**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN - ĐHQG TPHCM**

**KHOA CÔNG NGHỆ PHẦN MỀM**

**ĐỒ ÁN MÔN HỌC CHUYÊN NGÀNH**

ĐỀ TÀI: HỆ THỐNG KHUYẾN NGHỊ VIỆC LÀM

*GVHD:*

*TS HUỲNH NGỌC TÍN*

*SINH VIÊN THỰC HIỆN:*

*12520248 - Trần Minh Luận*

*12520492 – Nguyễn Thanh Anh Tuyên*

MỤC LỤC

[**Chương 1** **GIỚI THIỆU HỆ THỐNG** 1](#_Toc456070407)

[1.1 Vấn đề 1](#_Toc456070408)

[1.2 Phát biểu bài toán 1](#_Toc456070409)

[**Chương 2** **XÁC ĐỊNH VÀ PHÂN TÍCH YÊU CẦU** 2](#_Toc456070410)

[**2.1** **Danh sách các yêu cầu** 2](#_Toc456070411)

[**2.1.1** **Upload dataset** 2](#_Toc456070412)

[**2.1.2** **Xem thống kê dữ liệu đầu vào:** 2](#_Toc456070413)

[**2.1.3** **Cho phép chạy thực nghiệm các thuật toán khuyến nghị.** 4](#_Toc456070414)

[**2.1.4** **So sánh các thuật toán khuyến nghị** 4](#_Toc456070415)

[**2.2** **PHÂN TÍCH YÊU CẦU** 5](#_Toc456070416)

[2.1 Biểu đồ phân rã chức năng hệ thống 5](#_Toc456070417)

[2.2 Sơ đồ tổng thể hệ thống 5](#_Toc456070418)

[2.3 Sơ đồ các thành phần xử lý của hệ thống 6](#_Toc456070419)

[**Chương 3** **THIẾT KẾ HỆ THỐNG** 9](#_Toc456070420)

[**2.1** **KIẾN TRÚC HỆ THỐNG** 9](#_Toc456070421)

[2.1.1. Kiến trúc tổng quát 9](#_Toc456070422)

[2.1.2. Mô tả các thành phần trong kiến trúc 10](#_Toc456070423)

[2.1.3. Kiến trúc triển khai 10](#_Toc456070424)

[**2.2** **THIẾT KẾ DỮ LIỆU** 11](#_Toc456070425)

[2.2.1. System DB 11](#_Toc456070426)

[2.2.2. User dataset 13](#_Toc456070427)

[**2.3** **THIẾT KẾ GIAO DIỆN** 13](#_Toc456070428)

[2.3.1. Giao diện đăng nhập 13](#_Toc456070429)

[2.3.2. Trang upload dataset 15](#_Toc456070430)

[2.3.3. Giao diện đăng ký tài khoản 17](#_Toc456070431)

[2.3.4. Giao diện thống kê dữ liệu 18](#_Toc456070432)

[2.3.5. Giao diện chạy thực nghiệm các thuật thuật toán 18](#_Toc456070433)

[2.3.6. Giao diện xem kết quả thực nghiệm 20](#_Toc456070434)

[2.3.7. Giao diện thống kê dữ liệu 21](#_Toc456070435)

[2.3.8. Giao diện quên mật khẩu 21](#_Toc456070436)

[**2.4** **THIẾT KẾ XỬ LÝ** 22](#_Toc456070437)

[**Chương 4** **LẬP TRÌNH VÀ KIỂM THỬ** 28](#_Toc456070438)

[**KẾT LUẬN** 29](#_Toc456070439)

# **GIỚI THIỆU HỆ THỐNG**

## **Vấn đề**

Có rất nhiều hướng giải bài toán khuyến nghị việc làm như sử dụng phương pháp collaborative filtering, sử dụng phương pháp content-based và rất nhiều hướng giải khác. Tuy nhiên làm sao để tìm được một hướng giải tốt là một vấn đề khó khăn. Vì vậy, hệ thống này được xây dựng nhằm cung cấp các đánh giá, so sánh kết quả thực nghiệm của những thuật toán khuyến nghị khác nhau. Từ đó chúng ta có thể có những quyết định phù hợp.

