

**Demographically-Enhanced Movie Recommendation System for  
Personalized Book Suggestions in Big Data**

**Project Plan**

**Version 1.0**

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

## Table of Contents

1.	Introduction	3
1.1	Purpose of this document	3
1.2	Intended Audience	4
1.3	Scope	4
1.4	Definitions and acronyms	5
1.4.1	Definitions	5
1.4.2	Acronyms and abbreviations	5
1.5	References of dataset	6
2.	Background and Objectives	6
3.	Organization	7
3.1	Project group	7
3.2	Customer	7
4.	Development process	8
5.	Deliverables	8
6.	Project risks	8
7.	Communication	9
7.1	Canvas	9
7.2	Git	9
8.	Project plan	9
8.1	Time schedule	10
8.2	Test plan	10
9.	References	11

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

## 1. Introduction

### 1.1 Purpose of this document

The purpose of this document is to provide a description of the application **Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestion in Big Data**.

This project offers recommendations based on the user's tastes and preferences. By utilizing AWS's tools and technology, a recommendation engine is built that suggests movies and books. This document includes details about organization, roles, deliverables, project risks, time plans and financial plans.

### 1.2 Intended Audience

This document shall be used in all phases of the project as a guideline. Intended audiences of this project are all Project Stakeholders:

1. Project supervisor
2. Project leader
3. Team members
4. Movie enthusiasts
5. Book enthusiasts

### 1.3 Scope

An extensive and intricate project called "Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestion in Big Data" is being done. It specifically targets the junction of movie preferences and book suggestions based on demographic data, aiming to improve the personalization of content recommendations by utilizing big data analytics.

### 1.4 Definitions and acronyms

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

#### **1.4.1 Definitions Table Showing The Keywords Used In The Project**

<b>Keyword</b>	<b>Definitions</b>
Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestion in Big Data	The name of the project
Project Supervisor	A person in charge of supervising the project
Project Leader	A person in charge of organizing the team and communicating with the project supervisor
Team Member	An active member of the team responsible for making the job done
Milestone	A time in a project that marks the end of a project phase or the completion of an important deliverable.
Git	Version control system that will be used in this project
Product owner	Responsible for product management and its quality

#### **1.4.2 Acronyms and abbreviations**

N/A

#### **1.5 References of dataset:**

1. [https://www.tensorflow.org/datasets/catalog/movie\\_lens#movie\\_lens1m-ratings](https://www.tensorflow.org/datasets/catalog/movie_lens#movie_lens1m-ratings)
2. <https://www.kaggle.com/datasets/justinnguyen0x0x/best-books-of-the-21st-century-dataset>

