

E-Learning Platform

Name: Ayush Katoch
Student ID: 201647931

Supervisor: Mr. Henry Forbes

Project Description

Computer science is one of the major fastest-growing fields in the industry and many students are opting for this course for better employment opportunities. In the universities, it's very common to teach programming to Computer Science students and also to students from non-technical courses. Specially for students from non-technical background it's really challenging to complete programming assignment as they first have to download the assignment, code editor / IDE, compilers for the programming language that is to be used, additional libraries for that programming language before they can begin coding. Students often face a lot of problems in setting the development environment due to different versions or operating system. Not only this, when they complete their assignment, they have to upload the files so that they can be graded. For the professors, it's a tiring process to grade these assignments, as they too will first have to setup the development environment which might not work with all the assignments due to different versions of software used by the students. After setting up the development environment professors have to download the student code and run them individually which is a very time-consuming process.

Due to this time-consuming process Universities tend to give few programming assignments even though it's the best way to teach programming to the students. All this has been discussed in a Tufts University paper dated 2001 which talked about the issues with manual grading in more depth (Tufts, 2001 [3]).

In order to solve this issue, this e-learning software will be developed that will allow the professors to upload assignment, lecture videos, slides, notes and much more for the student. This software is specially designed for the programming courses as it allows students to do their programming assignments directly on the website without downloading any extra software and they can even test their code along with the correctness using predefined test cases (designed by the professor). This software also reduces the workload of the professors as it can automatically grade the students code using predefined parameters like code compiles, code runs without any errors, testing student code against predefined test cases and hidden cases.

Key Literature and Background Reading

In order to understand the features that should be implemented, a research on Computing Pedagogy was read. Pedagogy is a way of teaching a subject. Computing Pedagogy means effective ways to teach computing to people. It involves 12 NCCE's principles, Lead with concepts, Work together, Get hands-on, unplug unpack and repack, model everything, foster program comprehension, create projects, add variety, challenge misconceptions, make concrete, structure lessons lastly, read & explore code first. This software could help with the three principles of Computing Pedagogy,

- Lead with concepts: "graphical tools for organising and representing knowledge" (Florida Institute for Human and Machine Cognition , 2006). Concepts maps are used to represent a knowledge visually. It consists of an interconnected network of concepts. Links are used between the concepts to specify the relationships between them. This site will support uploading video lectures and notes that can be used by the students to better understand the concept.
- Work Together: Pair programming is an approach that can be used in a classroom. It involves students working together on a problem while sharing a computer. Students can work together on a programming assignment using this software.
- Get hands-on: Physical computing is a term to describe activities where learners write programs to interact with the real world using specialized hardware. This software could be the best way to practice programming, as it allows students to start coding without having to install / setup any 3rd party software.
- Unplug, unpack, repack: Programming is a subject that has lots of technical terms having a precise technical meaning. In order for students to understand and master

these terminologies is to follow semantic wave (Making semantic waves, 2018). This involves using metaphors, analogies to teach.

- Model Everything: When learning students are always exposed to the completed program and are not shown the troubleshooting that took place while writing the program. Live coding is when a teacher develops the solution to a problem in front of the class to allow learners to follow.

Apart from this, Canvas will be used as a reference for designing the User Interface (UI) of the website. This project is pretty similar to Canvas and using Canvas as reference would make designing UI much easier.

Aims & Objectives

Main aim of this project is to teach programming in a classroom. It also helps to achieve the following aims

- For students
 - o Providing a platform / tool to the students to write and test their code without having to setup any development environment
 - o Providing real time scores on submitted code.
 - o Easy access to all the uploaded lecture notes and videos
- For Professors
 - o Easy way to monitor class performance using visual tools like graphs
 - o Automatic grading of programming assignments using predefined parameters
 - o Plagiarism checker to ensure that academic integrity is maintained

In order to achieve the above aims, following objectives needs to be completed.