## **Phát biểu bài toán**

Hệ thống “Đánh giá các thuật toán khuyến nghị” (Recommender algorithm evaluation) là hệ thống cho phép người dùng đưa vào tập dữ liệu thực nghiệm. Sau đó tiến hành phân tích, thống kê dữ liệu đầu vào và chạy thực nghiệm các thuật toán khuyến nghị.

Hệ thống này phục vụ cho việc giải quyết bài toán khuyến nghị việc làm bằng cách cung cấp các đánh giá, so sánh các kết quả thực nghiệm. Từ đó tìm ra giải pháp hiệu quả để giải bài toán khuyến nghị việc làm.

* Đầu vào của hệ thống: Tập dữ liệu dataset.
* Đầu ra của hệ thống:
  + Biểu đồ thống kê dữ liệu.
  + Kết quả chạy của các thuật toán khuyến nghị dựa trên tập dataset đầu vào.
  + Các đánh giá so sánh kết quả thực nghiệm.

# **XÁC ĐỊNH VÀ PHÂN TÍCH YÊU CẦU**

## **Danh sách các yêu cầu**

Bảng danh sách các yêu cầu:

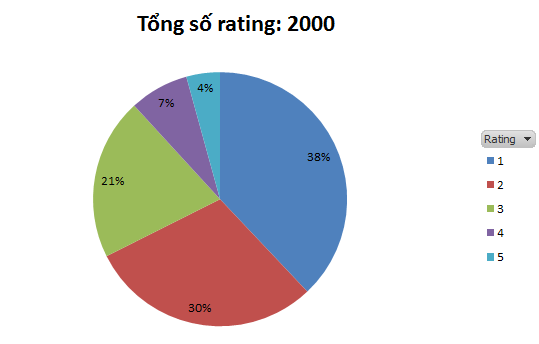
|  |  |  |
| --- | --- | --- |
| Mã Yêu cầu | Tên yêu cầu | Ghi chú |
| 1.1 | Upload dataset |  |
| 1.2 | Xem thống kê dữ liệu |  |
| 1.3 | Chạy thực nghiệm các thuật toán khuyến nghị |  |
| 1.4 | So sánh các thuật toán khuyến nghị |  |

### Upload dataset

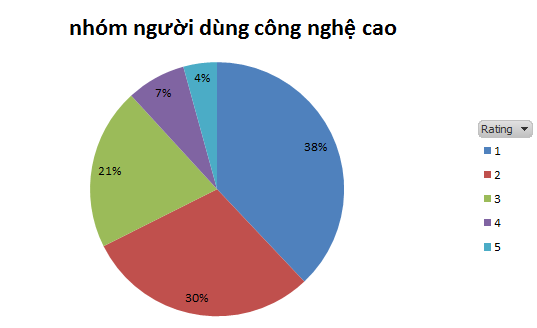
* Để bắt đầu sử dụng hệ thống. Người dùng phải cung cấp tên dataset cho hệ thống. Dataset được upload bao gồm 03 file.
  + File Cv.txt: Chứa dữ liệu về người dùng bao gồm các thông tin về cv người dùng.
  + File Job.txt: Chứa dữ liệu về việc làm sẽ được khuyến nghị.
  + Score.txt: Chứa dữ liệu về sở thích của người dùng trong file Cv.txt đối với những việc làm trong file Job.txt
* Sau khi chọn file người dùng chọn chức năng upload để bắt đầu upload dữ liệu. Hệ thống sẽ lưu trữ dataset vào thư mục có dạng: mã người dùng/tên dataset và thông báo kết quả thực hiện cho người dùng.