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

## 2 Background and Objectives

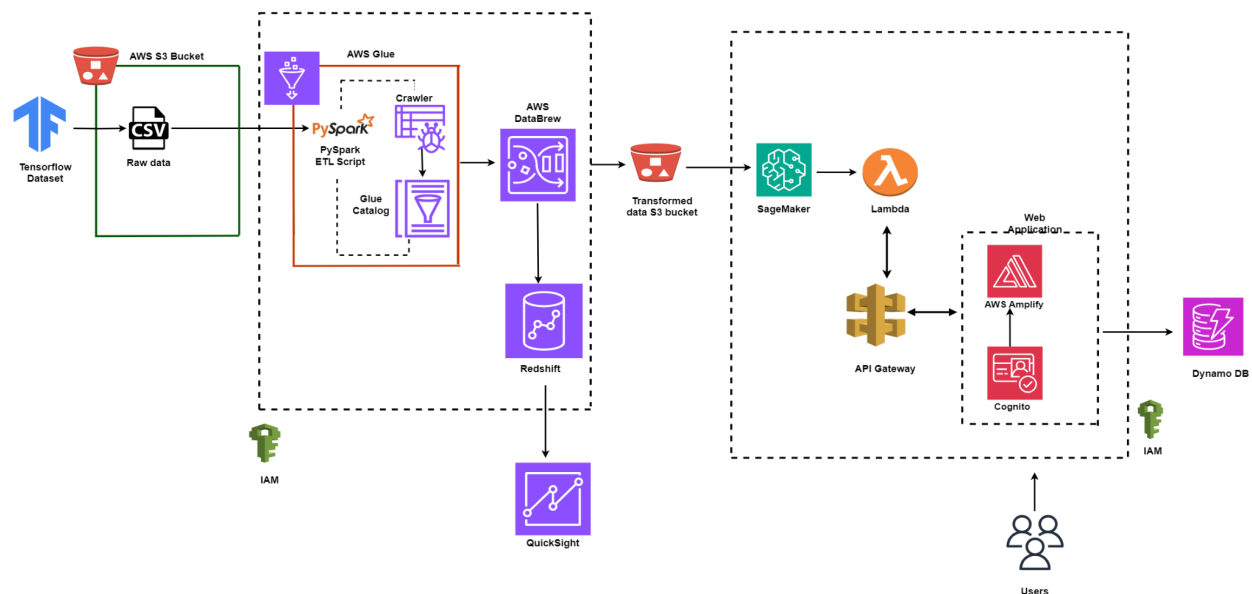
In the current digital world, the overflow of data poses a challenge for users to scroll through so much content, which leads to information overload. There are very few recommendation engines that are tailored to the user's preferences, especially in the combination of movies and books. Traditional systems do not use the demographics to provide suggestions.

Our motivation to build this project comes from recognizing that personalized recommendations are more significant now than ever. User's need help to navigate through the content available to them that aligns with their interests. By leveraging Big data and AWS tools, we aim to address the above mentioned problem by giving recommendations for both movies and books. This improves user experience and gives a solution to the problem of "too much information". We want to bridge the gap between the overflow of data and personalized recommendations.

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

### 3 Architecture & High Level Design

*Figure showing the architecture of the model*



#### 3.1 Organization

FA23: DATA-228 Sec 12 - Big Data Tech and App

#### 3.2 Project group

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

*Table listing the names of team members and their respective roles*

Name	Responsibility (roles)
Priya Varahan	ETL,PySpark script,AWS
Priyanka Akula	Python Analysis, Cleaning, AWS
Pooja Manjunatha	AWS (S3, SageMaker,Lambda)
Shreya Chikatmarla	AWS( QuickSight)
Nandini Sreekumaran Nair	AWS (Amplify, Cognito, API Gateway)

### 3.3 Customer

The target customers are listed below:

1. Supervisor
2. Project Team
3. Book Enthusiast
4. Movie Enthusiast

## 4 Development process

The project's and workflow are structured as follows:

### 4.1 Data Preparation and Cleaning:

Python: Utilization of PySpark scripts for the initial cleaning and preprocessing of network slice data.

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

## 4.2 Cloud Infrastructure and Services:

### Amazon Web Services (AWS)

AWS Cognito: Establishing user authentication mechanisms to ensure secure access to the project's resources.

AWS Amplify: Development of the frontend interface to interact with the network slicing system.

API Gateway: Creating and managing the backend APIs that the frontend will consume.

AWS Lambda: Serverless computing services to execute backend processes triggered by API calls.

Amazon SageMaker: Implementing and training machine learning models for network slice recognition and allocation.

AWS Identity and Access Management (IAM): Managing permissions and connecting components securely within AWS.

Amazon DynamoDB: Storing and retrieving the preprocessed data required for machine learning model training and inference.

## 4.3 Visualization and Interface:

Quick Sight: Building interactive dashboards for visualizing network slicing data and insights.



Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

HTML: Crafting the web application's markup structure for displaying information and receiving user input.

## 5 Deliverables

*Table showing the deliverables*

Deliverables	Output	Deadline
Abstract	Abstract and Data Source Selection	Sep 3, 2023
Design Plan	Data Cleaning and System Architecture	Oct 12, 2023
Coding and Testing	ETL Pipeline	Nov 29, 2023
Documentation	Data documentation and Visualization	Dec 10, 2023

## 6 Project risks

The development of a Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestion in Big Data faces significant risks

- 1.Data privacy and security concerns, especially when handling personal and demographic information.
- 2.Biases or inaccuracies in the dataset can adversely affect the system's recommendation

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

accuracy.