Objectives that are to be achieved are as follows:

- 1) Account Management System
 - a. Allowing users to create a new account and log in to their existing account.
 - b. There would be two types of users in the software, Teacher and Student. Both the account type will have different abilities / functionalities.
 - c. On successful login, an authorization token will be generated that needs to be passed in the headers by the user in order to authenticate themselves.
- 2) Course Management System
 - a. Allows professors to create, modify and delete existing courses (courses that they own)
 - b. This system, allows students to join an existing code using a course invite token which is given to a student by the professors.
- 3) Module System
 - a. This allows professors to upload lecture videos, notes and slides and also modify and delete the existing modules
 - b. Students can use this to download / view the modules
- 4) Assignment System
 - a. Allows professors to upload an assignment which can be of report type, programming type or a combination of both.
 - b. Students can use this to submit their assignments. They can also code and test their written code on the website itself.
- 5) Grading System
 - a. Allows professors to grade student assignments using a predefined rubric cube for marking.

- b. Students will be able to see their grades
- 6) Plagiarism Checker
 - a. Checks for plagiarism in the submitted reports and the code
- 7) Code Executor
 - a. Runs user code in the backend, maintains the virtual environment and the installed libraries.

Development & Implementation

Multiple languages will be used to develop this software. The entire project is divided into two parts, Frontend and Backend and both the parts will be using different programming languages.

Technologies Used

- 1) Backend
 - a. Python (Django framework)
 - b. PostgreSQL (Database)
- 2) Frontend
 - a. JavaScript (ReactJS)
 - b. Tailwind CSS (CSS Library)

Backend

It is much easier to develop a RestAPI using Python. Python also have thousands of libraries available which makes the development process very smooth. It is widely used in all fields of programming from Web development to Game development and even Machine Learning & AI making it one of the promising programming languages as compared to the others like NodeJS. It also has a huge community which makes it easier to get help if there are any issues during the development of project.

Django is a Python framework used to write a web server. It has its own Object-Relational Mapper (ORM) that makes interacting with the databases much easier to perform actions like creating / modifying tables and querying the database.

PostgreSQL is a relational database management system (RDBMS). It's highly scalable and quick making it one of the best options for both small-scale and large-scale projects. Compared to other databases like SQLite, PostgreSQL won't slow down when populated with huge amount of data.

Frontend

Similar to Django, ReactJS is one of the most used JavaScript libraries for developing web apps as per Stack Overflow survey of 2023. Out of 71,802 responses 40.58% respondent selected/preferred ReactJS. It was developed by Facebook and is being widely used in many websites like Twitter, Instagram, Facebook and many more. One of the main principles of ReactJS is reusability and it achieves that by using components which significantly reduces the number of lines of code that has to be written and allows reusing the same component in different section of the website without having to code it again.

In order to style the website and make it responsive, Tailwind CSS is will be used. Alternatively, Bootstrap could have been used as well but unlike bootstrap, Tailwind CSS makes the website look very neat and modern without having to significantly modify the UI.

Implementation

In order to implement this project, first the backend will be developed. Backend development would involve creating a database. It also involves coding a RestAPI to handle different HTTP requests like creating/updating/modifying/deleting different resources example, courses. Backend will have two major key components called, 'Code Runner' and 'Plagiarism Checker'.

Code Runner module accepts the code from the user and runs it in the backend, providing appropriate error and success messages. This module is also responsible for grading the programming assignment

of a student by granting marks to the student once they have reached certain stage of execution like code compiles, code runs without error, code passes all the tests (including the hidden tests)

Major drawback of the current Code Runner module is that it's blocking. Meaning, when one user is running their code then other users will have to wait for that user to finish before their code starts executing. This problem can be fixed by implementing Django Consumers (WebSocket) and Multi-threading. WebSocket enables two-way communication between the server and the client. This way, even when the code takes a long-time user will receive their code output immediate as its available without worrying about server timeout issue. Multi-threading will be used for running long operations like executing user code. This way multiple users can run their code simultaneously without having to wait for the others. One issue with this approach is that it will have a significant impact on the performance of the server specially when multiple users are running their code.

Another approach could be by running the code in another thread and save the output to a file. This way users won't have to wait for others code to complete executing. But since HTTP doesn't support two-way communication, user will have to send a request to the server multiple times in order to receive the executed code output which could again slow down the server when large number of users are waiting for their code's output.