### Xem thống kê dữ liệu đầu vào

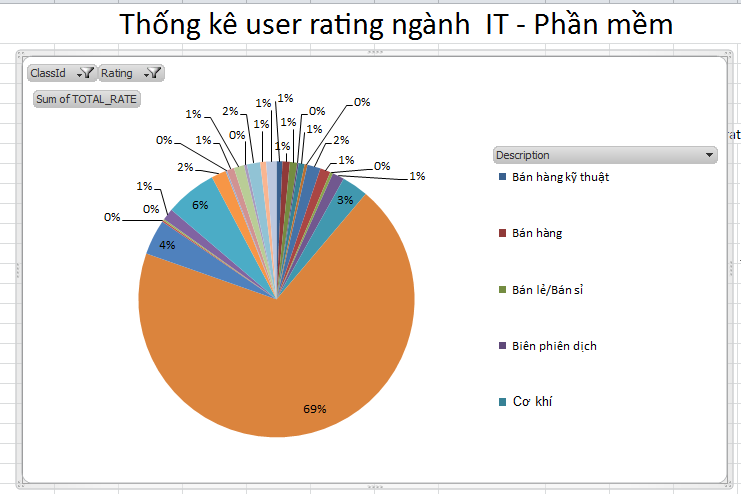
* Người dùng sẽ chọn tập dữ liệu cần phân tích để đưa vào module phân tích dữ liệu. Nếu chưa có thì người dùng có thể upload lên hệ thống. Sau khi người dùng chọn dữ liệu đầu vào. Người dùng chọn chức năng phân tích để bắt đầu phân tích dữ liệu. Sau khi hệ thống phân tích xong sẽ hiển thị ra màn hình cho người dùng.
* Các thống kê cần thực hiện.
  + Dựa vào file score.txt tiến hành thống kê tỉ lệ rating theo từng mức từ 1 – 5. Biểu đồ được vẽ là biểu đồ hình quạt có dạng bên dưới.



* + Thống kê người dùng theo từng ngành nghề(NND). Với mỗi NND thống kê tổng số rating mà thuộc NND đó. Bao gồm tỉ lệ % rating 1 điểm, tỉ lệ % rating 2 điểm. Tỉ lệ % rating 3 điểm. Tỉ lệ % rating 4 điểm và tỉ lệ % rating 5 điểm. Biểu đồ được vẽ như bên dưới.



Với mỗi nhóm rating như ở trên. Khi người dùng chọn một nhóm rating có giá trị x thì sẽ xuất thống kê tỉ lệ % các ngành mà NND đã cho x điểm. Biểu đồ được vẽ như hình bên dưới.



### Cho phép chạy thực nghiệm các thuật toán khuyến nghị.

* Người dùng tạo chọn dataset cần thực nghiệm, chọn thuật toán cần chạy. Nếu tập dữ liệu chưa có thì người dùng có thể upload lên hệ thống. Sau đó người dùng bấm chạy. Hệ thống sẽ thực hiện chạy thuật toán khuyến nghị. Sau khi chạy xong. Hệ thống sẽ thông báo cho người dùng kết quả thực nghiệm và lưu trữ lại vào file output.txt.
* Cấu trúc file output gồm 3 cột: UserId, JobId, Score. Các cột ngăn cách nhau bằng dấu kí tự tab.

### So sánh các thuật toán khuyến nghị

* Người dùng sẽ nhập những file output.txt sinh ra khi chạy thực nghiệm vào hệ thống. Hệ thống sẽ thực hiện tính toán và trả về kết quả kết quả so sánh các thuật toán khuyến nghị cho người dùng.

