology:**

The loop will be terminated when the sub array which is sorted to the size of actual array so I indicated the size of original array*.*

**Dry Run:**

|  |
| --- |
| Consider The Following Array:  3  1  4  5  9    Beginning from the primary list, analyse the first and the subsequent components.  If the primary component is more noteworthy than the subsequent component, they are swapped.  Presently, analyse the second and the third components. Swap them in case they are not All Together. The above cycle happens until the last component. A similar cycle continues for the remaining iterations. After every iteration, the biggest component among the unsorted components is set towards the end. In every iteration, the examination happens up to the last unsorted component. The array is sorted when every one of the unsorted components are put at their right positions. |

**Time Complexity:**

Bubble Sort is an easy-to-implement, stable sorting algorithm with a time complexity given below.

The Time complexity for bubble sort algorithm is given below:

* Worst Case Time Complexity [ Big-O ]: **O(n2)**
* Best Case Time Complexity [Big-omega]: **Ω (n)**
* Average Case Time Complexity[Big-theta]: **Θ(n2)**

**Three Strengths:**

* The space requirement is at a minimum.
* Elements are swapped in place without using additional temporary storage.
* Bubble sort is very popular and easy to implement.

**Three Weaknesses:**

* The main disadvantage of the bubble sort is the fact that it does not deal well with a list containing a huge number of items.
* The bubble sort requires n-squared processing steps for every n number of elements to be sorted.
* The bubble sort is mostly suitable for academic teaching but not for real-life applications.

# Cube Sort:

**Description of Algorithm:**

Cube sort is a parallel sorting algorithm that builds a self-balancing multi-dimensional array from the keys to be sorted. As the axes are of similar length the structure resembles a cube. After each key is inserted, the cube can be rapidly converted to an array.

**Code in Python:**

|  |
| --- |
| def sortArr(arr, n):  arr = [(i \* i \* i, i) for i in arr];  arr.sort()  for i in range(n):  print(arr[i][1], end = " ");  if \_\_name\_\_ == "\_\_main\_\_" :  arr = [ 4, -1, 0, -5, 6 ];  n = len(arr);  sortArr(arr, n); |

**Time Complexity:**

The Time complexity for Cube sort algorithm is given below:

* Worst Case Time Complexity [ Big-O ]: **O(n log n)**
* Best Case Time Complexity [Big-omega]: **Ω (n)**
* Average Case Time Complexity[Big-theta]: **Θ(nlog n)**

**Three Strengths:**

* It is a parallel sorting algorithm
* It uses a specialized binary search.
* It is very popular

**Three Weaknesses:**

* When an axis grows too large it is split.
* It is costly
* It is not easier to implement

# Heap Sort:

**Description of Algorithm:**

Heap sort is a comparison-based sorting technique based on Binary Heap data structure. It is similar to selection sort where we first find the minimum element and place the minimum element at the beginning. We repeat the same process for the remaining elements. It works by visualizing the elements of the array as a special kind of complete binary tree known as heap.

**Pseudo Code:**

|  |
| --- |
| void heapify(int arr[], int n, int i) {  // Find largest among root, left child and right child  int largest = i;  int left = 2 \* i + 1;  int right = 2 \* i + 2;  if (left < n && arr[left] > arr[largest])  largest = left;  if (right < n && arr[right] > arr[largest])  largest = right;  // Swap and continue heapifying if root is not largest  if (largest != i) {  swap(&arr[i], &arr[largest]);  heapify(arr, n, largest);  }  } |

**Code in Python:**

|  |
| --- |
| def heapify(arr, n, i):  # Find largest among root and children  largest = i  l = 2 \* i + 1  r = 2 \* i + 2  if l < n and arr[i] < arr[l]:  largest = l  if r < n and arr[largest] < arr[r]:  largest = r  # If root is not largest, swap with largest and continue heapifying  if largest != i:  arr[i], arr[largest] = arr[largest], arr[i]  heapify(arr, n, largest)  def heapSort(arr):  n = len(arr)  # Build max heap  for i in range(n//2, -1, -1):  heapify(arr, n, i)  for i in range(n-1, 0, -1):  # Swap  arr[i], arr[0] = arr[0], arr[i]  # Heapify root element  heapify(arr, i, 0)  arr = [1, 12, 9, 5, 6, 10]  heapSort(arr)  n = len(arr)  print("Sorted array is")  for i in range(n):  print("%d " % arr[i], end='') |

**Time Complexity:**

Bubble Sort is an easy-to-implement, stable sorting algorithm with a time complexity given below.