Code runner will function by taking the code of the user and running it in the server using subprocess library. It first compiles the user code, if failed it will award scores to the users (if they earned any) and stops the executing while returning the error message. Similarly, this module then runs the code, tests the code against test cases perform the necessary steps based on the outcome. If code passes all the tests user will receive a success message and will be granted some marks as defined by the professor when creating the assignment. Below is the diagram for the Code Runner algorithm.

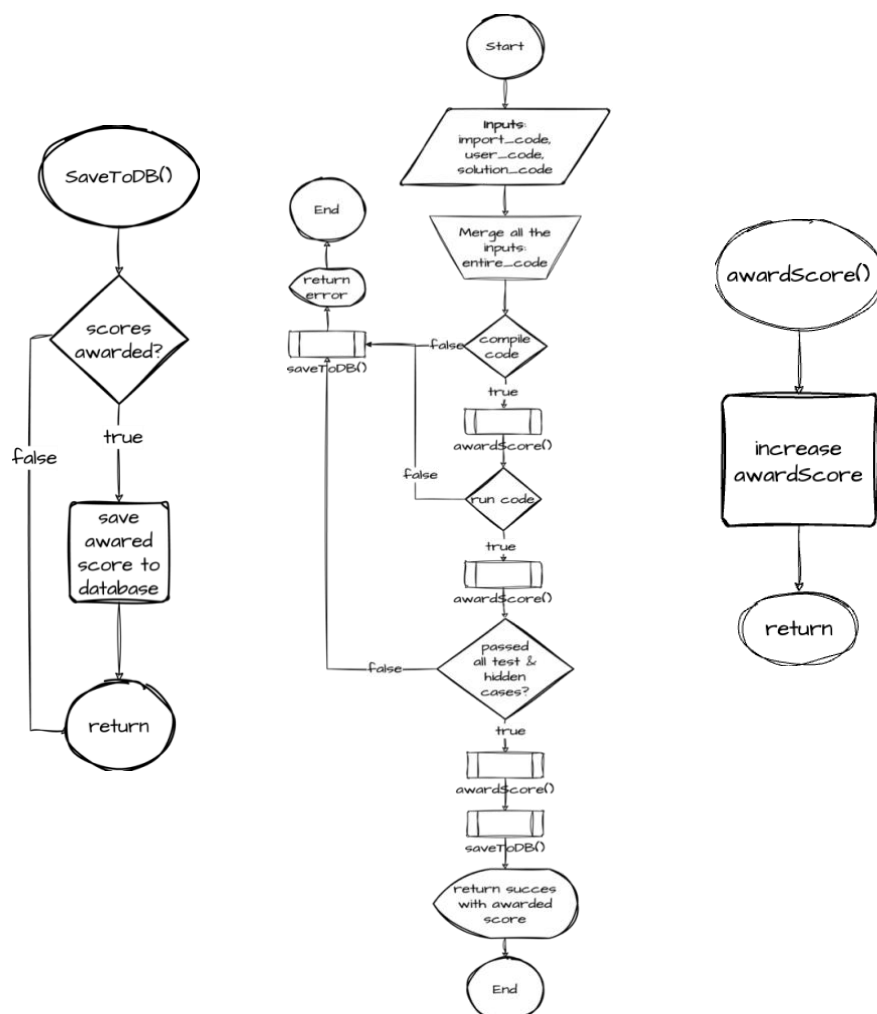


Figure 1: Code Runner Algorithm

To implement plagiarism checker, 3 methods were considered

- Using API
- String comparison
- Machine Learning

After carefully considering the limitations and advantages of all the three methods, String comparison was selected to implement the plagiarism checker. API wasn't selected because all of them were paid. Machine learning worked similarly to the String comparison method but required a large dataset with relevant articles in order to find plagiarism. There wasn't any dataset available that met the project's requirement and scrapping data from the required website would have taken a huge amount of time. Because of these limitations, String comparison method was selected.