## **Phân tích yêu cầu**

### Biểu đồ phân rã chức năng hệ thống

Recomender evaluation system

Upload dataset

View dataset Statistic

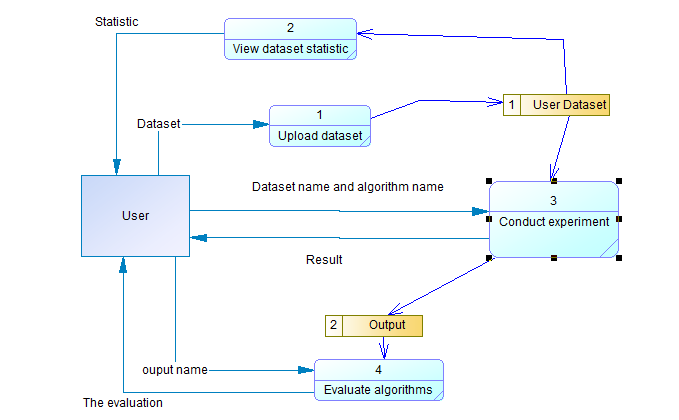
Conduct experiment

Evaluate algorithms

### Sơ đồ tổng thể hệ thống



### Sơ đồ các thành phần xử lý của hệ thống

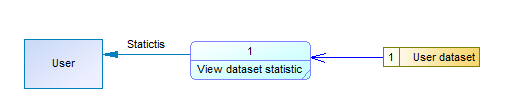


1. Mô tả process upload dataset.



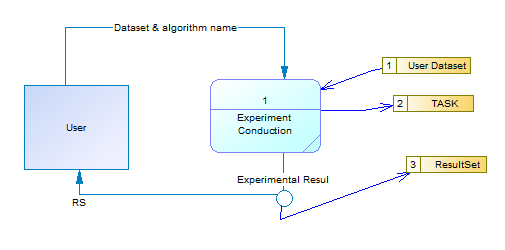
* Mô tả: người dùng đưa dataset gồm 3 file job.txt, Cv.txt và Score.txt vào hệ thống. Hệ thống kiểm tra và lưu dataset lại.
* Các biểu mẫu liên quan:
  + File Job.txt: Định dạng file text, gồm các cột JobId, JobName, JobLocation, Salary, Category, Requirement, Tags, Description
  + File Cv.txt: Định dạng file text gồm các cột: AccountId, ResumeId, User name, Cv title, Address, ExpectedSalary, Category, Education, Language, Skill, CareerObjective.
  + File Score.txt: Định dạng file text gồm các cột: AccountId, JobId, Score

1. Mô tả process view dataset statistic.



* Mô tả: từ dataset mà người dùng đã upload. Hệ thống sẽ phân tích. Kết quả phân tích sẽ được lưu lại vào kho Data Statistic và trực quan hóa dữ liệu cho người dùng.

1. Mô tả process conduct experiment.



* Mô tả: Người dùng nhập tên của thuật toán cần chạy thực nghiệm và tên của dataset. Hệ thống sẽ lấy dataset tương ứng từ kho User Dataset và lưu trữ lại các thông tin vào kho . Sau đó chạy thực nghiệm dựa thuật toán mà người dùng chọn. Sau khi chạy xong thì lưu trữ vào kho ResultSet và lưu lại trạng thái vào kho TASK là đã chạy xong.

1. Mô tả process evaluate algorithms



* Mô tả: người dùng đưa vào tên các kết quả thực nghiệm cần so sánh. Hệ thống sẽ lấy từ kho result set các kết quả thực nghiệm tương ứng. Sau đó hệ thống sẽ so sánh và trực quan hóa kết quả cho người dùng.

# **THIẾT KẾ HỆ THỐNG**

## **Kiến trúc hệ thống**

### Kiến trúc tổng quát

**SYSTEM DB**

**User DataSet**

**Spring MVC**

**Apache Mahout**

**Apache Hadoop**

**JobRec** **Task Management**

**Dataset Analyzation**

**Experimental conduction & comparision**

**Web module**

**Core**

**System DB**

**Libraries**

### Mô tả các thành phần trong kiến trúc

* + Web module
* JobRec Task Management: Mô-đun này cung cấp cho người dùng chức năng quản lý dataset, xem thống kê dataset, quản lý các tiến trình thực thi các thuật toán khuyến nghị, xem kết quả và các đánh giá kết quả khuyến nghị.
  + Core
* Dataset analyzation: Đây là mô-đun phân tích dataset của người dùng.
* Experimental conduction & comparision: chạy thực nghiệm các thuật toán khuyến nghị và so sánh kết quả thực hiện.
  + Database
* System DB: database dùng để tổ chức lưu trữ thông tin người dùng, các tác vụ thực hiện của người dùng và các thông tin thiết lập chạy thuật toán.
* User dataset: hệ thống lưu trữ files dataset do người dùng upload cùng với đó là các file kết quả khuyến nghị, kết quả so sánh và kết quả thống kê.
  + Libraries
* Spring MVC: dùng hỗ trợ xây dựng website tương tác người dùng.
* Mahout: sử dụng các thư viện recommend có sẵn trong Mahout để thực hiện khuyến nghị.
* Hadoop: sử dụng để lưu trữ dữ liệu lớn, hỗ trợ cho việc phát triển trong tương lai.

### Kiến trúc triển khai

Hệ thống được thiết kế dựa vào mô hình Client-Server. Bởi vì tiến trình thực hiện các thuật toán khuyến nghị, tiến trình phân tích dữ liệu là các tiến trình tốn nhiều thời gian nên không phù hợp cho các ứng dụng web. Chính vì lí do này hệ thống được chia thành nhiều project con. Và các project con này sẽ gọi lẫn nhau và trao đổi thông tin cho nhau. Gồm các sub-project như sau.

