Bookstore Database Design

Part 4

Manisha Goyal (mg7609)

Rhea Chandok (rc5397)

Sanam Palsule (sp7940)

Table of Contents

Business Use Cases for Workflow-Based Application

As part of our end-to-end solution for the bookstore application, we previously established a range of business use cases in Part 3. These use cases were meticulously selected and designed to support the creation of a data-driven, workflow-based application. Here, we summarize these business use cases and the processes integral to our application.

Business Use Cases Summary

Employee Management:

- Use Cases: Employee login, book restocking, and new employee registration.
- Processes: These processes involve inventory management, and employee data handling within the application.

Customer Engagement:

- Use Cases: Customer login, searching books by title/author, accessing personalized book recommendations, browsing the book catalog, and completing book purchases.
- Processes: Customer interactions encompass search functionality, personalized services based on user history and preferences, catalog navigation, and transaction processing.

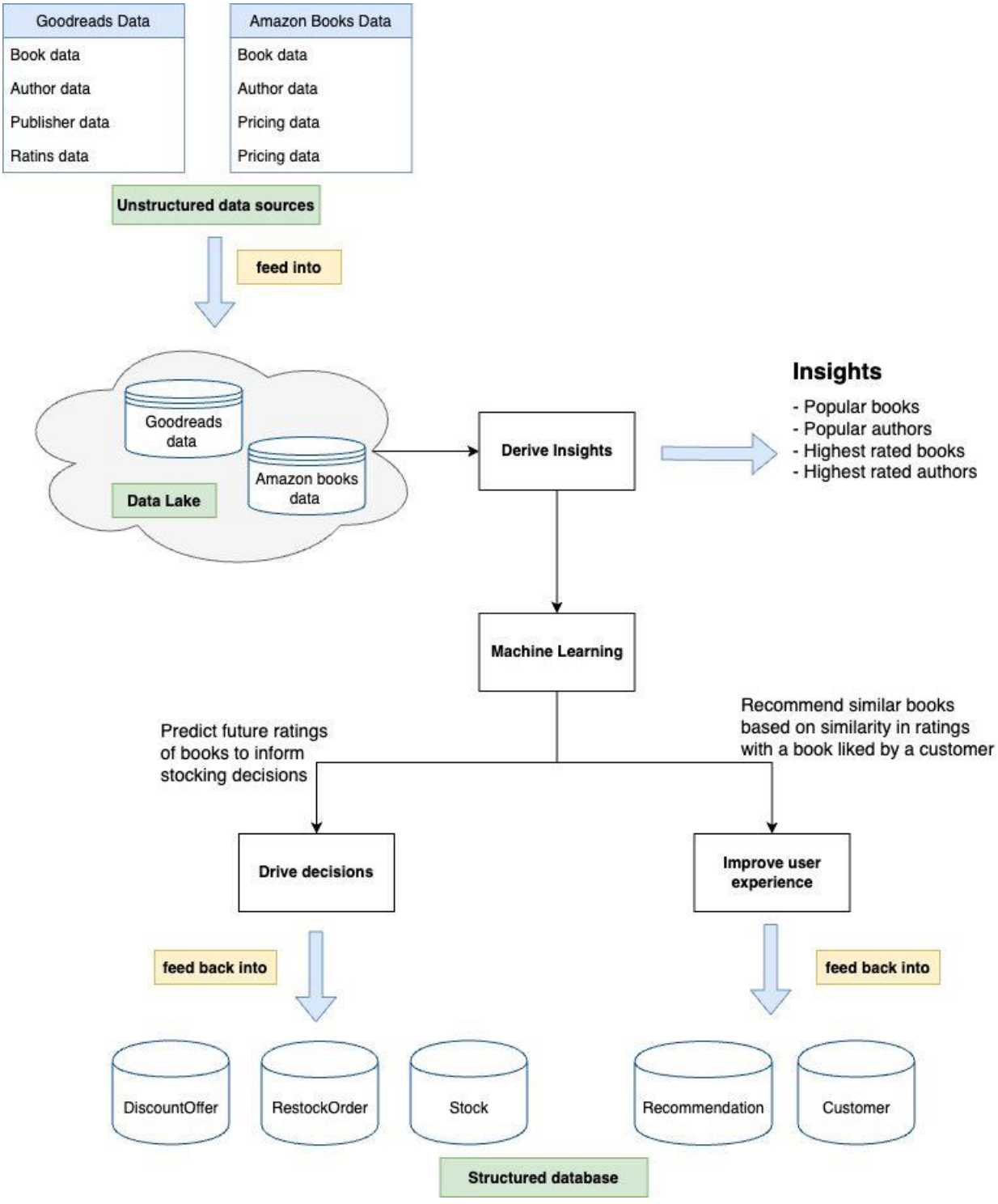
Data-Driven Recommendations:

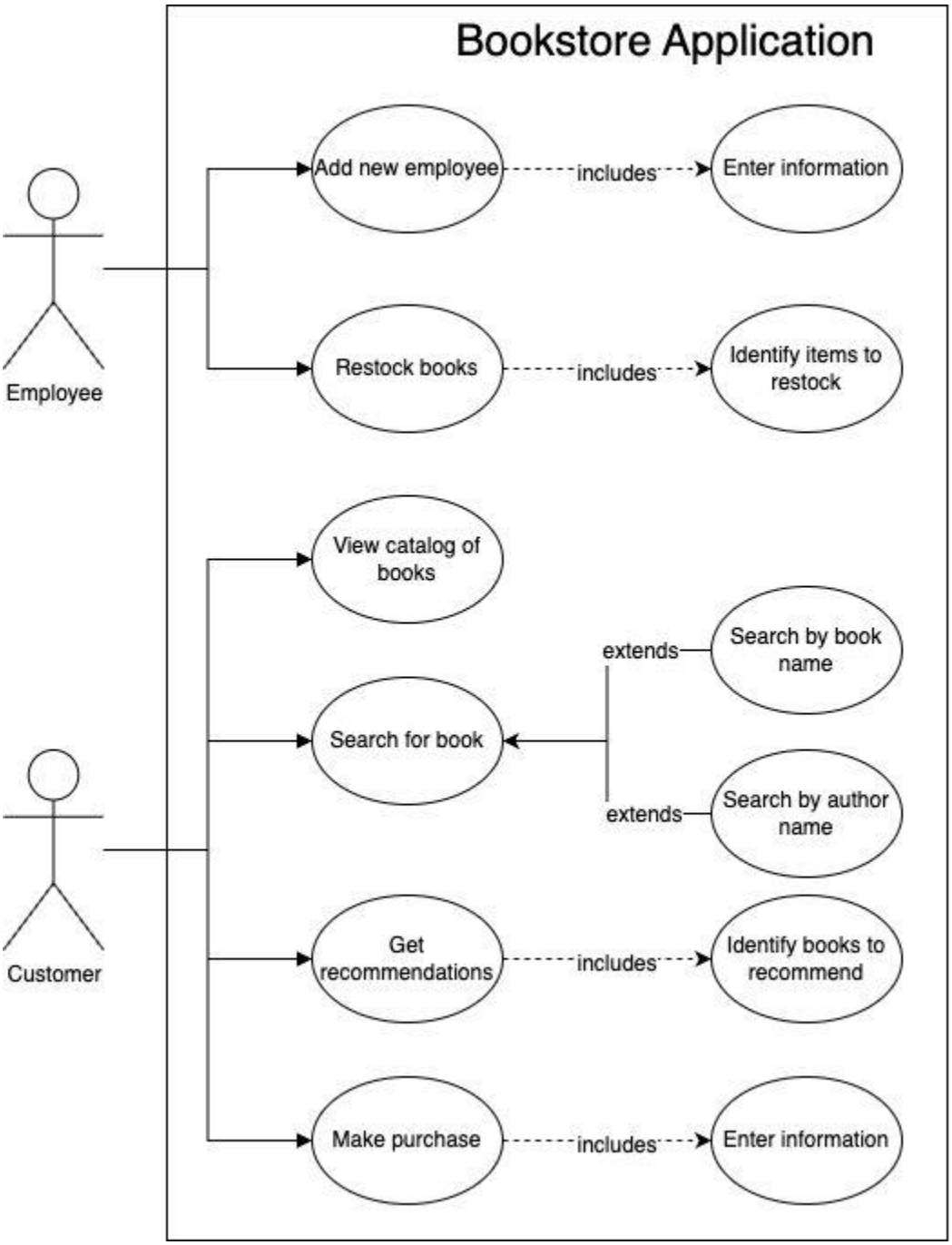
- Use Case: Leveraging machine learning models to provide personalized book recommendations.
- Process: Involves analyzing customer data and preferences using machine learning models to generate relevant book suggestions.

Workflow Integration

Seamless Integration: The application workflows are designed to seamlessly integrate these business use cases, ensuring a smooth and intuitive user experience.

Data-Driven Decision Making: Each use case is backed by data-driven logic, ensuring that the application not only serves operational needs but also enhances decision-making through insights derived from data analytics and machine learning models.


