# UNIVERSITY OF PISA



Artificial Intelligence and Data Engineering

# Distributed Systems and Middleware Technologies

 $FL console\ documentation$ 

Authors: Çolak F. Messina F. Nocella F.

Academic Year 2023/2024

# Contents

1	Intr	roduction and Project Overview 2
	1.1	Context and Project Objective
	1.2	Project Key Points
2	Ana	alysis 3
	2.1	Requirements
	2.2	Use Case Diagram
	2.3	Analysis Class Diagram
	2.4	Sequence Diagrams
3	Des	sign 8
	3.1	Introduction
	3.2	Software Architecture
	3.3	Database Design
	3.4	Message Handler
4	Imp	plementation 13
	4.1	Development Environment
	4.2	Main Modules
	4.3	Configuration
	4.4	Data Access
	4.5	Data Transfer
	4.6	Service
	4.7	User Interface
	4.8	Adopted Patterns and Techniques
5	Tes	ting 16
-	5.1	Structural Testing
	5.2	Functional Testing
6	Cor	nclusion 21

# Introduction and Project Overview

## Context and Project Objective

The goal of this project is to develop a system to manage Federated Learning (FL) experiments. FL is a decentralized machine learning approach where multiple devices collaborate to train a shared model while keeping their data locally. The project aims to provide a graphic interface with a web console to run FL experiments, enabling users to monitor their progress and analyze results. The system will use the Federated Learning Director (FL Director) to coordinate the execution of experiments among the devices. The FL Director is an Erlang node that manages the communication between the devices and the Web Console. The system will be designed to support concurrent execution of multiple FL experiments, real-time analytics, and flexible storage of experiment statistics. The project will adopt the Model-View-Controller (MVC) pattern to structure the Web Console, promoting separation of concerns and maintainability. The system will utilize DocumentDB for flexible storage of experiment statistics and its horizontal scalability. Concurrent execution of experiments will be implemented using Java threads and ExecutorService to optimize resource utilization. WebSocket communication will be established for real-time data exchange, enabling seamless interaction between the frontend and backend. The project will define message formats and outline the structure of Erlang nodes for efficient communication. The system will provide a user-friendly Web Console to initiate and manage FL experiments, as well as centralized access to experiment statistics for easy monitoring and analysis.

## **Project Key Points**

- Adopt the MVC (Model-View-Controller) pattern to structure the Web Console, promoting separation of concerns and maintainability.
- Utilize DocumentDB for flexible storage of experiment statistics and its horizontal scalability.
- Implement concurrent execution of experiments using Java threads and ExecutorService to optimize resource utilization.
- Establish WebSocket communication for real-time data exchange, enabling seamless interaction between the frontend and backend.
- Define message formats and outline the structure of Erlang nodes for efficient communication, ensuring reliability and scalability.
- System with a graphical interface and web console for executing FL experiments.
- Coordination of experiments through FL Director, an Erlang node.
- Real-time analytics and flexible storage of experiment statistics.
- Centralized access to experiment statistics for monitoring and analysis.

# **Analysis**

## Requirements

In this section, we outline the functional and non-functional requirements necessary for the successful implementation of the project.

### **Functional Requirements**

#### For Administrators:

- Administrators must be able to log in to the application.
- Administrators must be able to log out of the application securely.
- Administrators must be able to create new configurations for experiments.
- Administrators must be able to create experiments based on the configurations they have defined.
- Administrators must be able to initiate experiments and oversee their execution.
- Administrators must be able to view other administrators' experiments.
- Administrators must possess the authority to perform CRUD operations on configurations and experiments.

#### For Users:

- Users must be able to log in to the application.
- Users must be able to log out of the application securely.
- Users must be able to register for a new account within the application.
- Users must be able to search for experiments based on configuration and experiment names.
- Users must be able to monitor real-time progress of experiments.