Client computers

JobRec Task Management

Analyzation

Experimental

Conduction &

comparision

Database

## **Thiết kế dữ liệu**

### System DB

Database này gồm có 2 bảng sau:

* USER (UserId, UserName, Email, Password)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| STT | Tên cột | Mô tả | Kiểu dữ liệu | Miền giá trị | Ghi chú |
| 1 | UserId | Mã người dùng | Integer | 11 | Not null, primary key |
| 2 | UserName | Tên người dùng | Nvarchar | 50 | Not null |
| 3 | Email | Email đăng nhập người dùng | Varchar | 50 | Not null |
| 4 | password | Mật khẩu đăng nhập | Varchar | 50 | Not null |

* Task (TaskId, UserId, TaskName, TimeCreate, Status, Algorithm, InputFolder, OutputFile)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| STT | Tên cột | Mô tả | Kiểu dữ liệu | Miền giá trị | Ghi chú |
| 1 | TaskId | Mã của task | Integer | 11 | Not null, primary key |
| 2 | UserId | Mã người dùng | Integer | 11 | Not null, foreign key đến User |
| 3 | TaskName | Tên task | Nvarchar | 100 | Not null |
| 4 | TimeCreate | Thời gian tạo task | Datetime |  | Not null |
| 5 | Status | Trạng thái task | Nvarchar | 10 | Not null |
| 6 | Algorithm | Tên thuật toán | Varchar | 20 | Not null |
| 7 | InputFolder | Đường dẫn đến thư mục chứa file input | Varchar | 100 | Not null |
| 8 | OutputFile | Đường dẫn file output | Varchar | 100 | Not null |

### User dataset

* Dữ liệu đầu vào để chạy các thuật toán gồm 3 file Job.txt, CV.txt và Score.txt. Quy ước trong các file .txt mỗi cột cách nhau bằng một ký tự tab.
* File Job.txt gồm các cột sau: JobId, JobName, Location, Salary, Category, Requirement, Tag, Description.
* File CV.txt gồm các cột sau: UserId, CVId, UserName, CVName, UserAddress, ExpectedSalary, Category, Language, Education, Skill, CareerObjective.
* File Score.txt gồm các cột sau: UserId, JobId, Score.
* Dữ liệu đầu ra

File output.txt có cấu trúc gồm 3 cột: UserId, JobId, Score.

## **Thiết kế giao diện**

### Giao diện đăng nhập

* Giao diện:



* Mô tả giao diện:

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thuộc tính | Kiểu | Chức năng |
| 1 | Email | HTML Form Input(text) | Chứa địa chỉ email đăng nhập. |
| 2 | Mật khẩu | HTML form Input(password) | Chứa password của người dùng. |
| 3 | Đăng nhập | Html form Input(Đăng nhập) | Gửi dữ liệu lên trang xử lý |
| 4 | Lấy lại mật khẩu | HTML link | Chuyển sang trang lấy lại mật khẩu. |
| 5 | Đăng ký | HTML link | Chuyển sang trang đăng ký |

* Mô tả biến cố:

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Biến cố | Xử lý | Mã số xử lý |
| 1 | Người dùng Bấm chọn nút đăng nhập | Kiểm tra email và mật khẩu đã được đăng ký hay chưa. Nếu chưa tồn tại trong hệ thống thì thông báo và yêu cầu nhập lại. Nếu đã tồn tại thì chuyển sang màn hình trang chủ. | XL1 |
| 2 | Người dùng bấm link lấy lại mật khẩu | Chuyển sang trang lấy lại mật khẩu | XL2 |
| 3 | Người dùng bấm link đăng ký | Chuyển sang trang đăng ký | XL3 |
| 4 | Load trang | Kiểm tra trạng thái đăng nhập của user. Nếu user đã đăng nhập thì chuyển về trang chủ. Ngược lại thì hiển thị trang đăng nhập | XL4 |

### Trang upload dataset

* Giao diện:



* Mô tả giao diện

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thuộc tính | Kiểu | Chức năng |
| 1 | Danh sách dataset | HTML TABLE | Thông tin danh sách các dataset đã upload |
| 2 | Xem thống kê | HTML link | Hiển thị đường dẫn tới trang thống kê dataset |
| 3 | Tên dataset | HTML form input(text) | Chứa tên dataset do người dùng nhập |
| 4 | Chọn File score.txt | HTML form upload field | Chứa đường dẫn tới file score.txt do người dùng chọn |
| 5 | Chọn file Cv.txt | HTML form upload field | Chứa đường dẫn đến file cv.txt do người dùng chọn |
| 6 | Chọn file Job.txt | HTML form upload | Chứa đường dẫn đến file job.txt do người dùng chọn |
| 7 | Upload | HTML form input(submit) | Nút dùng để kích hoạt chức năng upload. |