Goodreads Data
Book data
Author data
Publisher data
Ratins data
Amazon Books Data
Book data
Author data
Pricing data
Pricing data
Unstructured data sources
feed into
Goodreads data
Amazon books data
Data Lake
Derive Insights
Insights
- Popular books
- Popular authors
- Highest rated books
- Highest rated authors
Machine Learning
Predict future ratings of books to inform stocking decisions
Recommend similar books based on similarity in ratings with a book liked by a customer
Drive decisions
Improve user experience
feed back into
feed back into
DiscountOffer
RestockOrder
Stock
Recommendation
Customer
Structured database



The business use cases selected for our bookstore application are central to its functionality, catering to both internal management and customer-facing services. Their implementation within the application reflects our commitment to creating a data-driven, efficient, and user-friendly environment. This summary encapsulates the essence of the use cases and workflows designed in Part 3, setting the stage for their integration into the final application developed in Part 4.

Design of data-driven workflow-based database application

Our bookstore application has adopted a complex n-tier architecture, seamlessly integrating structured and unstructured data management with advanced machine learning capabilities. This ensures the application's scalability, security, and robustness, catering to the diverse needs of both business operations and user interactions.

N-Tier Architecture Adaptation

Presentation Layer:

- Responsible for displaying information to the user, including book catalogs and reviews, and facilitating data entry and interactive elements like searching and purchasing books.
- Built with HTML and CSS, it offers an intuitive and responsive user interface across various devices, ensuring a seamless user experience.

Business Logic Layer:

- Acts as the intermediary that enforces business rules, manages application state, controls user flow, and processes requests from the presentation layer.
- Handles data retrieval and persistence with the database services layer and directs advanced processing tasks to the machine learning layer.

Machine Learning Layer:

- Hosts, executes, and manages machine learning models for tasks like predictive modeling, and personalized book recommendations.
- Interacts with both structured and unstructured data, undertaking preprocessing and feature extraction for model input.

Database Services Layer:

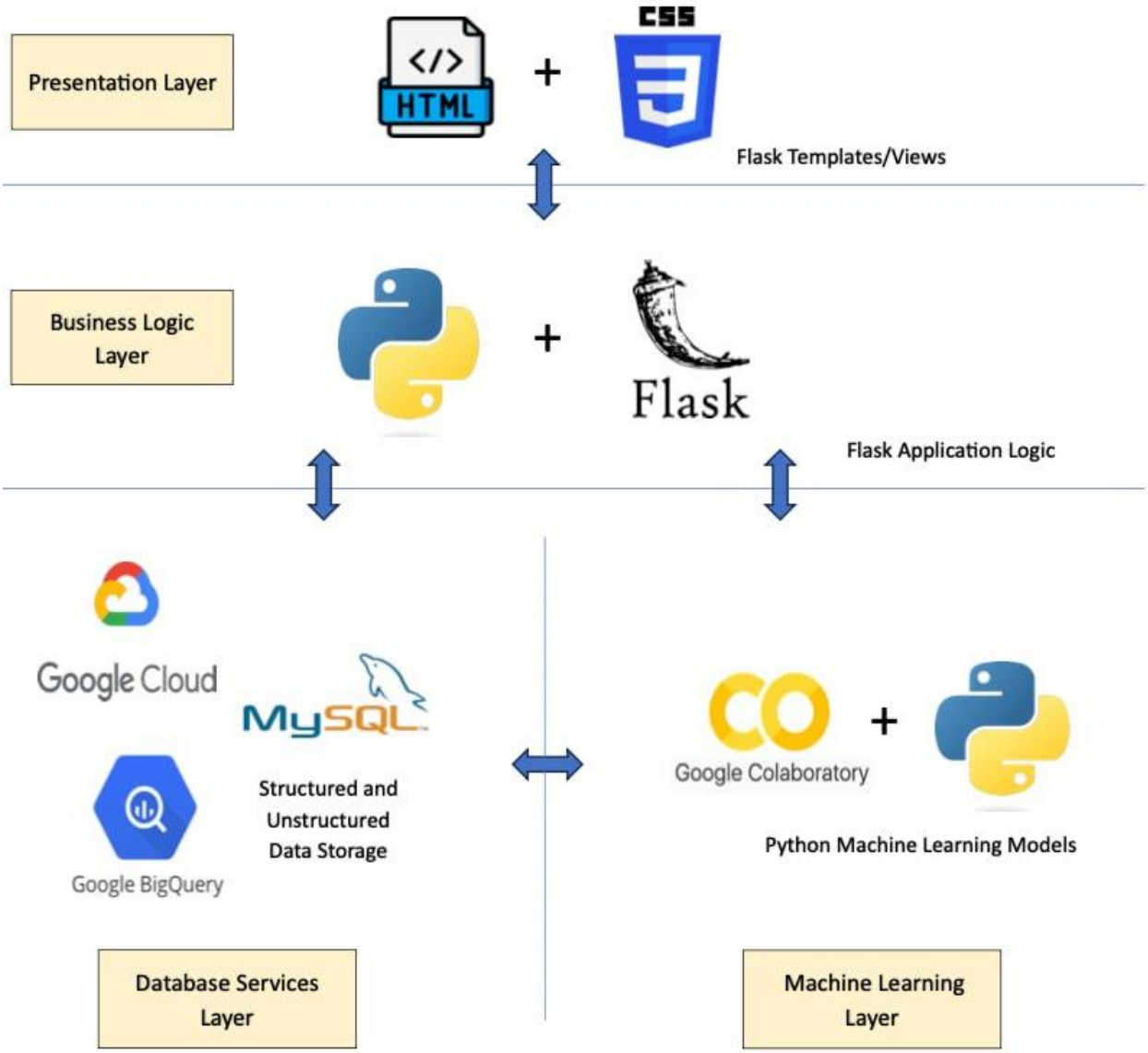
- The foundational layer for data storage and management, optimized to support a logical schema accommodating both transactional operations and machine learning data requirements.
- Manages interactions with the data lake, which stores unstructured data critical for machine learning insights.