## Non-Functional Requirements

- 1. **Performance:** The system must handle a large number of concurrent users without significant performance degradation. Response times for critical operations should be kept within acceptable limits.
- 2. Reliability: The system should be highly available with minimal downtime.
- 3. Security: User authentication and authorization mechanisms must be implemented.
- 4. **Scalability:** The system should scale horizontally to accommodate increasing user loads and data volumes.
- 5. Usability: The user interface must be intuitive and error messages must be informative.
- 6. **Maintainability:** The codebase must be well-organized and documentation must be comprehensive
- 7. **Compatibility:** The application must be compatible with a wide range of web browsers and devices. Integration with external systems must be seamless.

## Use Case Diagram

#### Actors

The actors who can interact with the web console system consist of the following:

- User: The user is the actor who can browse the system to view running and completed experiments and their results.
- **Admin:** The admin is the actor who can manage the system, including creating and deleting configurations and experiments, and viewing the results of experiments.

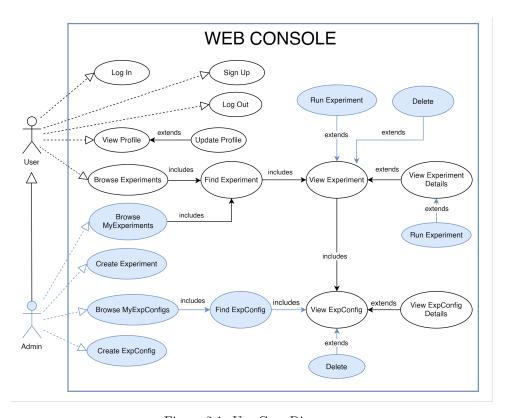


Figure 2.1: Use Case Diagram

### **Scenarios**

In the following tables, we present several key use cases related to the management and execution of experiments within the application. These use cases cover actions performed by different actors, including users and administrators, and describe the steps involved in each scenario, along with pre-conditions and post-conditions.

Table 2.1: Use Case: Find Experiment

Use Case	Find Experiment
Primary Actor	User, Admin
Secondary Actor	_
Description	Allows the actor to find a specific experiment
Pre-Conditions	Actor must be logged in
Main event steps	1. The actor navigates to the "Search" feature
	2. The actor enters the Experiment and/or the configuration name
	3. The system searches for the list of experiments in database for
	matching results
Post-Conditions	The actor views a list of experiments matching the
	search criteria if there are any
Correlated Use cases	
Alternative event steps	

Table 2.2: Use Case: Create Experiment

Use Case	Create Experiment	
Primary Actor	Admin	
Secondary Actor	_	
Description	Allows the admin to create a specific experiment	
Pre-Conditions	Actor must be logged in and has the admin privileges	
Main event steps	1. Admin selects the option to create a new experiment.	
	2. Admin fills in the name and configurations for the experiment.	
	3. Admin confirms the creation of the experiment.	
Post-Conditions	The experiment is successfully created.	
Correlated Use cases	Run Experiment	
Alternative event steps	_	

Table 2.3: Use Case: Run Experiment

Use Case	Run Experiment	
Primary Actor	Admin	
Secondary Actor	_	
Description	Allows the admin to start a specific experiment	
Pre-Conditions	Actor must be logged in and have admin privileges	
Main event steps	1. Admin selects the experiment and reaches the details page.	
	2. If the experiment has not started yet	
	2.1 when the start button is clicked	
	the system will start the experiment.	
Post-Conditions	3. The experiment statistics are shown on the experiment details page	
	and saved in the database.	
Correlated Use cases		
Alternative event steps	_	