3. Technical challenges in scalability and integration with existing platforms present considerable hurdles.

## 7 Communication

The communication and collaboration for the project were facilitated through a variety of platforms, each serving different purposes to ensure a seamless workflow

*Table Showing Platform And Purpose Used For Communication*

Platform	Purpose
Zoom	Formal meetings, presentations, and discussions with stable video conferencing capabilities.
Google Meet	Quick, impromptu calls, and integration with Google Workspace tools.
In-Person Collaboration	Hands-on sessions, brainstorming, face-to-face interaction for more effective communication.
Discord	Day-to-day communication, informal chats, file sharing, continuous communication.

### 7.1 Collaboration

Table below provides a clear overview of the collaboration tools and their specific functions in the project's workflow.

Table discussing platform and purpose of their use

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

Platform	Purpose
AWS Management Console	Central hub for managing and monitoring the Amazon Web Services infrastructure, enabling the team to access, configure, and control AWS services.
Collab	Could refer to various collaboration platforms or tools. Google Colab, used for collaborative coding in Python for data analysis and ML model development. Otherwise, it represents general collaborative efforts among team members.
Google Docs	Creating, editing, and sharing documents in real-time, facilitating simultaneous contributions to text-based content like project plans, reports, and notes.
Google Slides	Developing presentations collaboratively, allowing multiple team members to contribute to slide creation, content refinement, and presentation design.

## 7.2 Git

All source code and finished documentation will be uploaded to Github repository.

Repository URL: <https://github.com/Priyankaakula/Medusa/new/main>

## 8. Project plan

### 8.1 Time schedule

*Table showing the milestones*

Id	Milestone	Responsible /Roles	Finished week
1	Abstract	Team	Sept 3

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

2	Design	Team	Oct 12
3	ETL- AWS Glue	Priya, Pooja	Oct 15
4	Analysis- Python	Priyanka	Oct 20
5	Cleaning- Python	Priya	Oct 30
6	Code Testing	Nandini, Priyanka	Nov 15
7	Visualization	Shreya	Nov 20
8	Website	Pooja , Nandini	Nov 27
9	Presentation	Team	Nov 29
10	Documentation	Team	Dec 10

## 8.2 Test plan

### Unit Testing

- Test individual modules or components for functionality.

### Integration Testing

- Test the integration of different modules and services (AWS Lambda, API Gateway).
- Validate data flow and interaction between components.
- Ensure APIs and endpoints are correctly functioning.

### System Testing

Demographically-Enhanced Movie Recommendation System for Personalized Book Suggestions in Big Data	Version: 1.0
Project Plan	Date: 12/10/ 2023

- Test the complete and integrated system.
- Validate end-to-end functionality, including front-end and back-end integration.
- Check system behavior under various scenarios and data sets.

## Performance Testing

- Assess the system's performance under different loads (e.g., multiple simultaneous requests).
- Test scalability and response time of the machine learning models.
- Ensure the system meets performance benchmarks.

## 9. References

1. <https://github.com/Priyankaakula/Medusa/blob/main/README.md>
2. <https://us-east-2.quicksight.aws.amazon.com/sn/dashboards/0b1b3d74-f84b-4e13-af08-4c ff4d4fcbfe>
3. <https://dev5777.d3sksph6uzn2dy.amplifyapp.com>
4. <https://app.diagrams.net/?src=about>
5. [https://www.tensorflow.org/datasets/catalog/movie\\_lens#movie\\_lens1m-ratings](https://www.tensorflow.org/datasets/catalog/movie_lens#movie_lens1m-ratings)
6. <https://ianv3z58mg.execute-api.us-east-2.amazonaws.com/cinereads>