Leveraging Hybrid Data and Cloud Services

Data Lake Integration: Central to managing unstructured data, like Good Reads reviews, vital for machine learning models.

Cloud-Based ML Services: Utilizes Google Cloud's AI and machine learning capabilities for efficient model processing, training, evaluation, and deployment.

Security and Data Transfer: Implements advanced encryption technologies for secure data movement between layers, ensuring protection of sensitive information.



The adoption of an n-tier architecture enhances the bookstore application's data analytics and personalization capabilities. This sophisticated architecture not only delineates clear separation of responsibilities but also provides a flexible and scalable foundation, ready to accommodate future technological advancements and data-driven business intelligence needs.

Implementation of Data-Driven Program Module and Workflow-Based Application

Application Overview

Our Flask-based application represents the culmination of our project, integrating data-driven insights with an end-to-end workflow for both employees and customers. The application leverages the machine learning models developed earlier, handling unstructured data pipelines and retraining mechanisms seamlessly.

Data-Driven Program Module

MySQL Database: Our application leverages a MySQL database, chosen for its robustness, scalability, and widespread industry adoption. MySQL effectively handles both the transactional data for the bookstore operations and the data required for machine learning insights.

Integration of Machine Learning Models:

- The application includes a module (‘recommendation.py’) that utilizes our trained machine learning model to provide book recommendations. The Book Recommendation Engine makes personalized book suggestions based on the user’s input of a title. It utilizes a pre-trained machine learning model to find semantic similarities between books and generate relevant recommendations.

- The (‘rating.py’) module contains the logic to generate a rating prediction for a book based on attributes like number of pages, and reviews using a pre-trained machine learning model. It takes an ISBN as input and passes the book details from the dataset into the model to predict the expected average rating. After making suggestions, it also updates the inventory stock levels to increase the count for the books which are less than 20 in quantity and have a rating higher than 3.85.

- Overall, the unstructured data powers personalization of likely bestselling items, while integration with inventory management.

Workflow-Based Application Implementation

Flask Application Structure:

'app.py' serves as the backbone of the Flask application, routing various functionalities and integrating the data-driven modules with the user interface.

Database Connectivity and ORM:

Utilizing SQLAlchemy (an ORM framework), the application achieves efficient database interactions though automated mapping between Python classes and database tables. This allows interacting with database rows as Python objects for cleaner code.

Some advantages for our system:

- Simpler interface to SQL storage
- Handles object model and relational mapping
- Change tracking and caching
- Easier modeling of recommendations domain

For example, in our case, we access the Employee table using ORM, where the mapping is from Employee table in the database to the Employee.py file.

User Interface:

The 'templates' directory contains the HTML templates, providing a user-friendly interface for both customers and employees, as outlined in the design.

Static Assets:

The 'static' directory contains CSS files, ensuring that the application is not only functional but also aesthetically pleasing and intuitive to use.

