High Level Design (HLD)

Book Recommender System

Version 1.0

Date of Revision 15 April 2022

Contents

Document Version Control

Date Issued	Version	Description	Author
27th February 2022	1	Initial HLD – V1.0	Savan Javia

Abstract

A recommendation system is one of the top applications of data science. Every consumer Internet company requires a recommendation system like Netflix, Youtube, a news feed, etc. What you want to show out of a huge range of items is a recommendation system.

A book recommendation system is a type of recommendation system where we have to recommend similar books to the reader based on his interest. The books recommendation system is used by online websites which provide books like google play books, open library, goodReads, etc.

1. Introduction

1.1 Why this High-Level Design Document ?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project
- List and describe the non-functional attributes like:
    - Security
    - Reliability
    - Maintainability
    - Portability
    - Reusability
    - Application compatibility
    - Resource utilization
    - Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and the technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

1.3 Definitions

Terms	Description
Database	Collection of books information processed by system
IDE	Integrated Development Environment

2. General Description

2.1 Product Perspective

The Book Recommender System is a machine learning based recommendation system which recommends books to users based on their interest in the books.

2.2 Problem statement

To create a machine learning solution implement the following use cases.

- Take input from users for their interest in books.
- Recommend similar books using trained machine learning models.

2.3 Proposed Solution

For building machine learning models use a content based filtering approach because we have to recommend books to the users according to their interest. Data preprocessing is done on datasets like removing stop words, and applying stemming, so the model can be built in a reliable and stable way in terms of accuracy. Finally we use that model to recommend books to the users.

2.4 Technical Requirements

- build a machine learning web application.

- Use cassandra database
- Use cloud platform for deploying an application

2.5 Data Requirements

Books Database will be used for training the ML model, and for recommendation. Admin can change the database by uploading a CSV file containing book informations.

2.6 Tools Used

i. **Jupyter Notebook** has been used for EDA
ii. For data cleaning and preprocessing, **Pandas** and **NLTK** have been used.
iii. **Python**, **Flask**, **Scikit Learn** use for building Machine Learning Web Application.
iv. Front End Development is done using HTML, CSS, Javascript, JQuery.
v. **GitHub** has been used as the version control system

2.7 Constraints

The Book Recommendation System must be user friendly, well maintainable.