String Comparison compares the strings between the two documents and returns a score between 0 to 1 where 1 means 100% identical and 0 means no plagiarism detected. In order to compared the strings, the submitted reports by the students are first converted to vectors using TF-IDF algorithm. These vectors are then compared using cosine similarity formula. Advantage of this method is that it's completely free and also doesn't rely on a dataset to function. Disadvantage of this method is that it only looks for plagiarism between the submitted reports and doesn't check the online websites for the plagiarism.

In this project this algorithm will work in the following manner. Let's say that we have 30 submissions for an assignment. All the 30 submissions will first be converted to vectors using TF-IDF algorithm after converting the submissions into vectors 1 submission is picked and is compared with the other submissions using Cosine similarity algorithm. This occurs for all the submissions one at a time. These scores are then showed to the professors while grading the assignments where they can take the appropriate action.

The database consists of 10 tables

Table Name	Description
User	Holds information like first & last name, email, username and password
Profile	Holds users' personal information like profile picture, user type (teacher/student) a reference to the user table and date & time of account creation.
Course	Used to store the details of the students who are in any course.
CourseDetail	Used to store the details of the students who are in any course.
ModuleGroup	Stores information about different groups for organizing modules like week 1, week 2 and so on. Each group belongs to a Course
Assignment	Stores information about an assignment (report portion of the assignment). It has a reference to AssignmentCode that stores the 'programming' portion of the assignment and to Course which indicate the course this assignment belongs to.
AssignmentCode	Used for storing programming part of an assignment
AssignmentSubmission	Whenever an assignment is submitted, the details are stored into this table. It has a reference to Assignment table and Profile Table (user who uploaded the assignment)
AssignmentRemark	Remarks submitted by the professor is stored in this table. It has a reference to AssignmentSubmission and to Profile table (user who graded the assignment)