Application Optimization

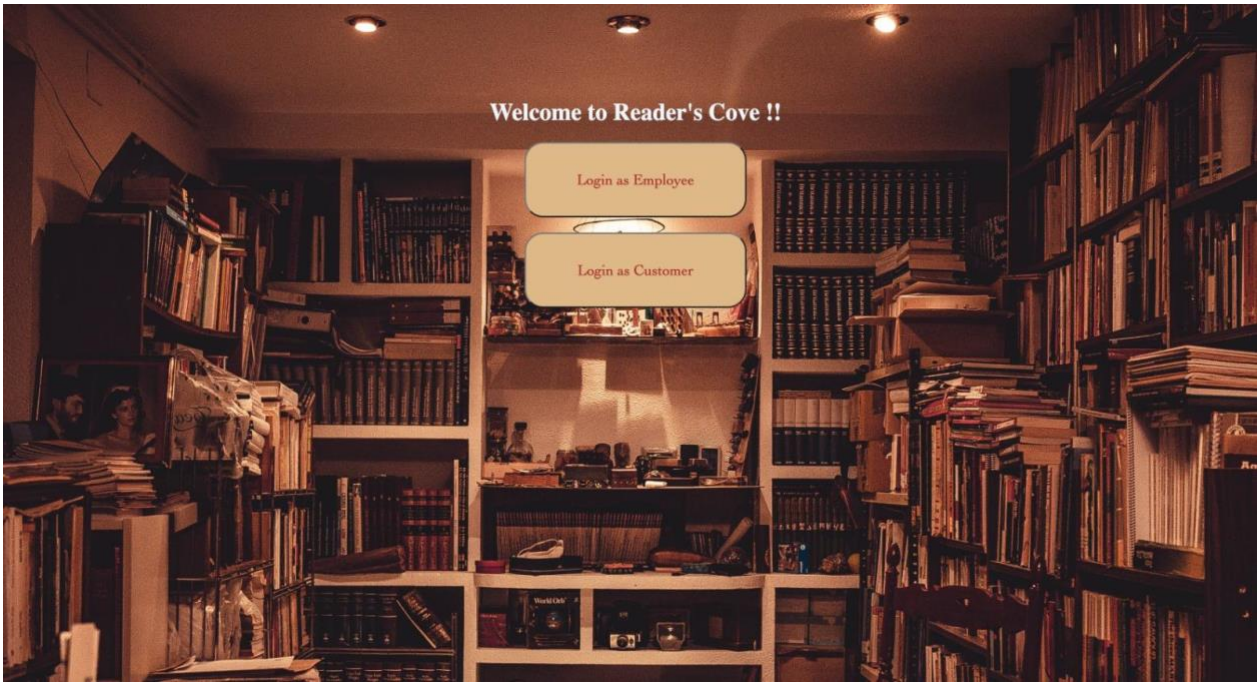
Query Optimization: The application employs optimized database queries, reducing response times and improving overall performance. Techniques like indexing and query restructuring are used to enhance efficiency.

ORM Optimization: The use of SQLAlchemy ORM allows for optimized data access patterns, reducing the overhead of database interactions and abstracting complex queries.

Demonstration and Validation

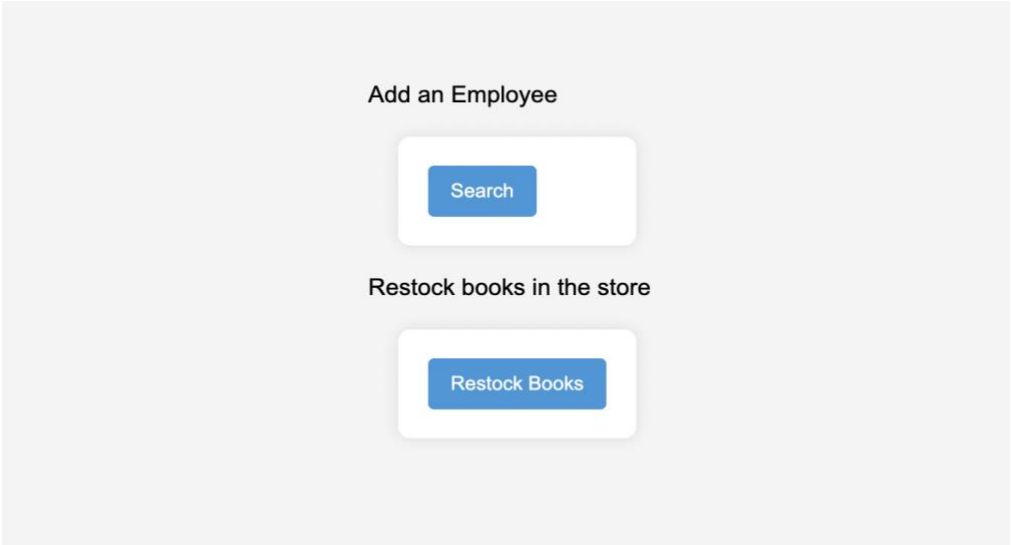
Walkthrough of the Application:

Landing Page:

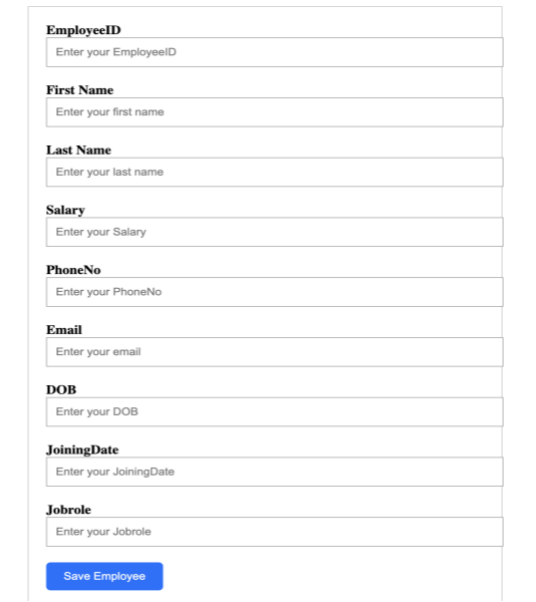


Employee Login: As an employee you can do the following:

1. Add Employee
2. Restock Books

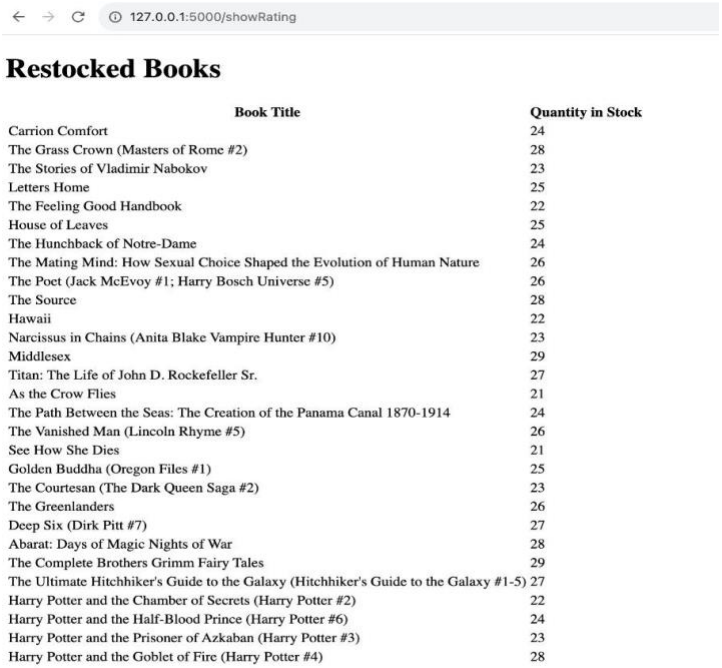


1. **Add an Employee:**



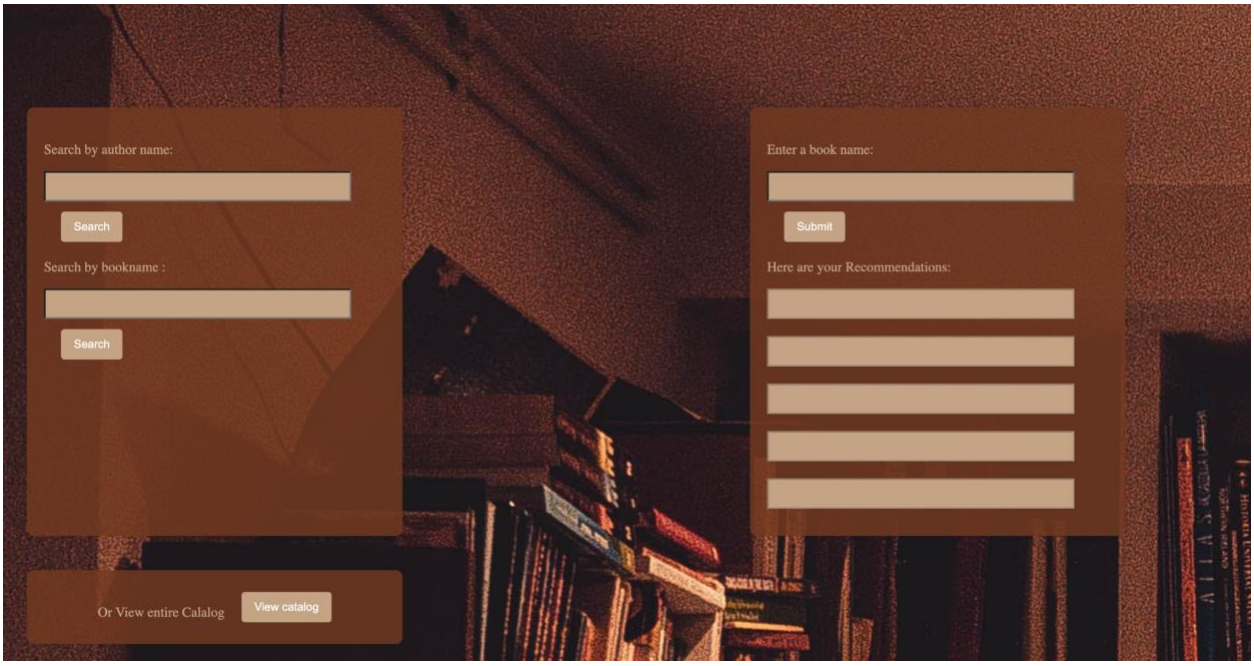
2. **Restock books:**

Based on the ratings from our machine Learning model, we restock the books.