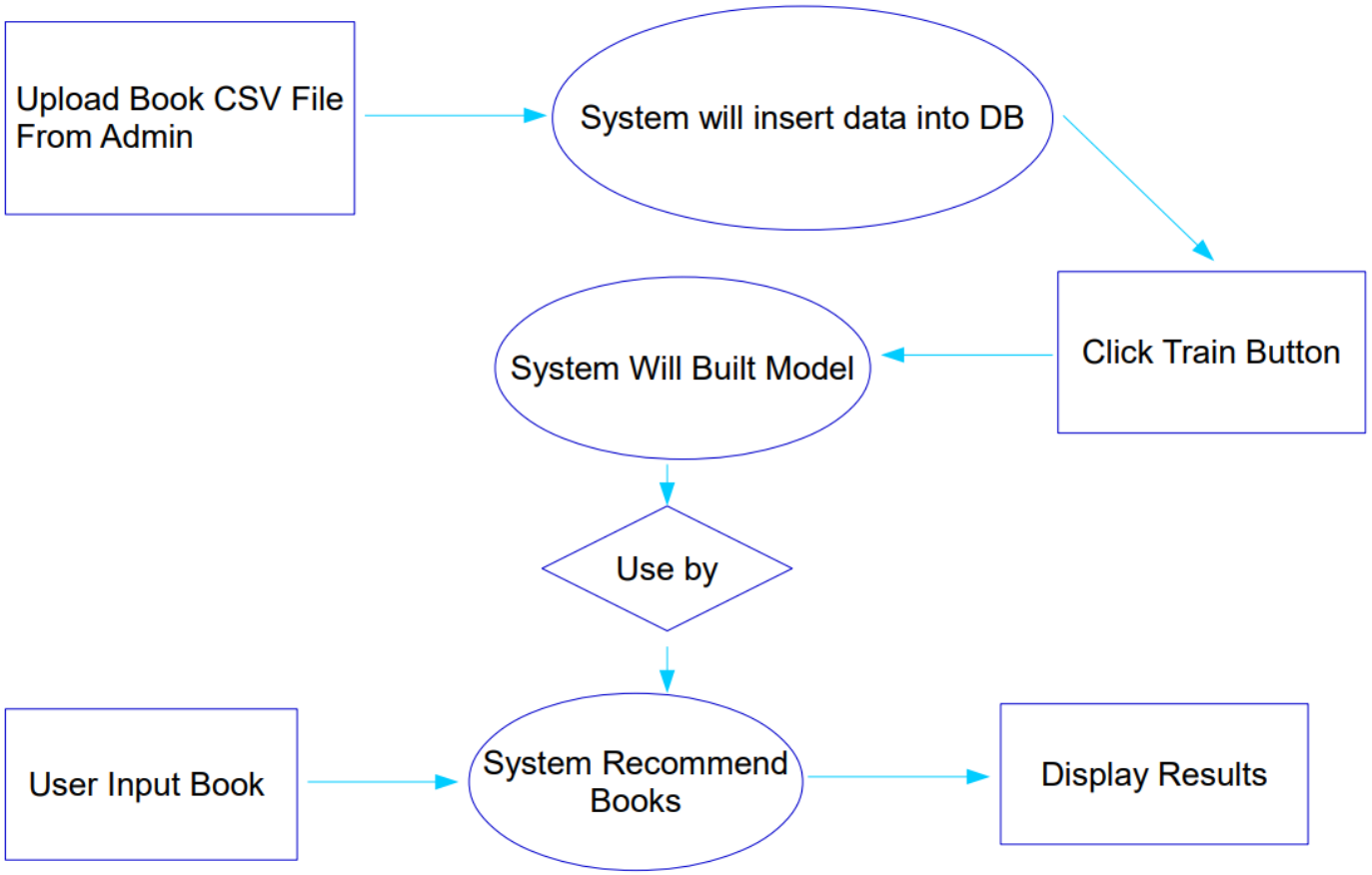
2.8 Assumptions

The main objective of the project is to implement use cases of book recommendation systems, and finding improvement strategies.

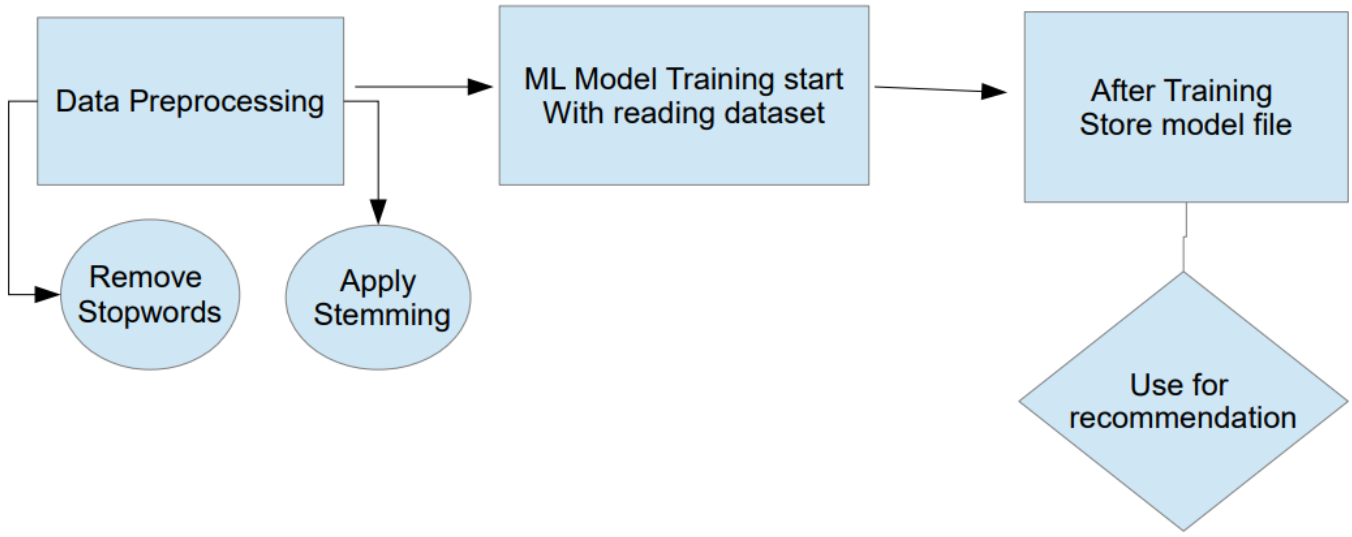
3 Design Details

3.1 Application Process Flow

For the book recommendation system we will use a machine learning based model. Below is the process flow diagram.



3.2 Model Training



3.3 Event Logs

The System should log every event so that the user will know what process is running internally.

Initial Step-By-Step Description

1. The System identifies at what step logging required
2. The System should be able to log each and every system flow.
3. Developers can choose logging methods. You can choose database logging/ File logging as well.
4. System should not hang even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

3.4 Error Handling

Should errors be encountered, an explanation will be displayed as to what went wrong? An error will be defined as anything that falls outside the normal and intended usage.

4 Performance

Evaluate performance of model for book recommendation system.

4.1 Reusability

The code written and the components used should have the ability to be reused with no problems.

4.2 Application Compatibility

The different components for this project will be using Python as an interface between them. Each component will have its own task to perform, and it is the job of the Python to ensure proper transfer of information.

4.3 Resource Utilization

When any task is performed, it will likely use all the processing power available until that function is finished.

4.4 Deployment

5 Conclusion

To recommend books based on user interest, build Machine Learning Model by doing training on more and more varieties of books, and test it and use it for users.