# Analysis Class Diagram

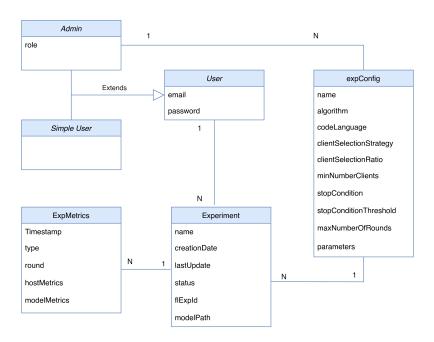


Figure 2.2: Class Diagram

# Sequence Diagrams

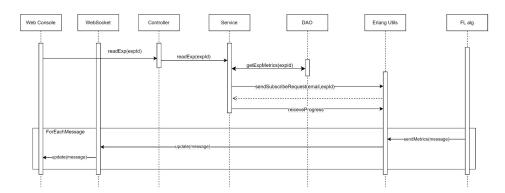


Figure 2.3: Sequence Diagram

# Design

## Introduction

This chapter aims to provide a detailed overview of the software architecture and database design of the project. It is essential for understanding the organization and structure of the system, as well as the design choices made to ensure the efficiency, scalability, and robustness of the software.

The design of the software architecture focuses on the organization and distribution of software components, defining roles, responsibilities, and interactions among them. Key architectural decisions guiding the project's development will be presented within this context.

Additionally, the database design will be examined, with particular attention to the decision to use a NoSQL database like MongoDB. This decision was motivated by the need to adapt to the specific requirements of the project, including flexible management of unstructured data and horizontal scalability.

## Software Architecture

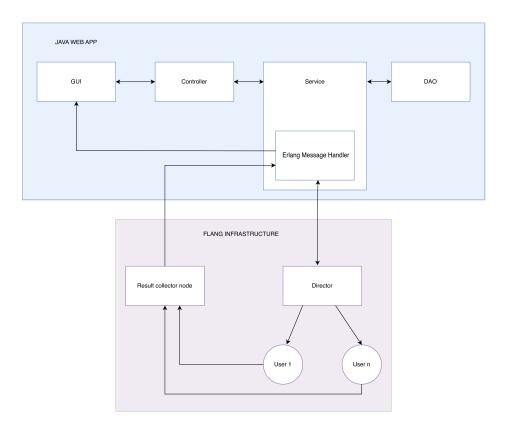


Figure 3.1: System Architecture

## Database Design

## MongoDB

#### Collections

ExpConfig document example:

```
{
    "_id": {
      "$oid": "6613f8b7aed2e52b006dea10"
    "name": "TestConfig",
    "algorithm": "fcmeans",
    "codeLanguage": "python",
    "clientSelectionStrategy": "probability",
    "clientSelectionRatio": 1,
    "minNumberClients": 2,
    "stopCondition": "max_number_rounds",
    "stopConditionThreshold": 5,
    "maxNumberOfRounds": 10,
    "parameters": {
      "targetFeature": "16",
      "lambdaFactor": "2",
      "numFeatures": "16",
      "seed": "10",
      "numClusters": "10"
    },
    "creationDate": {
      "$date": "2024-04-08T14:01:27.232Z"
 }
```

#### Experiment document example:

```
{
    "_id": {
      "$oid": "661c3d780bb4be3bd9b891b9"
    },
    "name": "ExpTest",
    "expConfig": {
      "_id": {
        "$oid": "6613f8b7aed2e52b006dea10"
     },
      "name": "TestConfig",
      "algorithm": "fcmeans"
   },
    "creationDate": {
      "$date": "2024-04-14T20:32:56.022Z"
    "status": "FINISHED",
    "flExpId": "\"d9d1bc7c-d733-4219-b4fb-16a3849db323\"",
    "modelPath": "\\FL_models\\exp_661c3d780bb4be3bd9b891b9.bin"
```

### ExperimentMetrics document example:

```
{
         "_id": {
           "$oid": "66144b5337a2fd7f67582f67"
         },
         "expId": "661c3d780bb4be3bd9b891b9",
         "type": "STRATEGY_SERVER_METRICS",
         "hostMetrics": {
           "cpuUsagePercentage": 5,
           "memoryUsagePercentage": 9.27
         },
         "modelMetrics": {
           "FRO": 845.7339394664009
         "timestamp": {
           "$date": "1970-01-20T19:43:26.034Z"
         "round": 1,
       }
User document example:
     "_id": {
      "$oid": "6611252030f96a50aebda458"
     "email": "admin@example.com",
     "password": "P@sswOrd",
     "creationDate": {
      "$date": "2024-04-06T10:34:08.669Z"
     },
     "role": "admin",
     "configurations": [
       "6613f8b7aed2e52b006dea10"
     ],
     "experiments": [{
         "_id": {
           "$oid": "661c3e800bb4be3bd9b891da"
         },
         "name": "ExpTest",
         "config": "TestConfig",
         "creationDate": {
           "$date": "2024-04-14T20:32:56.022Z"
         }
       }]
   }
```

