**ONESTOP CODING DASHBOARD USING WEB SCRAPING**

**A Project Report submitted in partial fulfillment of the requirements for the award**

**of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

**Submitted by**

**Bhange Ravi Shankar, 121810311001**

**Kasa Prathyush Kumar, 121810311036**

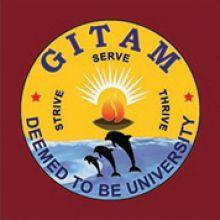
**Madarapu SNSSM Meghana, 121810311044**

**Seelam Sowmith, 121810311049**

**Under the esteemed guidance of**

**Dr.S.V.Lakshmi**

**Assistant Professor**

****

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**GITAM**

**(Deemed to be University)**

**VISAKHAPATNAM**

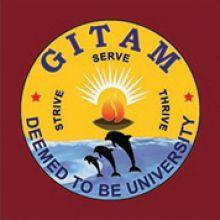
**DECEMBER 2021-APRIL 2022**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**GITAM INSTITUTE OF TECHNOLOGY**

**GITAM**

**(Deemed to be University)**

****

**DECLARATION**

We, hereby declare that the project review entitled **“ONESTOP CODING DASHBOARD USING WEB SCRAPING”** is an original work done in the Department of Computer Science and Engineering, GITAM Institute of Technology, GITAM (Deemed to be University) submitted in partial fulfillment of the requirements for the award of the degree of B.Tech. in the Computer science and Engineering. The work has not been submitted to any other college or university for the award of any degree or diploma.

Date: 05-04-2022

**Registration No(s). Name(s) Signature(s)**

121810311001 Bhange Ravi Shankar A black and white sign

Description automatically generated with low confidence

121810311036 Kasa Prathyush Kumar A picture containing diagram

Description automatically generated

121810311044 Madarapu SNSSM Meghana 

121810311049 Seelam Sowmith A white surface with writing on it

Description automatically generated with low confidence

****

**TABLE OF CONTENTS**

|  |  |  |  |
| --- | --- | --- | --- |
| 1. | Abstract | | 01 |
| 2. | Introduction | | 02 |
| 3. | Literature Review | | 03 |
| 4. | Problem Identification & Objectives | | 04 |
| 5. | System Methodology | | 05-06 |
| 6. | Overview of Technologies | | 07-08 |
| 7. | Implementation | | 09-21 |
|  | 7.1 | Coding |  |
|  | 7.2 | Testing |  |
| 8. | Results & Discussions | | 22-28 |
| 9. | Conclusion & Future Scope | | 29 |
| 10. | References | | 30 |

1. **ABSTRACT**

As you know, it is difficult for users to navigate each website. This website creates users’ coding portfolios and provides user activity which includes his/her no of problems solved in a day. By taking scores from different coding websites, we create the overall score and provide the leaderboard with global ranking and institute ranking which makes users analyze where his/her position is. For calculating the overall score of a user, we used a standard mathematical equation that gives an accurate score according to his/her individual scores on individual platforms. The website also shows the daily activity status. And coming to the activity of a user will be tracked every 24 hours. It is implemented by creating a background thread that runs and performs tasks like updating user scores if anything changed. We considered websites like Codechef, Codeforces, GeeksforGeeks, and InterviewBit. We retrieved users’ data like no of problems solved, rating, score, streak. For retrieving data from websites, we used web scraping technology. Backend is designed using Java 11 and for web scraping, we used JSOUP API.

1. **INTRODUCTION**

There are great platforms such as CODECHEF, CODEFORCES, GEEKSFORGEEKS, and INTERVIEWBIT, but it is very complicated for students and users to navigate from each platform to see scores, ratings, and upcoming contests. Our project is a website that can track all coding results from these platforms, provide an overall score, and maintain users’ daily and weekly activity regardless of whether the user is active or not. It is designed to solve this kind of navigation problem faced by students/users when there is only one, whether to join a particular platform.

This website creates a user coding portfolio. By collecting scores from different coding platforms, we summarize the overall score and provide the leaderboard with a global and institutional ranking. This allows users to analyze their position. There is a contest page where you can see upcoming contests from various platforms such as CODECHEF, CODEFORCES, GEEKSFORGEEKS, and INTERVIEWBIT. This website also shows your daily activity status, such as the number of problems solved in one day and your weekly activity status. Data is retrieved from the live websites and stored in our database to track user activity and we are using JSOUP API for web scraping. User activity is updated every 24 hours.