The Time complexity for bubble sort algorithm is given below:

* Worst Case Time Complexity [ Big-O ]: **O(nlog n)**
* Best Case Time Complexity [Big-omega]: **Ω (n log n)**
* Average Case Time Complexity[Big-theta]: **Θ(nlog n)**

**Three Strengths:**

* Heap sort algorithm is very efficient.
* In-place sorting algorithm, less memory required.
* It is simpler to understand and can be implemented consistently.

**Three Weaknesses:**

* It is not much stable and the values may vary.
* It uses 0(1) memory space for the sorting operation.
* It takes more time.

# Counting sort:

**Description of Algorithm:**

Counting sort is a popular and efficient sorting algorithm. Counting sort algorithm is used to sort the small input by dealing them as a small keys. Counting sort is non-comparing algorithm. An algorithm in which we cannot simply compare the two objects.

**Pseudo code:**

|  |
| --- |
| COUNTING-SORT(A, B, k)  let C(0….. k) be a new array  for i =0 to k:  C[i]= 0  4 for j = 1 to A.length  5 C[A[j]] =C[A[j]]+1  6 C[i] now contains the number of elements equal to i.  7 for i=1 to k  8 C[i]=C[i]=C[i]+ C[i] -1  9 // C[i] now contains the number of elements less than or equal to i.  10 for j =A.length downto 1  11 B[C[A[j]]] = A[j]  12 C[A[j]]= C[A[j]]-1 |

**Code in python:**

|  |
| --- |
| def CountSort(A):      max = int(FindMax(A))      length=len(A)        C=[0 for I in range (max+1)]      B=[0 for i in range (length)]      for j in range(0,len(A)):            C[A[j]]=C[A[j]]+1      for i in range(0,len(C)):          C[i]=C[i]+C[i-1]      i = length -1      while i >= 0:          C[A[i]]-=1          B[C[A[i]]-1]=A[i]          i -= 1      return B |

**Dry Run:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. Let the input array is  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **5** | **5** | **1** | **0** | **2** | **0** | **4** | **2** | **3** | **8** |  1. Now find the Highest number in the array 2. Create another array having the range of highest number name as count 3. Initialize the whole array with 0  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |   **5.** This array will count the number of repetitive elements in the array  6.So the array will be   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **2** | **1** | **2** | **1** | **0** | **2** | **0** | **0** | **1** | **0** |   **7.** Now apply commutative frequency rule on the new array named counting array  8.create another array named output array and initialize the complete array with 0   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |   **9.** Find the index of each element of the original array in the count array. This gives the cumulative count. Place the element at the index calculated as shown in figure below  10.The sorted array is   |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | 0 | 0 | 1 | 2 | 2 | 5 | 5 | 7 | |

**Time Complexity:**

The Time complexity of the counting sort I O (n+k). Where n is the number of elements in the array and the k is the range of the elements. The Counting Sort is the best and an efficient algorithm to solve if the number of elements in the array almost equal to the range of the array.

* Time complexity of the counting sort is O (n+k)

**Three Strengths:**

* Stable Sorting Technique
* Efficient enough if the range of the array and the number of elements in the array is almost equal
* Best for the elements if the difference between the elements is small

**Three Weaknesses:**

* Non Comparison Algorithm
* Has big time complexity
* Has bid space complexity

**Proof of correctness:**

**Initialization:**

After for loop of initializes the array C to all zeros, for loop of lines 4–5 inspects each input element.

**Maintenance:**

If the value of an input element is i, we increment C[i]. Thus, C[i]. Holds the number of input elements equal to.

# Bucket Sort:

**Description of Algorithm:**

Bucket sort algorithm is a sorting algorithm which is done by dealing the elements if the array in the form of buckets. Firstly it deal with the elements and store all the elements in the different buckets and then you can solve the problem either by another sorting algorithm or by simply recursively calling the bucket algorithm.

**Pseudo Code:**

|  |
| --- |
| BUCKET-SORT(A):  1. let B(0 ….. n – 1)be a new array  2 .n = A.length  3. for i = 0 to n- 1:  4 make B[i] asan empty list  5 for i = 1 to n:  6 insert A[i]into list B[[nA[i]]]  7 for i = 0 to n – 1:  8 sort list B[i] with insertion sort  9 concatenate the lists B[0]; B[1]; : : : ; B[n - 1]together in order |

**Code in python:**

|  |
| --- |
| def bucketSort(array):  bucket = []    for i in range(len(array)):  bucket.append([])  for j in array:  index\_b = int(10 \* j)  bucket[index\_b].append(j)  for i in range(len(array)):  bucket[i] = sorted(bucket[i])  k = 0  for i in range(len(array)):  for j in range(len(bucket[i])):  array[k] = bucket[i][j]  k += 1  return array |