## Message Handler

### Erlang for Message Passing

The message handler is implemented using the Erlang programming language. Erlang is a functional programming language designed for building scalable and fault-tolerant systems. It is particularly well-suited for building distributed systems, thanks to its lightweight processes and built-in support for message passing. In this project, it's utilized the Jinterface library, which allows to write Java code that can communicate with Erlang processes to send and receive messages, arriving from the FLang Infrastructure and vice versa.

## Message Structure

• Experiment Queued:

```
{
    "type": "experiment_queued",
    "timestamp": "1712206255"
}
```

• Worker Ready:

```
{
    "type": "worker_ready",
    "timestamp": "1712206257",
    "client_id": "1"
}
```

• Strategy Server Ready:

```
{
    "type": "strategy_server_ready",
    "timestamp": "1712206257"
}
```

• All Workers Ready:

```
{
    "type": "all_workers_ready",
    "timestamp": "1712206257"
}
```

• Start Round:

```
{
    "type": "start_round",
    "timestamp": "1712206257",
    "round": "1"
}
```

• Woker Metrics:

```
{
  "type": "woker_metrics",
  "timestamp": "1712206258",
  "round": "1"
  "client_id": "1"
  "hostMetrics": {
      "cpuUsagePercentage": 39.4,
      "memoryUsagePercentage": 52.69
}
  "modelMetrics": {
      "ARI": "0.10764575464353877"
}
}
```

• End Round:

```
{
  "type": "end_round",
  "timestamp": "1712206260",
  "round": "1"
}
```

• Strategy Server Metrics:

```
{
  "type": "strategy_server_metrics",
  "timestamp": "1712206264",
  "round": "2",
  "hostMetrics": {
      "cpuUsagePercentage": 39.4,
      "memoryUsagePercentage": 52.69
  }
  "modelMetrics": {
      "FRO": "0.13237183370436725"
  }
}
```

## Description of the Erlang Message Handler Module

The Erlang message handler module is a crucial component of the system responsible for managing incoming messages, processing them accordingly, and facilitating communication between different parts of the distributed system. It encapsulates the logic for handling various types of messages, such as error notifications, stop signals, and data updates, ensuring proper routing and processing. Additionally, the module provides interfaces for sending and receiving messages, abstracting the underlying communication mechanisms and enabling seamless integration with other system components. Its robust design and fault-tolerant features contribute to the overall reliability and performance of the distributed system.

# **Implementation**

## **Development Environment**

To be able to have efficient and successful implementation of Federated Learning Web Console Project, having a well-chosen development environment is one of the most important aspects. In this section, it is specified that the necessary tools, frameworks, and configuration requirements of the project.

- **Programming Language:** Java is used for creating a Web Application. Erlang is used for facilitation the development of middleware component and FL director is an Erlang node. So that effective communication between the Web Application and the FL director is provided.
- Frameworks: Spring is used for Java framework. It ensures to integrate dependencies for Web-Socket communication and MongoDB support. WebSocket is implemented to provide real-time communication between frontend and backend components.
- Database Management: MongoDB is chosen as a database to ensure storing data for experiment statistics and user information.
- Version Control: Git is used for version control. It is used to manage the source code of the project. GitHub is used to provide a collaborative development with its version control system. Efficient code management and collaboration is ensured by using repositories which is provided by the platform itself.
- Integrated Development Environment: IntelliJ IDEA is used as an IDE. It is a Java integrated development environment for developing computer software. It is developed by JetBrains. It is used to write, compile, and run the code. It also provides a user-friendly interface for developers.
- Build Automation: Maven is used for build automation. It is a build automation tool used primarily for Java projects. It is used to manage the project's build, reporting, and documentation from a central piece of information. Maven is used to control project dependencies and build configurations.
- **Testing:** Junit testing is used for testing Java code.