We mainly focus on the overall score and activity of a user. To calculate the overall score of a user, we gave priority to contest rating and then problems solved in each platform. Graph of overall score of a user mostly depends on contest ratings of different websites. As his/her contest rating increases, the overall score also increases in the same way as a quadratic equation where contest ratings are real numbers and problems solved in each platform are considered constants. We used threads for each user to increase the performance of our web application while updating user activity data in the database.

1. **LITERATURE REVIEW**

The student performance tracker tool is very important to track his/her all performances done in their respective challenges or test. Here we are going to derive an efficient way to track student coding performance. Student coding performance can be calculated in many different ways but here we are concentrating on some factors to measure student performance. A pr study says that student performance is dependent on major factors like motivation, learning, experience, comfort, confidence level, etc. Our project includes these kinds of factors like motivation and confidence level. Overall problem solved by a student gives a kind of motivation and confidence level to students which helps them a lot to continue with their performance. Not only this but also overall score with global and institute-wise ranking gives us much better motivation to a student. In this project, we mainly focus on student coding performance in which he/she improves their coding experience and our project keeps track of all the no of problems solved in a week and the total no of problems solved in each website. Along with this we also add some features like viewing all contests in our page. We mainly focused on integrating four coding platforms and tracking the student performance by giving them overall scores and tracking their performance.

1. **PROBLEM IDENTIFICATION AND OBJECTIVES**

PROBLEM STATEMENT

Users have accounts on different coding platforms and each website has coding scores and contest ratings. But it’s difficult to analyze the overall coding score. There will be various contests held on various platforms like CODECHEF, CODEFORCES, GEEKSFORGEEKS and INTERVIEWBIT. It’s quite difficult to navigate to each website and check which contest is being held and when.

OBJECTIVES

Our objective is to make a single platform by integrating all four platforms (Codechef, Codeforces, GFG, InterviewBit) so that

* Users can view their all scores in one place
* User can have overall score and leaderboard
* Can check which contest is held at which platform
* Can check user daily activity and weekly activity

EXISTING SYSTEM

* We already have some mobiles apps like CodeClock, CPing, etc. But these apps are built for just notifying which contest is held on which platform and at what time and date. We will get notifications for every contest held on different platforms.

PROPOSED SYSTEM

* Along with the existing features, users can also check is activity and we proposed features like dashboards for each platform and overall score with leaderboards.

