

VIETNAM NATIONAL UNIVERSITY HO CHI MINH CITY  
HO CHI MINH CITY UNIVERSITY OF TECHNOLOGY  
FACULTY OF COMPUTER SCIENCE AND ENGINEERING



**ĐỒ ÁN TỔNG HỢP**

---

**HƯỚNG KỸ THUẬT DỮ LIỆU**

**BOOK STORE**

---

GVHD: LÊ THỊ BẢO THU

NHÓM: 5

HO CHI MINH CITY, SEPTEMBER 2024



## Danh sách thành viên

No.	Họ và tên	Mã số sinh viên	Phân công nhiệm vụ
1	Nguyễn Quang Minh	2212063	Các kỹ thuật xử lý dữ liệu
2	Đỗ Hoàng Quân	2212779	Các yêu cầu chức năng
3	Đỗ Minh Quân	2212781	Các yêu cầu phi chức năng
4	Nguyễn Anh Tài	2212980	Các công cụ giám sát hiệu năng của DBMS



## Contents

<b>1</b>	<b>Mô tả đề tài</b>	<b>4</b>
1.1	Mô tả hệ thống . . . . .	4
1.1.1	Yêu cầu chức năng (Functional Requirement) . . . . .	4
1.1.2	Quyền hạn của tài khoản . . . . .	4
1.1.3	Chức năng cho từng vai trò . . . . .	4
1.1.4	Yêu cầu phi chức năng (Non-Functional Requirement) . . . . .	5
1.2	Kỹ thuật dữ liệu . . . . .	6
1.2.1	Tiền xử lý dữ liệu (Data Preprocessing) . . . . .	6
1.2.2	Xử lý và truy xuất dữ liệu (Data Processing and Retrieval) . . . . .	6
1.2.3	Phân tích dữ liệu (Data Analysis) . . . . .	7
1.2.4	Các công cụ giám sát hiệu năng của DBMS . . . . .	7

# 1 Mô tả đề tài

## 1.1 Mô tả hệ thống

Hệ thống quản lý sách (Book Store) được thiết kế để hỗ trợ để hỗ trợ các hoạt động quản lý và kinh doanh sách, bao gồm việc theo dõi thông tin sách, tác giả, thể loại, sách tồn kho, cũng như xử lý đơn hàng và quản lý nhân sự, khách hàng, khuyến mãi. Với hệ thống quản lý này, cửa hàng sách có thể vận hành hiệu quả và cung cấp trải nghiệm đọc và mua bán sách tiện lợi cho khách hàng. Hệ thống đem lại các lợi ích như:

- Quản lý toàn bộ thông tin về sách và đơn hàng một cách dễ dàng và nhanh chóng.
- Tích hợp chức năng tìm kiếm và tự động gửi thông báo cho khách hàng.
- Tối ưu hoá việc quản lý hàng tồn kho, giúp đảm bảo sản phẩm được yêu cầu luôn có sẵn.
- Cung cấp các báo cáo thống kê về doanh thu và tình hình kinh doanh.
- Hỗ trợ quản lý các chiến dịch khuyến mãi.

### 1.1.1 Yêu cầu chức năng (Functional Requirement)

Hệ thống có các loại tài khoản tương ứng với các vai trò: Quản trị viên (Admin), Nhân viên và Khách hàng. Mỗi tài khoản có quyền truy cập và sử dụng các chức năng phù hợp với vai trò của mình.

### 1.1.2 Quyền hạn của tài khoản

- **Khách hàng:** Có thể xem thông tin sách mà không cần đăng nhập, hoặc đăng ký tài khoản để mua sách.
- **Nhân viên:** Được cấp tài khoản bởi Admin để quản lý sách, đơn hàng và hỗ trợ khách hàng.
- **Admin(Quản trị viên):** Quản lý toàn bộ hệ thống, từ nhân sự, khách hàng, các loại sách cho đến chính sách khuyến mãi.

### 1.1.3 Chức năng cho từng vai trò

Tất cả các tài khoản, trừ Guest (Khách không đăng nhập), đều có thể thực hiện các thao tác đăng nhập, đăng xuất, chỉnh sửa thông tin cá nhân và thay đổi mật khẩu.

- **Guest (Khách không đăng nhập):**

- Tra cứu thông tin về sách, tác giả và các thể loại có sẵn trong hệ thống.
- Đăng ký tài khoản khách hàng để mua bán sách trực tuyến.
- Xem thông tin các chương trình khuyến mãi.