Table 1: Different tables in the database and their uses

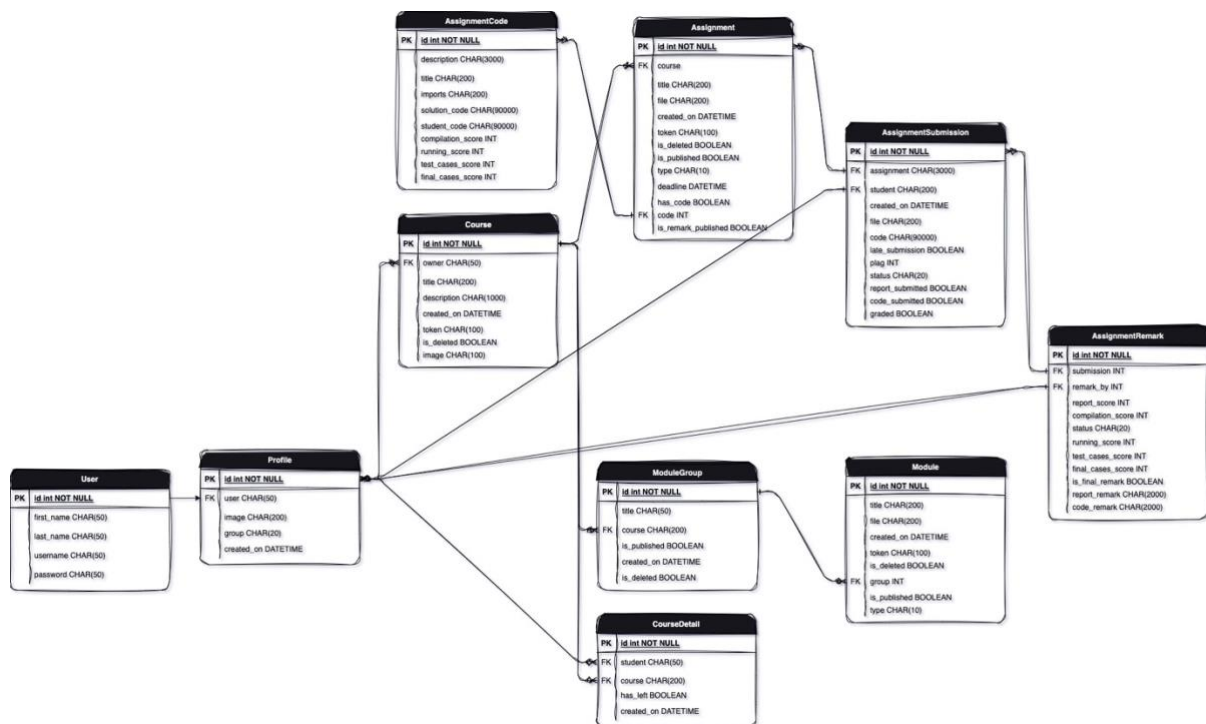


Figure 2: Database Diagram

Frontend development will start with designing the registration and login page, during the development frontend elements that are detected / identified as being repeated in multiple pages will then be turned into a component and will be reused throughout the website. One of the major components of the frontend would be FormGenerator. This project consists of large number of forms from login and registration forms to creating/modifying/deleting courses, assignments, modules, grading and much more. In order to make generating forms easier this component will be designed.

FormGenerator will take in a list of objects. Each object in the list defines an input type like email, text, number, radio box etc. It then loops through the list of objects and displays them on a webpage. This component would make the designing of forms very easy and will significantly reduce the number of lines of code that has to be written.

Use case diagram of the software based on different roles

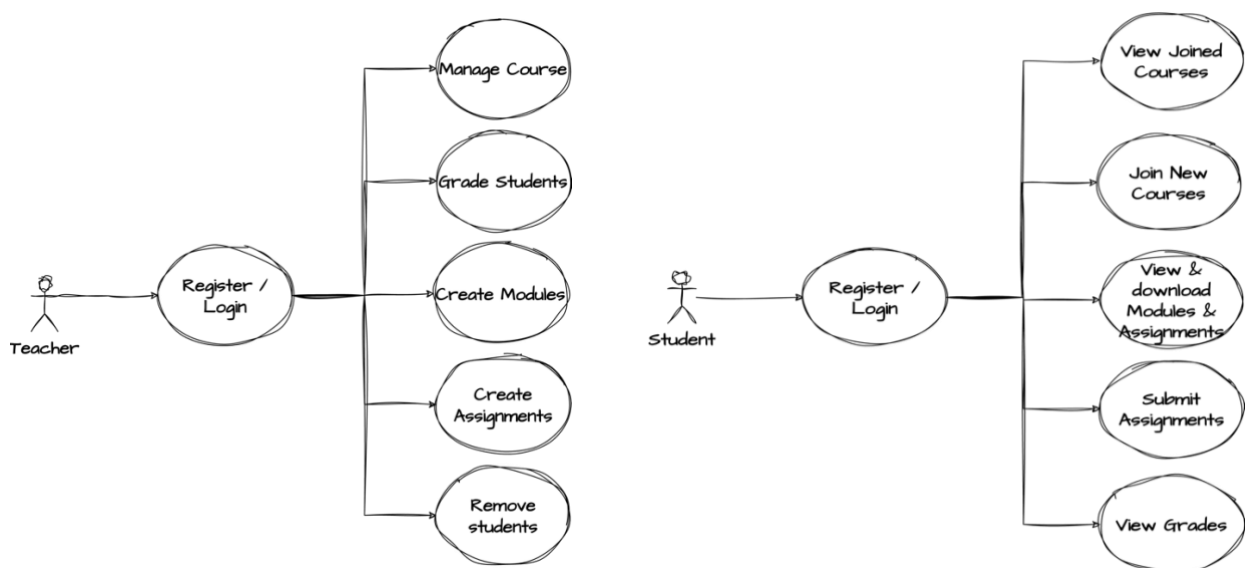


Figure 3: User case diagram for teacher & student

User Interface Mock-up

For designing the mock user interface, Canvas website was used as a reference. Couple mock screenshots of user interface are shown below. Final build will look similar but will have different color theme.

