**Data Structure and Algorithm**



Session: 2022 – 2026

**Submitted by:**

Syed Muhammad Faisal Ilyas 2022-CS-63

Gul-e-Zahra 2022-CS-75

**Supervised by:**

Sir Nazeef ul Haq

Department of Computer Science

**University of Engineering and Technology**

**Lahore Pakistan**

Table of Contents

Proposed Project Title 3

Description 3

Project Domain 3

Platform 3

Attributes3

Gitlab Repository link3

Algorithms3

UI Components4

# **Proposed Project Title:**

**Ecommerce Data Scrapper**

# **Description:**

An e-commerce data scraper is like a digital detective that helps to collect information from online shopping websites. It looks at the web pages of these websites and finds important details like product prices, descriptions, rating, delivery details, discount price, and customer ratings and reviews. People use these scrapers to compare prices, keep track of what’s available, and understand what customers are saying about products.

# **Project Domain**:

The project domain typically revolves around the e-commerce industry itself. It gathering information about products, prices, availability, reviews, and other related data.

# **Platform:**

**Flipkart**

**url**=https://www.flipkart.com/search?q=mobiles&amp;as=on&amp;as-show=on&amp;otracker=AS\_Query\_TrendingAutoSuggest\_1\_0\_na\_na\_na&amp;otracker1=AS\_Query\_TrendingAutoSuggest\_1\_0\_na\_na\_na&amp;as-pos=1&amp;as-type=TRENDING&amp;suggestionId=mobiles&amp;requestId=0dbbf0e0-bd6f-462f-b0c1-6abc15439d84&amp;page=

# **Attributes:**

|  |  |
| --- | --- |
| **Attributes** | **Definition** |
| Name | Contains product name |
| Price | Indicate price of Product |
| Description | Category of Product |
| Rating | Rating given by users |
| No. of Reviews | Reviews given by the users |
| Discounted Price | Price after discount |
| Delivery | Product delivery |
| Discount Percentage | OFF % |

# **Gitlab Repository link:**

**Link**: <https://gitlab.com/faisal-ilyas/mid-porject>

# **Algorithms:**

|  |  |
| --- | --- |
| Sorting Algorithm | |
| Bubble sort | Repeatedly compares and swaps adjacent elements if they're in the wrong order until the list is sorted, with a time complexity of O(n^2). |
| Insertion sort | Builds the sorted list one element at a time by inserting each element into its correct position among the already sorted elements, with an average time complexity of O(n^2). |
| Selection sort | Finds the smallest (or largest) element and swaps it with the first (or last) unsorted element in each pass until the list is sorted, with a time complexity of O(n^2). |
| Merge sort | Divides the list into smaller halves, recursively sorts them, and then merges the sorted halves to produce the final sorted list, with a time complexity of O(n log n). |
| Hybrid Merge sort | A modified merge sort that may switch to a different sorting algorithm (like insertion sort) for small sublists to improve efficiency. |
| Quick sort | Selects a "pivot" element, partitions the list into elements smaller and larger than the pivot, and recursively sorts the sublists, with an average time complexity of O(n log n). |
| Bucket sort | Bucket sort is a linear time sorting algorithm that distributes elements into a finite number of buckets, sorts each bucket, and concatenates them to achieve overall sorting. It is effective for uniformly distributed data but less efficient when data is highly skewed. |
| Radix sort | Radix sort is a sorting algorithm that processes integers digit by digit, sorting from the least significant digit to the most significant digit, resulting in linear time complexity for fixed-size integers. |

# **UI Components:**

|  |  |  |
| --- | --- | --- |
| **UI Component Name** | **Type** | **Purpose** |
| Enter URL | Text Box | It will take URL as input to scrap. |
| Export from CSV | Push Button | By pushing this button, it will load data from CSV file to data grid. |
| Sorting Algorithm | Drop Down Menu | Display a list to select sorting algorithm. |
| Time Taken | Text Box | Display time taken by algorithm. |
| Scrap | Push Button | Press this button to scrap data from URL. |
| Sort | Push Button | Press this button to sort. |
| Select column to search | Drop Down Menu | Display list of columns to select for sorting. |
| Search | Push Button | Press this button to search. |