- **Khách hàng:**

- Xem danh mục sách và thực hiện mua sách trực tuyến.
- Theo dõi lịch sử các đơn hàng đã thực hiện.
- Thực hiện thanh toán trực tuyến sau khi đặt hàng.
- Nhận thông báo về đơn hàng và các ưu đãi mới nhất.
- Quản lý giỏ hàng, điều chỉnh số lượng sách muốn mua.

- **Nhân viên:**

- Quản lý sách: Thêm mới, chỉnh sửa hoặc xóa thông tin về sách, thể loại, tác giả và cập nhật lượng sách tồn kho.
- Quản lý đơn hàng: Xử lý đơn hàng, cập nhật tình trạng giao hàng, và hỗ trợ khách hàng trong quá trình đặt sách.
- Quản lý thông tin khách hàng: Theo dõi, hỗ trợ và giải quyết các yêu cầu của khách hàng.
- Thống kê và báo cáo: Tạo báo cáo doanh thu, tình trạng kho sách và các số liệu kinh doanh khác.

- **Admin:**

- Quản lý nhân viên: Tạo tài khoản mới, chỉnh sửa thông tin hoặc xóa tài khoản nhân viên.
- Quản lý chính sách khuyến mãi: Tạo và quản lý các chương trình khuyến mãi dành cho khách hàng.
- Thiết lập và quản lý hệ thống: Cấu hình các chức năng chính của hệ thống và cơ sở dữ liệu.
- Thống kê tổng quan: Theo dõi doanh thu, tình hình kinh doanh và hiệu quả hoạt động của hệ thống.

#### 1.1.4 Yêu cầu phi chức năng (Non-Functional Requirement)

- Thân thiện với người dùng, dễ dàng sử dụng sau 3 lần thao tác

- Hoạt động ổn định trong trường hợp có tối đa 500 người dùng truy cập cùng lúc
- Tuân thủ theo các quy tắc bảo mật thông tin và bảo đảm quyền riêng tư của người dùng
- Khả năng phản hồi và thực hiện các thao tác nhanh chóng (trong vòng 2 giây)
- Hệ thống hoạt động liên tục 24/7
- Dễ dàng thay đổi, phát triển, mở rộng hệ thống trong tương lai

## 1.2 Kỹ thuật dữ liệu

### 1.2.1 Tiền xử lý dữ liệu (Data Preprocessing)

- Làm sạch dữ liệu (Data Cleaning): quá trình loại bỏ hoặc sửa các giá trị không chính xác, bị thiếu hoặc không phù hợp trong dữ liệu để đảm bảo tính chính xác và đáng tin cậy.
- Tích hợp dữ liệu (Data Integration): Kết hợp dữ liệu từ nhiều nguồn khác nhau để tạo thành một tập dữ liệu thống nhất. Kỹ thuật này quan trọng khi dữ liệu đến từ nhiều hệ thống hoặc định dạng khác nhau và cần phải hợp nhất để phân tích.
- Chuyển đổi dữ liệu (Data Transformation): Chuyển đổi dữ liệu từ dạng này sang dạng khác để phù hợp với phân tích hoặc mô hình. Điều này bao gồm việc chuẩn hóa, thay đổi định dạng, hoặc biến đổi dữ liệu để dễ dàng sử dụng trong các mô hình phân tích.
- Làm giảm dữ liệu (Data Reduction): Quá trình giảm thiểu kích thước của dữ liệu trong khi vẫn giữ lại thông tin quan trọng. Các kỹ thuật bao gồm chọn lọc các thuộc tính cần thiết, lọc bớt các bản ghi không quan trọng, hoặc sử dụng các thuật toán nén dữ liệu.

### 1.2.2 Xử lý và truy xuất dữ liệu (Data Processing and Retrieval)

- Chỉ mục (Indexing): Cấu trúc dữ liệu trong cơ sở dữ liệu giúp tăng tốc độ truy vấn bằng cách tạo các chỉ số trên các cột hoặc nhóm cột cụ thể. Khi có chỉ mục, cơ sở dữ liệu không cần phải quét toàn bộ bảng để tìm dữ liệu, mà thay vào đó sẽ tra cứu thông qua chỉ mục.
- Tối ưu hóa câu truy vấn (Query Optimization): Quá trình cải thiện cách câu truy vấn được viết hoặc thực hiện để tối thiểu hóa thời gian và tài nguyên sử dụng. Điều này bao gồm việc sử dụng các chỉ mục, lựa chọn thuật toán truy vấn tốt nhất và tránh các thao tác tốn kém.
- Phân mảnh dữ liệu (Data Partitioning): Kỹ thuật chia cơ sở dữ liệu lớn thành nhiều phần nhỏ hơn (phân mảnh) để cải thiện hiệu suất truy vấn và quản lý. Mỗi phân mảnh có thể lưu trữ ở các máy chủ hoặc vùng khác nhau, giúp giảm tải và tăng khả năng mở rộng.