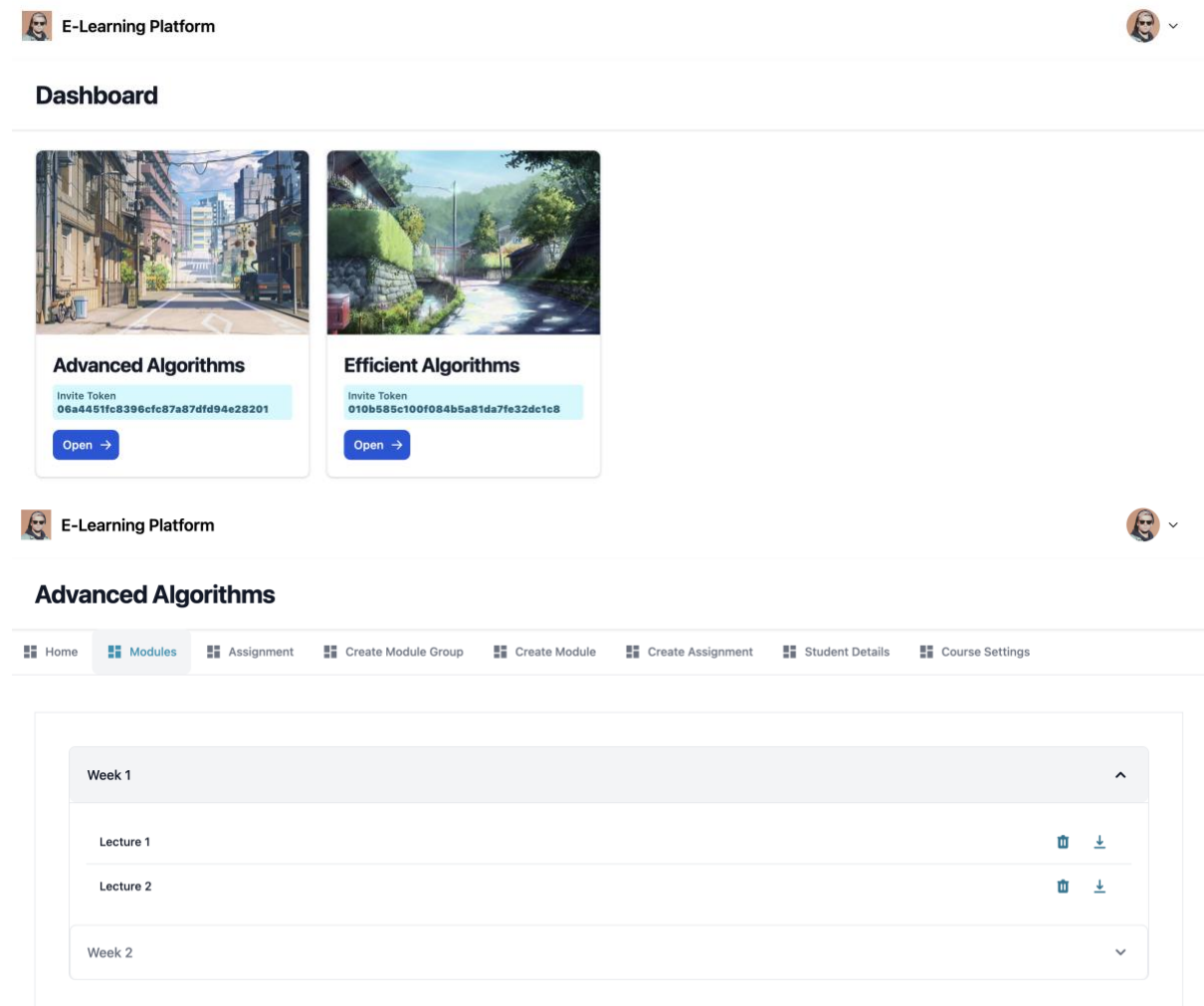


Figure 4: User interface

Data Sources

This project won't be using any real data, all the data that will be generated would be mock data to showcase the working of this project. The data that is being generated by the users are using the forms. Users' personal information that will be stored are as follow:

- Email Address: For login / registration / resetting password / communication by the professors.
- First Name & Last Name: For personalized experience and identification.
- Profile Picture: For personalization and can be deleted by the user.
- Courses: Courses that have been joined
- Grades: Users score in every assignment

Ethical Considerations

This software doesn't store any other personal information like address, phone number, gender, age etc. All the data obtained will be entered by the users themselves.