**Dry Run:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. Let the input array is  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 52 | 37 | 98 | 54 | 21 | 01 | 28 | 10 | 46 | 73 |  1. Now find the Highest number in the array andzthan count the number of digit in that elements 2. Create buckets having the range of highest number (number of digits) 3. Sort the array with the last digit  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 10 | 01,21 | 52 | 73 | 54 | 46 | 37 | 28,98 | 0 | 0 |   5. Now arrange the elements according to the 1st digit  6.So the array will be   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 01 | 10 | 21,28 | 37 | 46 | 52,54 |  | 73 |  | 98 |   8.Now the sorted array is   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 01 | 10 | 21 | 28 | 37 | 46 | 52 | 54 | 73 | 98 | |

**Time Complexity:**

Time Complexity is time taken by an algorithm to solve the given problem and produce a sorted output

* The best case time complexity of the bucket sort is O(n+k)
* The Worst time complexity of the bucket sort is O(n^2)

**Three Strengths:**

* Stable Sorting Technique
* Bucket sort id quicker than the bubble sort
* Putting the elements in to different buckets and then solve the problem reduce the number of the comparisons take place. Which reduce the time to run the program

**Three Weaknesses:**

* Non Comparison Algorithm
* Cannot apply on all the data types you need a best bucketing scheme for this
* Has big time complexity

# Radix Sort:

**Description of Algorithm:**

Radix sort is an integer sorting algorithm that sorts data with integer keys by grouping the keys by individual digits that share the same significant position and value (place value). Radix sort uses counting sort as a subroutine to sort an array of numbers.Radix sort is the algorithm used by the card-sorting machines you now find only in computer museums. In a typical computer, which is a sequential random-access machine, we sometimes use radix sort to sort records of information. Intuitively, you might sort numbers on their most significant digit, sort each of the resulting bins recursively, and then combine the decks in order.

**Pseudo Code:**

|  |
| --- |
| Radix-Sort(A, d)  for j = 1 to d do  int count[10] = {0};  for i = 0 to n do  count[key of(A[i]) in pass j]++  for k = 1 to 10 do  count[k] = count[k] + count[k-1]  for i = n-1 downto 0 do  result[ count[key of(A[i])] ] = A[j]  count[key of(A[i])]--  for i=0 to n do  A[i] = result[i] |

**Python code:**

|  |
| --- |
| def countingSort(array, place):  size = len(array)  output = [0] \* size  count = [0] \* 10  for i in range(0, size):  index = array[i] // place  count[index % 10] += 1  for i in range(1, 10):  count[i] += count[i - 1]  i = size - 1  while i >= 0:  index = array[i] // place  output[count[index % 10] - 1] = array[i]  count[index % 10] -= 1  i -= 1  for i in range(0, size):  array[i] = output[i]  def radixSort(array):  max\_element = max(array)  place = 1  while max\_element  countingSort(array, place)  place \*= 10  data = [12, 42, 54, 23, 1, 85, 78]  radixSort(data)  print(data) |

**Proof of correctness:**

**Initialization:**

The correctness of radix sort follows by induction on the column being sorted Maintenance

**Maintenance:**

The analysis of the running time depends on the stable sort used as the intermediate sorting algorithm. When each digit is in the range 0 to k, so that it can take on k possible values), and k is not too large, counting sort is the obvious choice. Each pass over n d-digit numbers then takes time, n C k.

**Termination:**

There are d passes, and so the total time for radix sort is d\*n C k. When d is constant and k D O (n), we can make radix sort run in linear time. More generally, we have some flexibility in how to break each key into digits

**Dry Run:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. Let the input array is  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 52 | 37 | 98 | 54 | 21 | 01 | 28 | 10 | 46 | 73 |  1. Now find the Highest number in the array and then count the number of digit in that elements 2. Create buckets having the range of highest number (number of digits) 3. Sort the array with the last digit  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 10 | 01,21 | 52 | 73 | 54 | 46 | 37 | 28,98 | 0 | 0 |   5. Now arrange the elements according to the 1st digit  6.So the array will be   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 01 | 10 | 21,28 | 37 | 46 | 52,54 |  | 73 |  | 98 |   8.Now the sorted array is   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 01 | 10 | 21 | 28 | 37 | 46 | 52 | 54 | 73 | 98 | |

**Three weaknesses:**

* A radix sort is cannot be used for the non –integers. It can only be applicable for the integer numbers
* Radix sort take more time than the Quick sort (space also)
* Radix sort is harder to generalize than all other algorithms

**Three Strengths:**

* Non comparative sorting Algorithm
* One of the most popular algorithm
* Having One time complexity for all inputs
* Space complexity of the radix sort is better than the counting sort.