* Mô tả biến cố

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Biến cố | Xử lý | Mã xử lý |
| 1 | Load trang | * Kiểm tra trạng thái đăng nhập của người dùng. Nếu chưa đăng nhập thì chuyển đến trang đăng nhập hệ thống. * Lấy danh sách dataset mà người dùng đã upload hiển thị lên bảng 1. | XL4,  XL5 |
| 2 | Người dùng bấm xem thống kê | Chuyển đến trang thống kê dataset của dataset tương ứng | XL6 |
| 3 | Người dùng bấm upload | Kiểm tra thông tin người dùng đã nhập. Lưu trữ lại dataset và thông báo kết quả thực hiện cho người dùng. | XL7 |

### Giao diện đăng ký tài khoản

* Giao diện:



* Mô tả giao diện

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thuộc tính | Kiểu | Mô tả |
| 1 | Tên tài khoản | HTML form input(text) | Chứa tên tài khoản do người dùng nhập |
| 2 | Email | HTML form input(text) | Chứa email do người dùng nhập |
| 3 | Mật khẩu | HTML form input(password) | Chứa mật khẩu do người dùng nhập |
| 4 | Nhập lại mật khẩu | HTML form input(password) | Chứa mật khẩu do người dùng nhập |
| 5 | Đăng ký | HTML form input(submit) | Kích hoạt chức năng đăng ký tài khoản |
| 6 | Đăng nhập | HTML link | Quay về trang đăng nhập |

* Mô tả biến cố

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Biến cố | Xử lý | Mã xử lý |
| 1 | Load trang | Kiểm tra cookie trên máy người dùng. Nếu người dùng đã đăng nhập thì yêu cầu đăng xuất trước khi tiếp tục thực hiện | XL8 |
| 2 | Người dùng chọn đăng ký tài khoản | Kiểm tra các thông tin mà người dùng nhập vào hệ thống. Nếu còn thiếu hoặc sai thì yêu cầu người dùng nhập lại. Nếu thông tin hợp lệ thì chuyển thông tin về trang xử lý. | XL9 |
| 3 | Người dùng chọn đăng nhập | Quay trở về trang đăng nhập hệ thống | XL10 |

### Giao diện thống kê dữ liệu

### Giao diện chạy thực nghiệm các thuật thuật toán

* Giao diện:



* Mô tả giao diện:

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thuộc tính | Kiểu | Mô tả |
| 1 | Danh sách task | HTML table | Hiển thị danh sách các tác vụ đã được tạo và trạng thái. |
| 2 | Task | HTML link | Hiển thị liên kết đến trang kết quả thực nghiệm. |
| 3 | Tên task | HTML form input(text) | Chứa tên task người dùng nhập |
| 4 | Chọn thuật toán | HTML select | Dùng để lựa chọn thuật toán cần thực nghiệm |
| 5 | Chọn dataset | HTML select | Dùng để lựa chọn dataset cần thực nghiệm. |

* Các biến cố và xử lý:

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Biến cố | Xử lý | Mã xử lý |
| 1 | Load trang | * Kiểm tra trạng thái đăng nhập của user. Nếu chưa đăng nhập thì yêu cầu đăng nhập. * Load danh sách các tác vụ mà người dùng đang đăng nhập hệ thống đã tạo. Tô màu xanh với các tác vụ đã chạy thành công. Tô màu vàng với các tác vụ đang chạy. Tô màu đỏ với các tác vụ bị lỗi. | XL11 |
| 2 | Người dùng bấm vào tên tác vụ. | Nếu tác vụ thực hiện thành công. Chuyển người dùng đến trang xem kết quả. Ngược lại không làm gì cả. | XL12 |
| 3 | Người dùng chọn tạo tác vụ | Kiểm tra tính hợp lệ dữ liệu mà người dùng nhập. Nếu có sai sót hoặc thiếu thì yêu cầu người dùng nhập lại. Nếu hợp lệ thì lưu trữ thông tin và bảng task và tiến hành thực nghiệm. | XL13 |