To protect the users and their data, all the confidential data like passwords will be stored after encrypting and hashing them making sure that even if the database gets hacked, hacker cannot access the user data like passwords.

While registering for an account, users would be informed that account once created cannot be deleted. As deleting the account could cause issues with the databases. Since there are a lot of references between the tables and most tables are connected to the Profile table, deleting a user account could affect the integrity of the database and could result in incorrect data being shown to the professors.

Testing & Evaluation

Postman will be used to perform Unit Testing and help from friends would be taken to test the software. All the modules in this software are pretty small and makes it very easy to implement Unit Testing.

In order to test a module, multiple HTTP requests will be sent to a specific module using Postman with different inputs to ensure it handles all incorrect inputs, duplicate data and invalid requests.

After performing Unit Testing, a friend will be asked to use the software. Without any interference from the developer their software usage would be noted and appropriate actions would be taken if the software breaks / crashes / unexpected behavior is observed during the test.

A module would be considered successful if it passes all the tests.

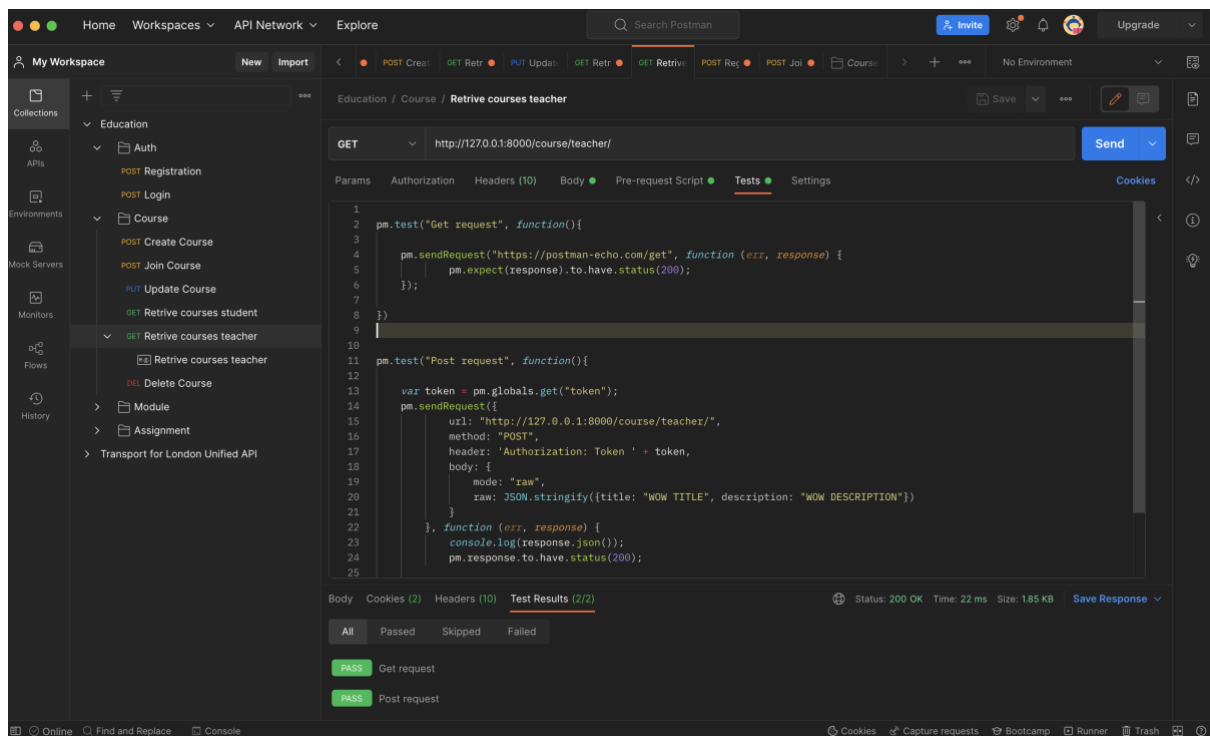


Figure 5: Unit testing in Postman

Project Plan

Plan for this project is as follow, finish one part of the project (backend / frontend) then test it and move on to the next item.