#### Main Modules

Implementation of the project is structured by diving the project into modules. Each module ensures specific requirements of the project architecture. The modules are:

- Configuration
- Controller
- DAO (Data Access Object)
- DTO (Data Transfer Object)
- Model
- Service
- Utils

## Configuration

Configuration classes of the Federated Learning Web Console project are created to provide responsibilities for configuring different parts of the application such as logging, execution, HTTP request handling, MVC setup and WebSocket communication. Efficient operation, security and scalability of the system can be ensured by those configuration properties.

#### **Data Access**

The data access classes are fulfilling the requirements of interacting with the database layers, providing data retrieval, storage, and manipulation. This module includes classes includes CRUD (create, read, update, delete) operations and query executions. With the Data Access classes such as ExpConfigDao, ExperimentDao, MetricsDao, UserDao the applications guarantee effective operations, management of experiments and tracking of the progress.

### **Data Transfer**

Data Transfer layer contains a ExpConfigSummary, ExperimentSummary and UserSummary classes to ensure the functionality of transferring data structure between different layers and components of the application. With the help of the DTO classes, related information will be able to be transferred between frontend, backend, and service layers. User information is transferred in a more standardized way for achieving better communication.

#### Service

Service module includes business logic and operations for ensuring the fully functional application. It provides data processing and interaction between different components. Service module includes:

- Cookie Service is for managing cookie operations such as cookie creation, retrieval, and deletion. The purpose of this service is ensuring session management and personalized user experience.
- Experiment Configuration Service is for implementing business logic for experiment configuration includes creation, deletion, retrieval and searching by some parameters.
- Experiment Service is for creating operations that are related with experiment like creation, running, deletion, retrieval and searching.
- User Service is implemented for ensuring business logic for user-based operations. Those operations include authentication of user, sign up, deletion of account, updating user information and retrieval of the user
- Message Service is implemented for managing communication with Erlang node for sending experiment configuration and monitoring the progress.
- Metrics Service is created to handle operations of retrieving experiment metrics from the database with the related experiment ID.

#### User Interface

User Interface module is responsible for providing a user-friendly interface for the users. This module makes application functionalities visible for the end-user. It includes the following components:

• Login and Sign Up Page: This page is for user authentication and registration. Users can log in to the system by providing their email and password. If the user does not have an account, they can sign up by providing their email, password, and description.

- User Dashboard: This page is for displaying the experiments to the user. Users can see experiments and their progress on this page.
- Experiment Page: This page is for displaying the details of the experiment. The page shows the details of the experiment and its progress on this page.
- Admin Dashboard: This page is for displaying all experiments. Admins can see all experiments and their progress on this page and also it provides creating new experiments for the admin.
- Profile Page: The profile page allows users to view and manage their account settings and profile information.

## Adopted Patterns and Techniques

During the implementation of the Federated Learning Web Console project, various patterns and techniques are adopted to ensure the efficiency, scalability, and maintainability of the application. These are some of the used patterns and techniques:

## Model-View-Controller (MVC) Pattern

The Federated Learning Web Console project is implemented by following the Model-View-Controller (MVC) pattern. This pattern is used to separate the application into three main components: Model, View, and Controller. The Model represents the data and business logic of the application, the View represents the presentation layer, and the Controller handles the user input and updates the model and view accordingly. This pattern ensures a clean separation of concerns and makes the application easier to maintain and extend.

#### WebSocket Communication

WebSocket communication is implemented to provide real-time communication between the frontend and backend. This allows the application to send and receive messages in real-time without the need for polling or long-polling. WebSocket communication is used to update the user interface with the latest data and provide a seamless user experience.