**Time complexity:**

Time Complexity is time taken by an algorithm to solve the given problem and produce a sorted output

* Worst Case Time Complexity [ Big-O ]: **O(n)**
* Best Case Time Complexity [Big-omega]: **Ω (n log n)**
* Average Case Time Complexity[Big-theta]: **Θ(nlog n)**

**Dry Run:**

# Shell Sort:

**Description of Algorithm:**

Shell sort is an optimization of insertion sort that allows the exchange of items that are far apart. The idea is to arrange the list of elements so that, starting anywhere, taking every element produces a sorted list. Such a list is said to be h-sorted Shell sort algorithm is also known as shell-sort or shell method. Shell sort algorithm is used to solve the algorithms in the form of the group form far apart from each other and then progressively start reducing the space between them.

**Pseudo Code:**

|  |
| --- |
| foreach (gap in gaps)  for (i = gap; i < n; i += 1)  temp = a[i]  for (j = i; j >= gap and a[j - gap] > temp; j -= gap)  a[j] = a[j - gap]  a[j] = temp |

**Code in python:**

|  |
| --- |
| def shellSort(array, n):  # Rearrange elements at each n/2, n/4, n/8, ... intervals  interval = n // 2  while interval > 0:  for i in range(interval, n):  temp = array[i]  j = i  while j >= interval and array[j - interval] > temp:  array[j] = array[j - interval]  j -= interval  array[j] = temp  interval //= 2  data = [9, 8, 3, 7, 5, 6, 4, 1]  size = len(data)  shellSort(data, size)  print('Sorted Array in Ascending Order:')  print(data) |

**Dry run:**

|  |
| --- |
| Consider The Following Array:  3  1  4  5  9    We are utilizing the shell's unique grouping (N/2, N/4,--1) as spans in our calculation.  In the primary circle, on the off chance that the cluster size is N = 8, the components lying at the timespan/2 = 4 are thought about and traded in case they are not all- together.  The 0th component is contrasted and the fourth component.  In the event that the 0th component is more prominent than the fourth one, the fourth component is first put away in temp variable and the 0th component (I-e. more noteworthy component) is put away in the fourth position and the component put away in temp is put away in the 0th position. This interaction continues for every one of the excess components. In the subsequent circle, a time frame/4 = 8/4 = 2 is taken and again the components lying at these spans are arranged. The components at fourth and second position are looked at. The components at second and 0th position are likewise analysed. Every one of the components in the exhibit lying at the current stretch are looked at. A similar cycle continues for residual components. At last, when the span is N/8 = 8/8 =1 then the cluster components lying at the time frame are arranged. The cluster is presently totally arranged. |

**Time Complexity:**

Time Complexity is time taken by an algorithm to solve the given problem and produce a sorted output

* Worst Case Time Complexity [ Big-O ]: **O(n)**
* Best Case Time Complexity [Big-omega]: **Ω ( n)**
* Average Case Time Complexity[Big-theta]: **Θ( n)**

**Three strengths:**

* Best efficient algorithm
* This algorithm avoids large shifts as in case of insertion sort
* One of the most popular algorithm

**Three weaknesses:**

1. Shell sort is that its only efficient for medium size lists
2. For bigger lists, the algorithm is not the best choice
3. Has large time complexity

# Hybrid Sort:

**Description of Algorithm:**

A Hybrid Algorithm is an algorithm that combines two or more other algorithms that solve the same problem, either choosing one (depending on the data) or switching between them throughout the algorithm. In computer science, hybrid algorithms are very common in optimized real-world implementations of recursive algorithms, particularly implementations of divide and conquer or decrease and conquer algorithms, where the size of the data decreases as one moves deeper in the recursion. In this case, one algorithm is used for the overall approach but deep in the recursion, it switches to a different algorithm, which is more efficient on small data.