1. **SYSTEM METHODOLOGY**

PROJECT PLANNING

RETRIVING DATA FROM CODING PLATFORMS

DESIGNING DATABASE SCHEMA

TESTNG

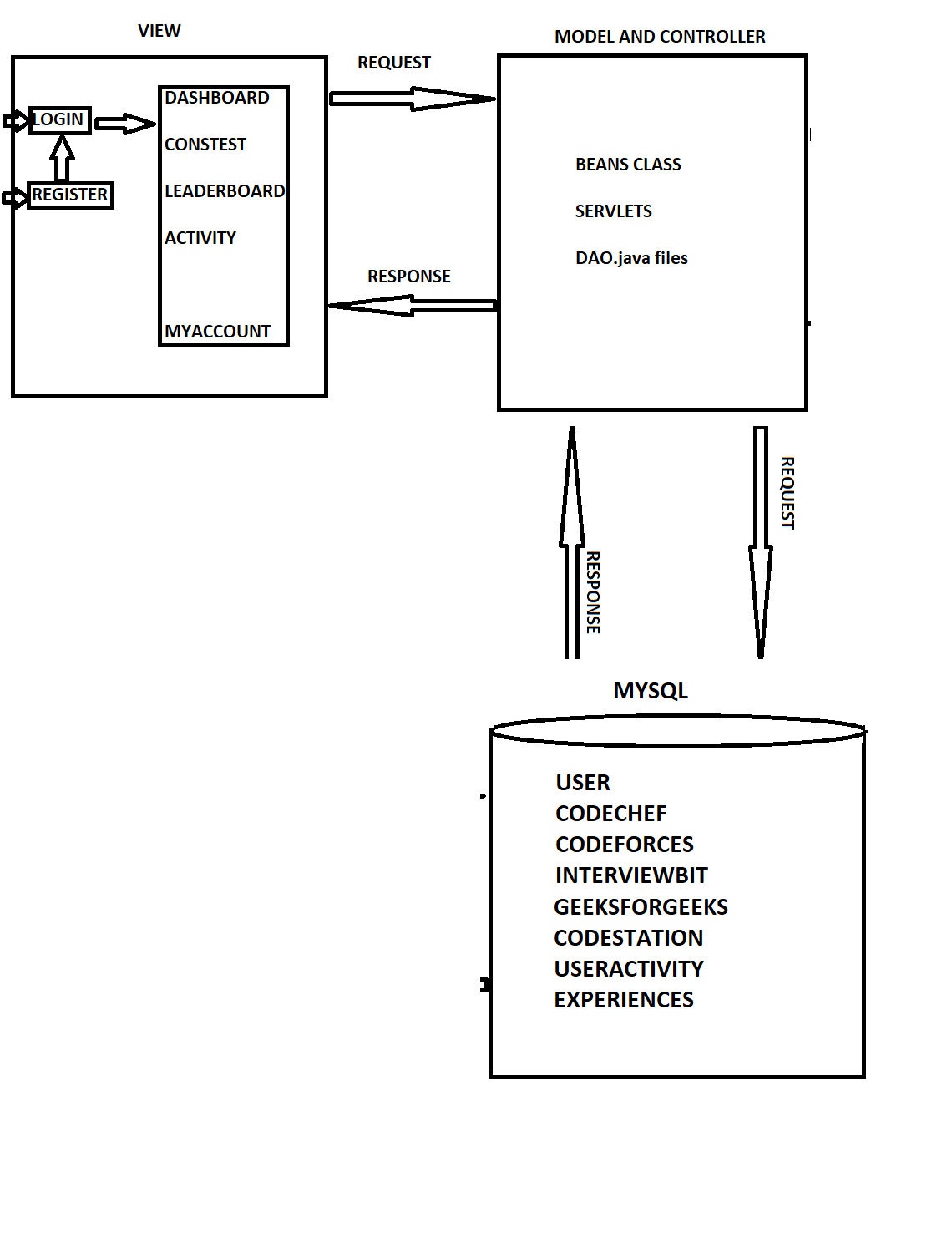
WEBSITE IMPLEMENTATIONN

DEPLOYMENT IN LOCALHOST

ARCHITECTURE

****

WORKFLOW



1. **OVERVIEW OF TECHNOLOGIES**

|  |  |
| --- | --- |
| FRONTEND | HTML 5, CSS, BOOTSTRAP 4, JAVASCRIPT, JSP |
| BACKEND | JAVA 11 |
| DATABASE | MYSQL 8 |
| SERVER | TOMCAT 9 |
| API | JSOUP |
| IDE | ECLIPSE |

1. HTML 5**:**  The Hypertext Markup Language which is used for creating web pages and applications. Hypertext means text within the text. A text has a link within it, which is a hypertext. The markup language is a computer language that is used to apply layout and formatting conventions to a text document.
2. CSS**:** Cascade stylesheets describe how HTML elements appear on screen, paper, or other media. It saves a lot of time and allows you to control the layout of multiple web pages at the same time. This is a simple design language that aims to simplify the process of creating a website.
3. BOOTSTRAP**:** Bootstrap is an HTML, CSS, and JS library that focuses on facilitating the development of useful websites (rather than web apps). This is a front-end framework used for easier and faster web development. The main purpose of adding to a web project is to apply Bootstrap's selected colors, sizes, fonts, and layouts to that project. You can also use the JAVASCRIPT plug-in to easily create responsive designs.
4. SERVLETS ANDJSP: Servlets are used to dynamically create data from users’ create forms and websites and present the results. JSP and Java Server Pages are Servlet-like technologies used to build web applications. A JSP page from HTML code, embedded in one of the Java codes. The server side is requested after the client's Java code has performed such processing, and the generated HTML page is returned to the client browser.
5. MYSQL**:** MySQL is a relational database management system based on SQL structured query language. Like other relational databases, MySQL stores data in row and column tables. Applications are used for a wide spectrum of purposes, including data warehouses, electronic commerce, and logging applications. However, MY SQL's most common use is to store data in the database.
6. JSOUP**:** JSOUP is an open-source Java library used primarily for extracting data from HTML. This is an HTML parser that can also edit and output HTML. It has a stable development line, excellent documentation, and a fluent and flexible API. You can also use JSOUP to parse and create XML.
7. TOMCAT**:** This is an open-source Java Servlet container. The main usage goal is to implement various Java Enterprise specifications such as website APIs, Java Server Pages, and Java Servlets. Developed and maintained by the open developer community with the support of the Apache Software Foundation, it is released under Apache License 2.0.
8. **IMPLEMENTATION**
   1. CODING
   2. TESTING