#### **Asynchronous Processing**

Asynchronous processing techniques like Java threads and ExecutorService are e used to manage concurrent execution of experiments and other tasks. This allows the application to handle multiple requests and tasks simultaneously and improve performance and scalability.

#### Message Passing Protocol

To achieve seamless communication between Erlang FL director and Java web application, a customized and well specified message passing protocol is defined. This protocol guarantees the reliable and well-defined exchange of messages and data.

## **Testing**

Testing methodologies are used to ensure about the reliability, correctness, functionality and quality of the Federated Learning Web Console. In this chapter, the testing methodologies used in the project are described. The testing methodologies are divided into two main categories: structural testing and functional testing.

## Structural Testing

Structural testing, also known as white-box testing is applied to the project to ensure that the implemented code is working as expected and evaluate the internal structure of the system. For primary structure testing JUnit testing is applied as a testing methodology.

### JUnit Testing

The JUnit testing performed on the project of FL Web Console. The JUnit testing is applied to various classes like DAOs and Services to check whether the implemented code is working as expected or not and specified requirement are hold by the methods. Some examples of the JUnit testing that performed on the classes:

#### **UserDAO**

The UserDAO class is an important component of the project which is responsible for interacting with the database to handle data related with users. With the help of the JUnit tests different scenarios are tested to ensure that the implemented code is working as expected and the requirements are fulfilled. This scenarios are including creating new user, deleting existing user, finding user by some criterias. These tests show the correctness of the CRUD operations of the UserDAO class. Below it can be seen an example of performing JUnit test for some methods in the UserDAO.

```
@Test ± winternox*
void findListOfConfigurationsByEmail() {

// 61ven
String userEmail = "admin@example.com";

// When
List<String> retrievedConfigurations = userRepository.findListOfConfigurationsByEmail(userEmail);

// Then
assertNotNull(retrievedConfigurations);
assertFalse(retrievedConfigurations.isEmpty());

System.out.println(retrievedConfigurations);

334
}
```

Figure 5.1: Testing the method of findListOfConfigurationsByEmail() in UserDAO class

```
▼ Vasification to a local section of their five 1800s

▼ Industrian principles of the local section of their five 1800s

■ Vindustrian principles of the local section of their five 1800s was the local section of the local section of their five 1800s was the local section of the lo
```

Figure 5.2: Testing result for the method of findListOfConfigurationsByEmail() in UserDAO class

With this JUnit test method, the findListOfConfigurationsByEmail() method of the UserDAO class is tested. The test is performed by finding all the related configuration that are belong to the user with that email. The test is successful and the expected result is returned as a list of configurations.

#### **ExperimentDAO**

Experiment DAO is another important class of the project which is responsible for interacting with the database to handle data related with experiments. JUnit tests are created to ensure that the experiment related functions are working as expected and the requirements are fulfilled. The tests are including creating new experiment, updating existing experiment, deleting existing experiment, finding experiment by some criterias. Below it can be seen an example test method for update an experiment.

Figure 5.3: Testing the method of update() in ExperimentDAO class

#### ConfigurationDAO

Test class for Configuration DAO is another example of JUnit test that is performed on the project. The Configuration DAO class is responsible for interacting with the database to handle storage and retrival of the system experiment configuration. Implementing JUnit for this class guarantees that system operates the data in an intended way for configuration class. Below there is an example test and test result for saveAndRetrieve() method of Configuration DAO class. The test is performed by saving a configuration and then retrieving it from the database. The test is successful and the expected result is returned.

Figure 5.4: Testing the method of saveAndRetrieve() in Configuration DAO class

Figure 5.5: Testing result for the method of saveAndRetrieve() in Configuration DAO class

## **Functional Testing**

Functional testing, also known as black-box testing is applied to the project to evaluate the system behaviour that needs to fulfill functional requirements. Functional testing helps to ensure that user expectations are provided in a right way. The functional testing is performed by creating test cases for the system.