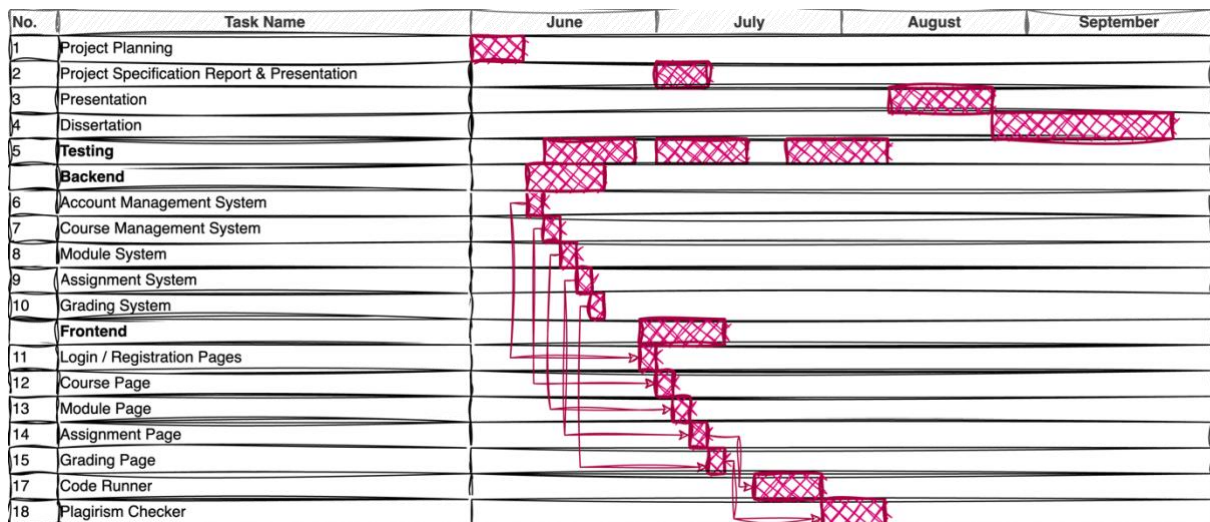


Figure 6: Gantt chart showing how the work is planned

Risks and Contingency Plans

Risks	Contingencies	Likelihood	Impact
Long waiting time when running code	In order to solve this issue, WebSocket and Multi-Threading can be used but due to increased complexity this wasn't implemented and was kept for future improvement.	Medium (depends on the number of people using the website at a time)	This will impact the performance of the website for all the user. This shouldn't have any significant impact on the performance if small number of users are using the software or if the program that is being executed is pretty small.
Server crash	When hosting the website, precautions should be taken to handle the situation when the server crashes.	Low	Website becomes inaccessible to the users unless restarted by the admin.
Database offline	Creating scripts to ensure if server ever goes offline, it gets restarted	Low	Website will be accessible but it will won't display any information from the backend.

Table 2: Risks and their impact

References

Teach Computing. (n.d.). *Pedagogy*. [online] Available at:

<https://teachcomputing.org/pedagogy>.

Draw.io (2023). *Flowchart Maker & Online Diagram Software*. [online] app.diagrams.net.

Available at: <https://app.diagrams.net/>.

Stack Overflow. (n.d.). *Stack Overflow Developer Survey 2023*. [online] Available at: <https://survey.stackoverflow.co/2023/#most-popular-technologies-webframe> [Accessed 3 Jul. 2023].

Cheang, B., Kurnia, A., Lim, A. and Oon, W.-C. (2003). On automated grading of programming assignments in an academic institution. *Computers & Education*, 41(2), pp.121–131. doi: [https://doi.org/10.1016/s0360-1315\(03\)00030-7](https://doi.org/10.1016/s0360-1315(03)00030-7)

Novak, J. & Cañas, A. (2008) The Theory Underlying Concept Maps and How to Construct and Use Them. Florida Institute for Human and Machine Cognition. Technical Report IHMC CmapTools 2006-01 Rev 01-2008.

Maton, K. (2013) Making semantic waves: A key to cumulative knowledge-building. *Linguistics and Education*. 24(1), 8–22.