**Code in python:**

|  |
| --- |
| def insertion\_sort(arr, low, n):  for i in range(low + 1, n + 1):  val = arr[i]  j = i  while j>low and arr[j-1]>val:  arr[j]= arr[j-1]  j-= 1  arr[j]= val    def partition(arr, low, high):  pivot = arr[high]  i = j = low  for i in range(low, high):  if arr[i]<pivot:  a[i], a[j]= a[j], a[i]  j+= 1  a[j], a[high]= a[high], a[j]  return j  def quick\_sort(arr, low, high):  if low<high:  pivot = partition(arr, low, high)  quick\_sort(arr, low, pivot-1)  quick\_sort(arr, pivot + 1, high)  return arr    def hybrid\_quick\_sort(arr, low, high):  while low<high:    # If the size of the array is less  # than threshold apply insertion sort  # and stop recursion  if high-low + 1<10:  insertion\_sort(arr, low, high)  break    else:  pivot = partition(arr, low, high)      if pivot-low<high-pivot:  hybrid\_quick\_sort(arr, low, pivot-1)  low = pivot + 1  else:  # If the right side of pivot is less  # than left, sort right side and  # move to the left side  hybrid\_quick\_sort(arr, pivot + 1, high)  high = pivot-1  a = [ 24, 97, 40, 67, 88, 85, 15,  66, 53, 44, 26, 48, 16, 52,  45, 23, 90, 18, 49, 80, 23 ]  hybrid\_quick\_sort(a, 0, 20)  print(a) |

**Time Complexity:**

Time Complexity is time taken by an algorithm to solve the given problem and produce a sorted output

* Worst Case Time Complexity [ Big-O ]: **O(n^2)**
* Best Case Time Complexity [Big-omega]: **Ω (n)**
* Average Case Time Complexity[Big-theta]: **Θ(n)**

**Three Weaknesses:**

* Non Comparison Algorithm
* Has big time complexity
* Has bid space complexity

**Three Strengths:**

* Stable Sorting Technique
* Hybrid sort is quicker than the bubble sort
* This algorithm avoids large shifts as in case of insertion sort

# Collaboration:

The Collaboration between Both the members went very well. The Ideas and discussions were held for the improvement of various functionalities in the project. A lot of efforts were put in by both the members. Special Attention was paid in making reports. All the commits were being done by both the members timely on The Git-hub Repository. There was as such no issue faced in collaboration. Everything was done in a professional way. This project is totally a result of mutual effort of both the members. Time was managed perfectly by both the members. Only the Algorithm report took more than expected amount of time. The other tasks were easily handled and were submitted even hours before the deadline.

# Task Division:

* Tasks were equally divided between both the Members.
* All The Reports were being divided. Like In Algorithm Report 6, 6 algorithms were divided.
* The Main Searching And Idea Was Given By Both The Members
* Scraping was done from 2 different websites. Firstly Both The Members Did It From Daraz.pk. Then after some issues (Mentioned in Issues Heading) we switched to Flipkart.
* The same issue was observed with this website. Then after Discussion with Our Supervisor we chose to move forward without scraping 1 million data.
* GUI was first made by 2020-CS-111. And then it was Further Updated, Icons And colours were added by 2020-CS-129
* Data Loading was done by both the members
* Implementation is in progress.

**Conclusion:**

As A Result the Project was collectively done by both the members.

# Issues Faced:

Despite all our efforts a lot of issues were being faced by us mostly in the data scraping part. The main issue faced by us was insufficient amount of data. Million data was not available even 1, 2 Lac is a huge amount. The data found online is just 2-3,000. So at first we tried scraping from Daraz.pk. We were actually being fooled as it was showing total result of 40 Lac something data. We just checked a few pages and were sure that data is available here. But during scraping we came to know that the data was minimal. It was Just 4000 Something and not just laptops. But Laptop Chargers and other accessories were also being included in the Laptop Section. We were dis-hearted due to this as we gave a lot of time to this and THE DATA WAS NOT SCRAPED. Then we decided to switch our website, we moved to Flipkart.pk. The same thing again happened it was showing 5 Lac Data and only 1000 was present. At First we checked 10, 20 pages and thought maybe data is present here. But we were panicked at that time and didn’t saw all the results that was a bit mistake at our side. So just a 1000 data of laptops is being scraped by us. We didn’t lose hope and tried scraping from amazon.pk but due to some security issues it was not easy for us to scrape data from a protected website. We tried searching other websites so that we can further merge the common entities. But unfortunately the data available on most of the websites was only in 200-500 range plus the entities were also not matching. So that were all the issues which were faced by us during the project. Despite Scraping issues all the other deadlines were smoothly met and no as such issue was faced.

# Implementation:

Sorting is in Process.

# Project Plan

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr.#** | **Feature’s** | **Member Name** | **Completion Date** |
| 1 | Project Proposal | Noor Atif & Aqsa Riaz | October 15, 2021 (EDUKO)  October 18, 2021 (Hard Form) |
| 2 | Scrapping | Noor Atif & Aqsa Riaz | October 19, 2021 |
| 3 | UI Implementation | Noor Atif & Aqsa Riaz | October 24, 2021 |
| 4 | Sorting Algorithms | Noor Atif & Aqsa Riaz | October 27, 2021 |
| 5 | Integration And Project Report | Noor Atif & Aqsa Riaz | November 3, 2021 |