Test cases are identified according to the functional requirements. It shows how the system should behave in different scenarios that are both normal and anormal. Those test cases are including user authentication, creating new configuration, creating new experiment, deleting experiment, finding experiment, finding configuration, deleting configuration, etc. Below there is a table that shows some examples of the test cases that are created for the project. As a result of the test cases, it can be said that the Federated Learning Web Console provides all the necessary functionalities and meets user expectations.

#### Test Cases

Table 5.1: Admin Test case

Id	Description	Input	Expected Out-	Output	Outcome
			put		
A_T_01	Admin Login	Email:	Login Success-	Redirected	Passed
A_1_01	Admin Login	admin@example.com	fully	to admin	
		Password: Adm1nP@ss		dashboard	
		(valid credentials)			
A_T_02	Admin Login	Email: in-	Error message	Unable to	Passed
A_1_02	2	valid@example.com	displayed	login	
		Password: invalid (in-			
		valid credentials)			
$A_{T_03}$	Creating New	Adding all necessary	Configuration	Configuration	Passed
	Configuration	values to the new FL	Created success-	Created	
		configuration form	fully	successfully	
A_T_04	Creating New	Entering all values ex-	Error of missing	Configuration	Passed
	Configuration	cept stop condition	value message	is not cre-	
	2		displayed	ated	
$A_T_05$	Creating New	Name is written and FL	Experiment Cre-	Experiment	Passed
	Experiment	configuration is selected	ated successfully	Created successfully	
A_T_06	Starting an	Press Start Experiment	Experiment	Experiment	Passed
	Experiment	button	starts	starts	
A_T_07	Search	Write name with exist-	Show the list of	List of con-	Passed
A_1_01	Configuration	ing configuration name	the configuration	figurations	
	by name		with that name	with that	
				name	
A_T_08	Search	Write name with exist-	Show the list of	List of ex-	Passed
Λ_1_00	experiment by	ing experiment name	the experiments	periments	
	name		with that name	with that	
				name	

Table 5.2: User Test case

Id	Description	Input	Expected Out-	Output	Outcome
			put		
U_T_01	User Login	Email:	Login Success-	Redirected	Passed
0_1_01		firstTest@example.com	fully	to user	
		Password: P@ssw0rd		dashboard	
		(valid credentials)			
U_T_02	User Login 2	Email:	Error message	Unable to	Passed
0_1_02	User Login 2	wrong@example.com	displayed	login	
		Password: invalid (in-			
		valid credentials)			
U_T_03	Haan Cianun	Email:	Sign up success-	Redirected	Passed
0_1_03	User Signup	new@example.com	fully	to user	
				dashboard	
		Password: P@ssw0rd			
		(valid input)			
U_T_04	User Signup 2	Email:	Error message	Unable to	Passed
0_1_04		new@example.com	displayed	Signup	
		Password: invalid (in-			
		valid password)			
U_T_05	Search	Write name with exist-	Show the list of	List of con-	Passed
0_1_00	Configuration	ing configuration name	the configuration	figurations	
	by name		with that name	with that	
				name	
U_T_06	Search	Write name with exist-	Show the list of	List of ex-	Passed
0_1_00	experiment by	ing experiment name	the experiments	periments	
	name		with that name	with that	
				name	

## Conclusion

In this project Federated Learning Web Console is introduced which is a decentralized platform for managing FL experiments with using Java and Erlang as primary programming languages, Spring as a Java framework and MongoDB as a database. The primary goal of this project is to implement and provide robust system that can coordinate FL experiment with machine learning approach and being in a collaboration with multiple devices while data locality is retained.

As a main project architectural structure, the project follows the MVC pattern. Various and functional models are implemented in the project such as Configuration, Data Access Object, Services and User Interface with their specific roles and functionalities for serving the application. MongoDB is chosen as a document DB database to have scalable and adaptive database.

With the help of this architecture FL Web Console project ensures a scalable design and having seamless functionality between different components. The user-centric user interface includes essential pages to provide fully functional experience for the user. Those pages are login/signup, user/admin dashboard, profile page and experiment details.