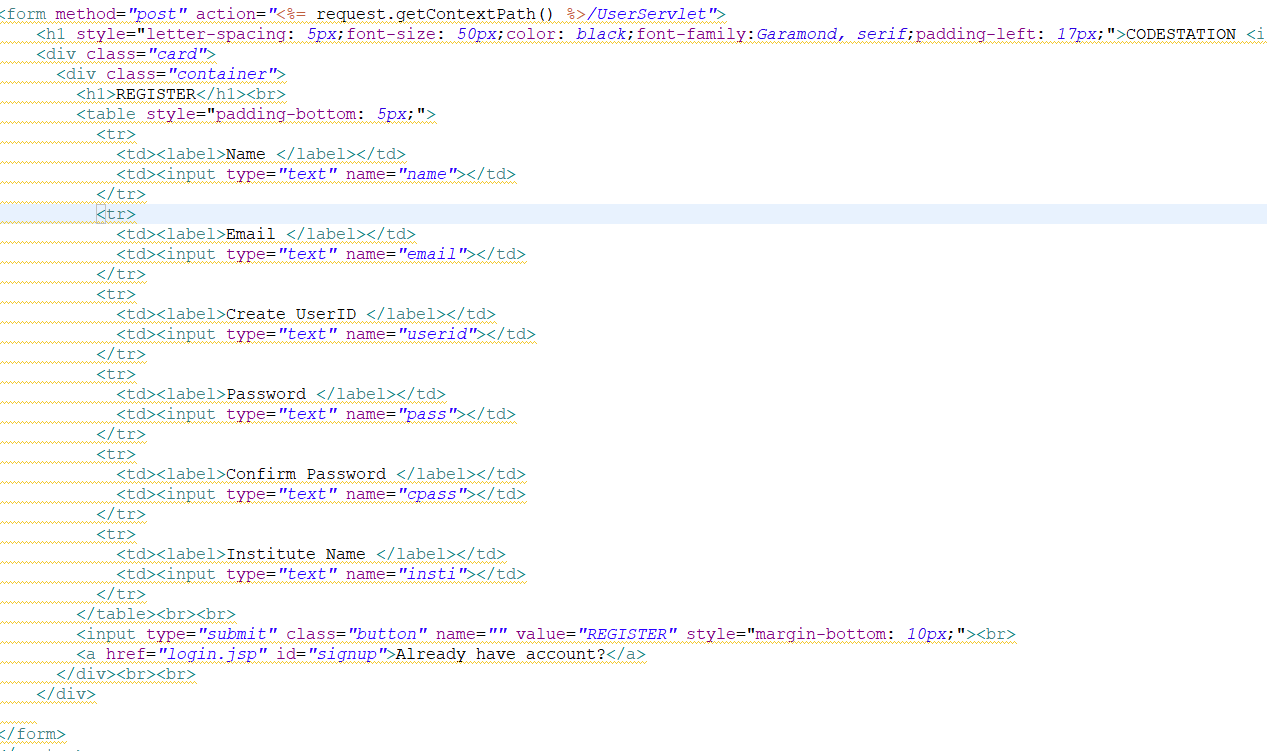
**LOGIN PAGE:**

* The user has to log in using his/her mail id and password.



**REGISTER PAGE:**

* The user has to give his details like Name, Mail Id, User Id, Password, Institute for creating an account on this website.



**USERS TABLE:**

Graphical user interface, text, application, email

Description automatically generated

**MY ACCOUNT PAGE:**

* To create an account on this website users should have an account on websites like CODECHEF, CODEFORCES, GEEKSFORGEEKS, and INTERVIEWBIT.
* After giving all user ids of all platforms, data will be fetched and stored in the database, and by default, activity will be considered Inactive.