Customer Login: As a customer you can do the following:

1. Search for books
2. Get book recommendations
3. View catalog of books
4. Checkout and purchase books



Search by Book/Author:

Customers can search the bookstore's catalog by book title or author name. This use case involves querying the database and presenting results in a user-friendly manner.

Search by Author - results:

Search Results

	Book Title	Author
☑	Harry Potter and the Chamber of Secrets (Harry Potter #2)	J.K. Rowling
☑	Harry Potter and the Half-Blood Prince (Harry Potter #6)	J.K. Rowling
☐	Harry Potter and the Prisoner of Azkaban (Harry Potter #3)	J.K. Rowling
☐	Harry Potter and the Goblet of Fire (Harry Potter #4)	J.K. Rowling

Proceed to Checkout

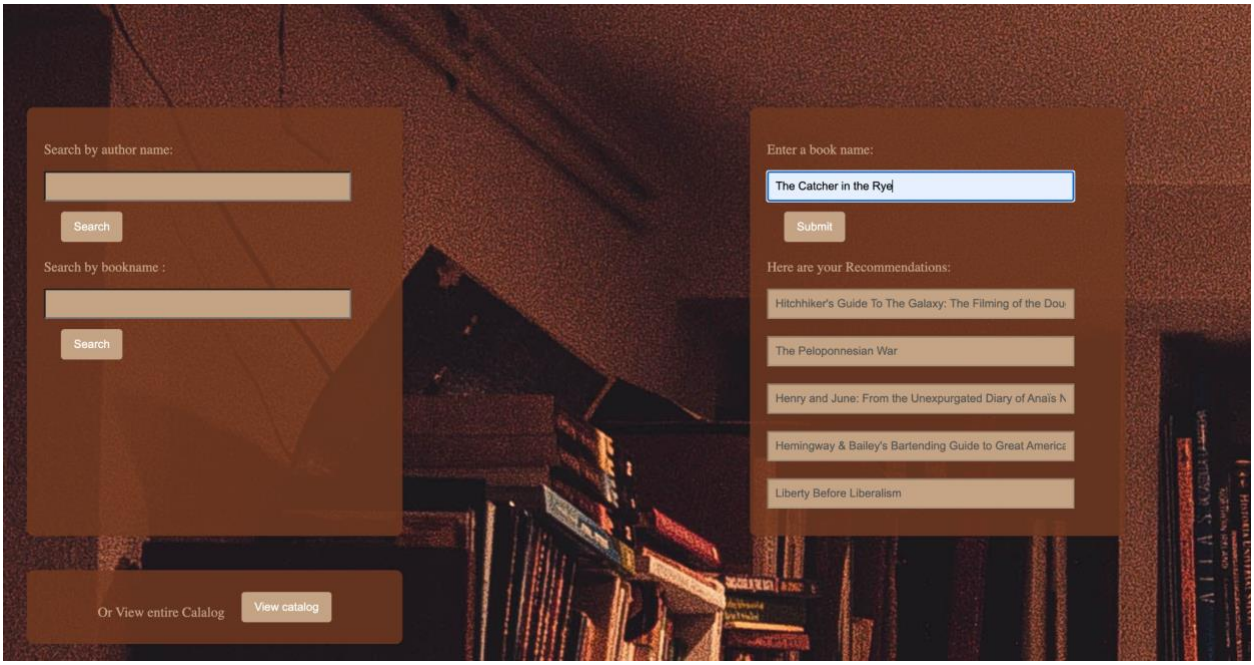
Search by Book: results:

Search Results

	Book Title	Author
☐	Carrion Comfort	Dan Simmons

Proceed to Checkout

Get Book Recommendations:



View Catalog of Books:

Search Results

Book Title	Author
Carrion Comfort	Dan Simmons
The Grass Crown (Masters of Rome #2)	Colleen McCullough
The Stories of Vladimir Nabokov	Vladimir Nabokov
Letters Home	Sylvia Plath
The Feeling Good Handbook	David D. Burns
House of Leaves	Mark Z. Danielewski
The Hunchback of Notre-Dame	Victor Hugo/Walter J. Cobb
The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature	Geoffrey Miller
The Valley of Horses (Earth's Children #2)	Jean M. Auel
The Poet (Jack McEvoy #1; Harry Bosch Universe #5)	Michael Connelly
The Source	James A. Michener
Hawaii	James A. Michener
Narcissus in Chains (Anita Blake Vampire Hunter #10)	Laurell K. Hamilton
Strangers	Dean Koontz
The Sigma Protocol	Robert Ludlum
Matthew Flinders' Cat	Bryce Courtenay
Wideacre (The Wideacre Trilogy #1)	Philippa Gregory
Middlesex	Jeffrey Eugenides
The Jester	James Patterson/Andrew Gross
Titan: The Life of John D. Rockefeller Sr.	Ron Chernow
The Secret History	Donna Tartt
As the Crow Flies	Jeffrey Archer
The Path Between the Seas: The Creation of the Panama Canal 1870-1914	David McCullough
The Vanished Man (Lincoln Rhyme #5)	Jeffery Deaver
See How She Dies	Lisa Jackson
Golden Buddha (Oregon Files #1)	Clive Cussler/Craig Dirgo
Winter's Tale	Mark Helprin
The Courtesan (The Dark Queen Saga #2)	Susan Carroll
The Greenlanders	Jane Smiley
Deep Six (Dirk Pitt #7)	Clive Cussler
Abarat: Days of Magic Nights of War	Clive Barker
The Complete Brothers Grimm Fairy Tales	Jacob Grimm/Wilhelm Grimm
The Glass Books of the Dream Eaters (Miss Temple Doctor Svenson and Cardinal Chang #1)	Gordon Dahlquist
The Ultimate Hitchhiker's Guide to the Galaxy (Hitchhiker's Guide to the Galaxy #1-5)	Douglas Adams
The Catcher in the Rye	J.D. Salinger
Agile Web Development with Rails: A Pragmatic Guide	Dave Thomas/David Heinemeier Hansson/Leon Breedt/Mike Clark/Thomas Fuchs/Andreas Schwarz
Harry Potter and the Chamber of Secrets (Harry Potter #2)	J.K. Rowling

Checkout and Purchase Books

Checkout

Total:
Your Total is: 93.13 Place Order
Back to Home

Your Order is Placed!

Validation

End-to-End Integration: The application demonstrates a seamless integration of data-driven insights with operational workflows, fulfilling the project's objectives.

Data Governance Compliance: Throughout the development, data governance principles have been adhered to, ensuring data quality, security, and compliance with enterprise goals.

In summary, the developed Flask application represents a comprehensive solution that successfully integrates machine learning insights into a practical, user-centric workflow. It exemplifies the power of combining data-driven programming with robust database management and intuitive user interfaces, thereby achieving the project's goals of enhancing user experience and promoting organizational excellence.

Github Link

The entire codebase along with supporting documentation can be found in this github repository:
https://github.com/manisha-goyal/database-systems-project

Finalization of End-to-End Reference Architecture (RA)

Foundational Principles	**Business Principles**	**Data Principles**	**Application and Technology Principles**
	- Ensure robust operations resistant to system disruptions. - Strict compliance with data regulations and intellectual property protection. - Foster collaborative and informed decision-making through enhanced information management.	- Recognition of data as a key asset, managed with precision and care. - Secure and controlled data sharing to enhance collaborative efforts. - Appointment of data trustees to ensure consistent data quality. - Uniform definition and documentation across the enterprise. - Rigorous data protection measures against unauthorized access.	- User-friendly and well-documented application interfaces. - Application portability and independence, using middleware for flexibility. - Balance in technological choices for optimal efficiency and usability. - Promote software interoperability across various application systems.
Organising Framework	The bookstore's organizing framework includes: - A blend of structured and unstructured data, along with machine learning algorithms, to derive business insights and drive decisions. - Business product components and solution modules that streamline operations and insight delivery. - Integration of both external and internal processes and applications, facilitating seamless interactions with suppliers, partners, and customer-facing platforms.		
Business Solutions	Comprehensive methods for planning, delivery, and operation of business solutions: - Solutions and applications crafted with clear objectives and user-centric designs. - Development pathways incorporating feedback mechanisms for continuous improvement. - Scalable, modular, and resilient application designs. - Compliance-focused development with thorough testing protocols. - Risk mitigation in production environments through strategic integration methods. - Establishment of continuous delivery pipelines following these principles.		
Data Governance	- Data Access and Usage: - Standardized and secure API access for efficient and automated data management. - Designated data trustees overseeing access to sensitive information, employing encryption where necessary. - Systematic tracking of data access and usage to ensure compliance and security. - Data Quality Management: - Continuous monitoring and assessment of data quality across both structured and unstructured datasets. - Implementation of corrective measures to address any discrepancies or inconsistencies detected during data processing. - Data Lifecycle Management: - Managing the data lifecycle from creation to deletion, ensuring data relevance and accuracy at all stages. - Effective data archival and retrieval processes to support business continuity and historical analysis. - Fairness, Accountability, and Transparency: - Ensuring fairness in data-driven decision-making processes, particularly in aspects like book stocking and customer recommendations. - Maintaining a high level of transparency and accountability in how data is used to drive business decisions.		