- Bộ nhớ đệm (Caching): Lưu trữ các kết quả truy vấn phổ biến hoặc dữ liệu thường xuyên truy cập trong bộ nhớ tạm thời (RAM) để giảm thời gian truy vấn khi cần dùng lại dữ liệu đó.

### 1.2.3 Phân tích dữ liệu (Data Analysis)

- Phân tích thống kê (Statistical Analysis): Sử dụng các phương pháp thống kê để mô tả, tóm tắt và tìm ra các mô hình trong dữ liệu.
- Phân tích hồi quy (Regression Analysis): Hồi quy là kỹ thuật dùng để dự đoán một giá trị liên tục dựa trên các đặc trưng đầu vào, một hàm số tuyến tính được sử dụng để mô tả mối quan hệ giữa biến độc lập và biến phụ thuộc.
- Phân tích phân loại (Classification Analysis): Kỹ thuật này dùng để gán nhãn cho dữ liệu và phân loại chúng vào các nhóm khác nhau.
- Phân tích phân cụm (Clustering Analysis): Phân cụm là kỹ thuật dùng để nhóm các đối tượng dữ liệu có đặc điểm tương tự nhau mà không cần gán nhãn trước.
- Trực quan hóa dữ liệu (Data Visualization): kỹ thuật sử dụng các biểu đồ, đồ thị và hình ảnh để trình bày dữ liệu, giúp người dùng dễ dàng nhận thấy các xu hướng, mô hình, và quan hệ trong dữ liệu.

### 1.2.4 Các công cụ giám sát hiệu năng của DBMS

- MySQL Enterprise Monitor:
  - Cung cấp khả năng giám sát liên tục hệ thống MySQL trong thời gian thực.
  - Hiển thị trực quan các thông tin về hiệu suất để phát hiện nhanh các vấn đề.
  - Có sẵn các cảnh báo tự động cho những sự kiện quan trọng như truy vấn chậm, khóa cơ sở dữ liệu, độ trễ truy vấn và các vấn đề tài nguyên.
  - Cải thiện và tối ưu hóa hiệu năng
  - Theo dõi thời gian thực thi truy vấn, tần suất chạy truy vấn, và sử dụng tài nguyên của từng truy vấn.
  - Cho phép có được thông tin đầy đủ về sao lưu.
  - Cung cấp cảnh báo về lỗi hỏng bảo mật và giúp theo dõi các quyền truy cập để đảm bảo cơ sở dữ liệu được bảo vệ an toàn.
  - Giám sát người dùng, quyền hạn, và các truy cập bất thường vào hệ thống.
- Percona Monitoring and Management:



- Cung cấp một phân tích dựa trên thời gian kỹ lưỡng cho MySQL. Phân tích này giúp bạn hiểu rõ hơn về sức khỏe và hiệu suất của cơ sở dữ liệu.
- Sử dụng mô hình client/server, có nghĩa là, bạn sẽ phải tải xuống và cài đặt cả máy khách và máy chủ.
- Công cụ phân tích truy vấn (QAN) luôn giám sát thời gian thực hiện các truy vấn. Có thể nhận được thông tin chi tiết về các truy vấn SQL mất nhiều thời gian để thực hiện và hiểu nguyên nhân gốc rễ của chúng.
- PMM tích hợp khả năng thiết lập cảnh báo tự động cho các chỉ số quan trọng của hệ thống, ví dụ như truy vấn chậm, sử dụng tài nguyên quá mức, hoặc tình trạng khóa dữ liệu.
- Lập bản đồ truy vấn dựa trên số liệu để cung cấp cho bạn dữ liệu thích hợp, giúp đưa ra những quyết định quan trọng.
- Giám sát quyền truy cập và các hoạt động liên quan đến người dùng, đảm bảo dữ liệu được bảo vệ an toàn.