Text

Description automatically generated

**DASHBOARD:**

* On this page users can be able to view their different platform scores, ratings as well as overall scores.
* Mathematical Equation for calculating the overall score of a user is given below

Overall Score = (CCPS\*10 +(CCR-1300) ^2/30) +(CFPS\*10 +(CFR-1200) ^2/30) +(IBS/3) + (GFG\*10)

Where,

CCPS = CODECHEF problems solved

CCR = CODECHEF rating

CFPS = CODEFORCES problems solved

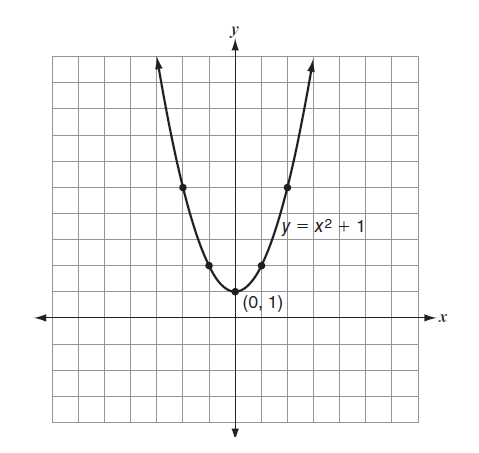
CFR = CODEFORCES rating

IBS = INTERVIEWBIT score

GFG = GEEKSFORGEEKS solved problems

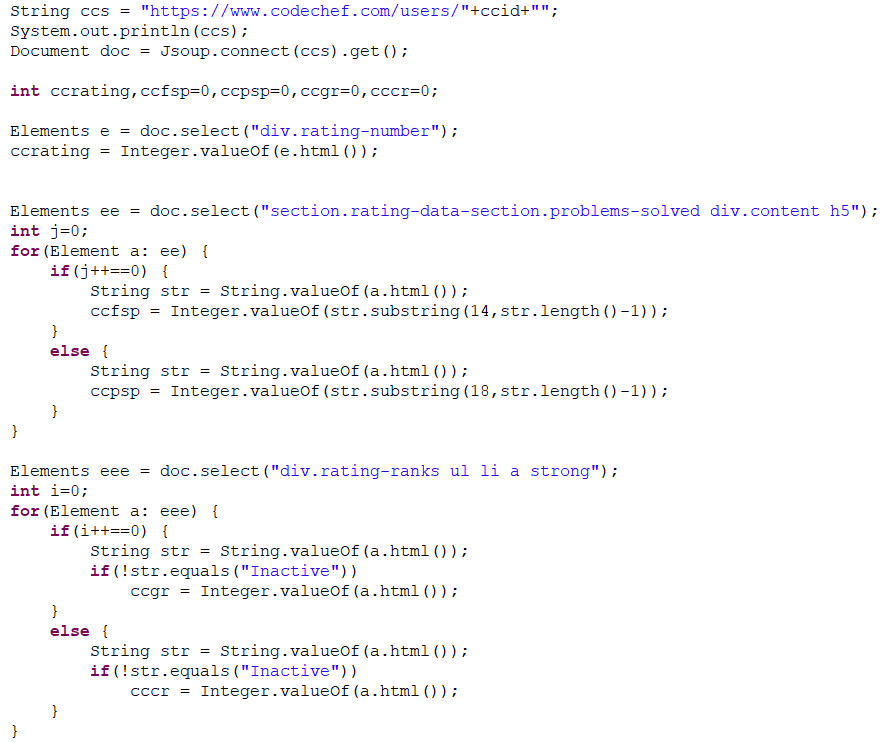
* For example, Codechef Score = CCPS\*10 +(CCR-1300) ^2/30)

That means the equation is in the form of y = x^2 + c where y stands for CodeChef score and x stands for rating and c stands for problems solved in the platform. Here the rate of increase in score mostly depended on the contest rating of a platform. The score increases drastically as the rating increases similar to a quadratic equation graph shown below





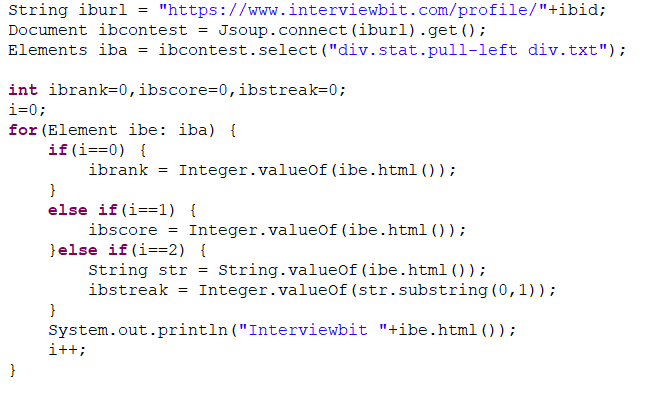
Web scraping of user’s CODECHEF data:



Text

Description automatically generated with medium confidence

Web scraping of user’s INTERVIEWBIT data:



Graphical user interface

Description automatically generated

Web scraping of user’s CODEFORCES data:

Text

Description automatically generated

Graphical user interface

Description automatically generated

Web scraping of user’s GEEKSFORGEEKS data:

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface

Description automatically generated

**LEADERBOARD:**

* We have two leaderboards (Global leaderboard and Leaderboard by Institute)

Graphical user interface, text, application

Description automatically generated

Table

Description automatically generated

**ACTIVITY PAGE:**

* On this page users can view their weekly activity and daily activity from 4 platforms
* In this application, we created a thread called SiteMapThread which runs in the background and performs some background tasks like updating users’ activity. The user activity table will be updated every 24 hours

Text

Description automatically generated

* In daily activity, data will be updated by storing current date data in day7 and previous date data will be stored in previous-> previous Data. Just like adding a node in LinkedList and deleting the first node in LinkedList.
* Along with these dates’ update, if any scores or ratings changed in users’ coding profiles they will be updated in the database to track weekly activity.



Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application

Description automatically generated

1. **RESULTS AND DISCUSSIONS**

**LOGIN PAGE:**

Graphical user interface

Description automatically generated

**REGISTER PAGE:**

Graphical user interface

Description automatically generated

**MYACCOUNT PAGE:**

Graphical user interface, application

Description automatically generated

**DASHBOARD PAGE:**

Graphical user interface, application

Description automatically generated

**CONTEST PAGE**:

Graphical user interface

Description automatically generated with medium confidence

**LEADERBOARD PAGE:**

Global Ranking:

A screenshot of a computer

Description automatically generated

Institute Ranking:

A screenshot of a computer

Description automatically generated

**ACTIVITY PAGE**:

Chart, line chart

Description automatically generated

**Codeforces Activity Table:**

Table

Description automatically generated with medium confidence

**GeeksforGeeks Activity Table:**

A picture containing graphical user interface

Description automatically generated

**CodeChef Activity Table:**

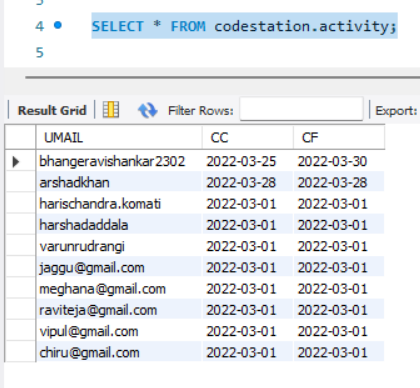
Table

Description automatically generated with medium confidence

**InterviewBit Activity Table:**

Table

Description automatically generated with medium confidence



1. **CONCLUSION AND FUTURE SCOPE**

As we have gone through different modules in the project we can say that this website will be very helpful for students who are willing to improve their skills sets in competitive programming. In this website, we have seen four platforms with ratings and scores of a user. Activity will be tracked every 24 hours. This website can be made more graphical interface by adding graphs and some other way of representation. This project can be extended by increasing the number of platforms and adding some more features like topic recommendations based on the type of problems solved in each platform. If this idea has been implemented then it will be very helpful to check on which topics they should concentrate and they can see how many problems they have solved on each topic. This can be implemented by problem tags in which each problem will be tagged with some topics. We can also add features like follow/following so that they can follow their friends to check their profiles easily. With this, we can also add one more leaderboard among friends for each user so that they can actively check their position among their friends. It would be great if add user submissions on this website so that users can view all submissions in one place and view problems according to tags, max success rate, favorite, etc. so that we will be redirected to that particular website to solve the problem. This website might become a hub if we implement all the above ideas. By adding all these features it will be helpful to students which makes them very easy to track their profiles and all. And even if they want to show their coding profiles they can just share their codestation profile which contains all scores from different platforms and having an overall score from that will create a portfolio

1. **REFERENCES**

* <https://www.w3.org/standards/webdesign/htmlcss>
* <https://en.wikipedia.org/wiki/Bootstrap_(front-end_framework)>
* <https://www.upgrad.com/blog/jsp-vs-servlet/>
* <https://www.123-reg.co.uk/support/servers/what-is-mysql-and-why-do-i-need-it/>
* <https://www.baeldung.com/java-with-jsoup>
* <https://n6host.com/blog/what-is-tomcat-6-reasons-you-should-use-tomcat/>