### Giao diện xem kết quả thực nghiệm

* Giao diện:



* Mô tả giao diện

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên thuộc tính | Kiểu | Mô tả |
| 1 | Resulte | HTML link | Hiển thị link download file result |
| 2 | Bảng kết quả | HTML table | Hiển thị kết quả thực nghiệm |

* Mô tả các biến cố và xử lý

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Biến cố | Xử lý | Mã xử lý |
| 1 | Load trang | * Kiểm tra trạng thái đăng nhập của người dùng. Nếu chưa đăng nhập thì về trang đăng nhập * Lấy mã tác vụ cần xem và load kết quả thực hiện vào bảng 1 | XL14 |
| 2 | Người dùng chọn result | * Chuyển đến trang download dữ liệu | XL15 |

### Giao diện thống kê dữ liệu

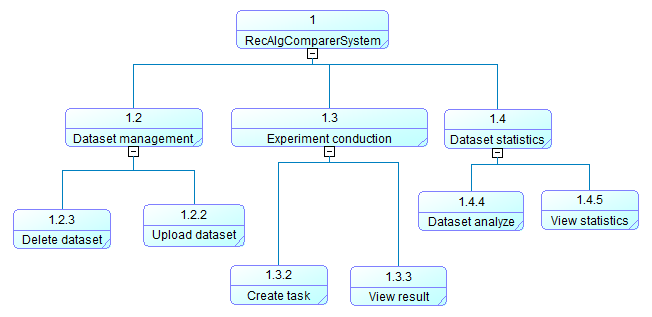
### Giao diện quên mật khẩu

## **Thiết kế xử lý**

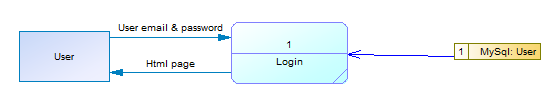
1. Danh sách các xử lý:

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Tên xử lý | Mô tả | Mã số xử lý |
| 1 | Đăng nhập hệ thống | Đăng nhập hệ thống trước khi sử dụng. | XL1 |
| 2 | Người dùng chọn quên mật khẩu | Mở trang khôi phục mật khẩu | XL2 |
| 3 | Người dùng chọn đăng ký | Mở trang đăng ký tài khoản | XL3 |
| 4 | Kiểm tra trạng thái đăng nhập | Kiểm tra trạng thái đăng nhập của người dùng. | XL4 |
| 5 | Hiển thị danh sách dataset | Liệt kê danh sách dataset từ thư mục lưu trữ của người dùng. | XL5 |
| 6 | Người dùng chọn mở trang xem thống kê | Chuyển đến trang xem thống kê. | XL6 |
| 7 | Xử lý Upload dataset | Kiểm tra dataset mà người dùng nhập. Sau đó lưu trữ vào thư mục dataset của người dùng. | XL7 |
| 8 | Load trang đăng ký tài khoản | Kiểm tra đăng nhập khi load trang đăng ký tài khoản. | XL8 |
| 9 | Người dùng chọn đăng ký tài khoản. | Kiểm tra & Lưu thông tin mà người dùng đăng ký vào hệ thống. Sau đó thông báo kết quả thực hiện. | XL9 |
| 10 | Người dùng chọn đăng nhập | Mở trang đăng nhập hệ thống. | XL10 |
| 11 | Load trang chạy thực nghiệm | Lấy danh sách dataset, danh sách các task đưa lên giao diện người dùng. | XL11 |
| 12 | Xem kết quả thực hiện | Lấy dữ liệu chạy thực nghiệm hiển thị cho người dùng. | XL12 |
| 13 | Tạo task | Lưu thông tin tác vụ vào hệ thống. Sau đó chạy task vụ. | XL13 |
| 14 | Load trang xem kết quả thực nghiệm | Lấy kết quả thực nghiệm của dataset tương ứng. | XL14 |
| 15 | Download kết quả thực nghiệm. | Lấy kết quả thực nghiệm gửi về cho người dùng. | XL15 |
| 16 |  |  | XL16 |
| 17 |  |  | XL17 |
| 18 |  |  | XL18 |

1. Biểu đồ phân cấp chức năng:

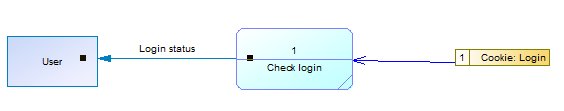


1. Mô tả xử lý
2. Xử lý đăng nhập hệ thống(XL1)
   * DFD xử lý:



* + Mã giả:
    - B1: Kiểm tra user và password mà người dùng nhập. Nếu có một trong hai giá trị bằng rỗng thì yêu cầu nhập lại. Nếu đầy đủ thì sang bước 2.
    - B2: Băm password bằng mã MD5. Sau đó đối chiếu user email và password được lưu trong database. Nếu giống nhau thì chuyển sang bước 4. Nếu khác nhau thì chuyển sang bước 3.
    - B3: Thông báo user email hoặc password không tồn tại. Và quay trở lại màn hình đăng nhập.
    - B4: Lưu vào bảng USER token md5(user email + datatime(“ddMMYYY”)) và lưu một cookie có nội dung tương tự với tên login\_staus trên máy người dùng.
    - B5: Chuyển đến trang chủ.

1. Kiểm tra trạng thái đăng nhập(XL4):
   * DFD xử lý:



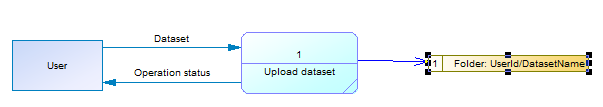
* + Mã giả:
    - B1: Lấy cookie có tên là login\_status từ gói tin http request
    - B2: So sánh token được lưu trong cookie và token được lưu trong bảng USER. Nếu bằng nhau thì tiếp tục thực hiện request. Nếu không bằng thì chuyển đến trang đăng nhập.

1. Xử lý load dataset(XL5)
   * DFD xử lý:



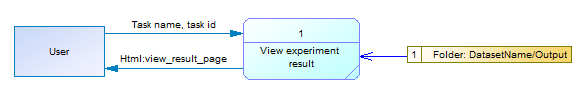
* + Mã giả:
    - B1: Lấy tất cả các thư mục con thuộc thư mục user và đưa lên bảng danh sách dataset
    - B2: Nếu dataset đã được chạy phân tích thì gắn link tới trang thống kê dữ liệu.

1. Xử lý upload dataset(XL7)
   * DFD xử lý:



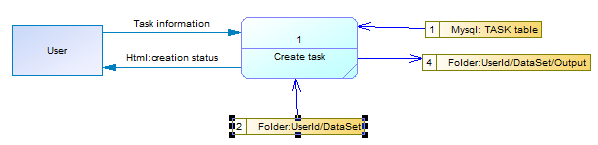
* + Mã giã:
    - B1: Kiểm tra dữ liệu mà người dùng nhập. Nếu hợp lệ thì chuyển sang bước 2. Nếu không thì bắt buộc người dùng nhập lại.
    - B2: Lấy dataset name mà người dùng nhập. sau đó tạo folder trong thư mục UserId/dataset\_name
    - B3: Lấy file mà người dùng upload lên hệ thống và tương ứng đổi tên thành job.txt, cv.txt, score.txt
    - B4: Load dataset vừa upload lên table danh sách dataset trong giao diện quản lý dataset.
    - B5: Thông báo kết quả thực hiện.

1. Xử lý xem kết quả thực hiện(XL12)
   * DFD xử lý:



* + Mã giả:
    - B1: Lấy task name và task id và kiểm tra tính hợp lệ. Nếu hợp lệ chuyển sang bước 2. Ngược lại thông báo không tìm thấy task.
    - B2: Lấy dữ liệu trong thư mục có tên là dataset name và đưa vào datatable thuộc giao diện xem kết quat thực nghiệm.
    - B3: Tạo link download file kết quả
    - B4: Load những đánh giá kết quả thực hiện chứa trong thư mục User/Datasetname/Ouput

1. Xử lý tạo task(XL13)
   * Xử lý tạo task:



* + Mã giả:
    - B1: Kiểm tra dữ liệu mà người dùng nhập. Nếu hợp lệ thì chuyển sang bước 2. Nếu không thì yêu cầu nhập lại.
    - B2: Chèn vào bảng TASK.
    - B3: Ghi file execuate.bat vào thư mục Datasetname/Ouput có nội dung java –jar “Datasetname” “Datasetname/Ouput”
    - B4: Chạy file execuate.bat

# **LẬP TRÌNH VÀ KIỂM THỬ**

# **KẾT LUẬN**