Project Planning & Management Document

Project Name: Personalized Recommendation System with Feedback Loop

Platform/Environment: Microsoft Azure

---

1. Project Proposal

Overview

In today's digital landscape, personalized recommendations are the cornerstone of user engagement and retention. This project aims to build a scalable, real-time personalized recommendation engine tailored for e-commerce or content platforms. Inspired by industry leaders, the system will not only generate highly relevant suggestions but also adapt dynamically to user feedback. The final solution will be hosted on Azure, complete with CI/CD, MLOps, and automated retraining pipelines.

Objectives

- **Generate Tailored Suggestions:** Develop high-performing collaborative, content-based, or hybrid models to match users with relevant items.
- **Real-Time Adaptation:** Create a feedback loop that adjusts recommendations based on live user interactions (clicks, purchases, ratings).
- **Scalable Deployment:** Containerize and deploy the model as a highly available REST API on Azure.
- **Automated MLOps:** Implement experiment tracking, real-time monitoring, and scheduled retraining to prevent model degradation.

Scope

- **In-Scope:** Data aggregation, exploratory data analysis (EDA), model training (Collaborative Filtering, Matrix Factorization), API development (FastAPI/Flask), Dockerization, Azure ML deployment, and MLflow integration.
- **Out-of-Scope:** Front-end UI/UX development, complex multi-region global deployment (beyond the initial MVP scale), and building the source platforms generating the raw data.

---

2. Project Plan

Timeline & Milestones (Gantt Chart Representation)

Assuming a standard 10-week project lifecycle.

Week	Phase	Milestone / Tasks	Deliverables
1-2	Phase 1	**Data Collection & Preprocessing:** Aggregate logs, clean data, create interaction matrices, perform EDA.	Cleaned Dataset, Preprocessing Doc, EDA Report
3-5	Phase 2	**Model Development:** Build CF/Content-based models, tune hyperparameters, evaluate (NDCG, MAP).	Trained Models, Evaluation Report
6-7	Phase 3	**Cloud Deployment:** Containerize with Docker, deploy to Azure ML, build FastAPI/Flask endpoints.	Deployed API, Integration Docs
8	Phase 4	**MLOps & Monitoring:** Setup MLflow, integrate real-time logging, schedule automated retraining.	MLOps Doc, Monitoring Setup, Retraining Pipeline
9-10	Phase 5	**Documentation & Handover:** Finalize	Final Report, Presentation Deck,

		reports, prepare presentation deck, conduct live demo.	Future Roadmap

Resource Allocation

- **Compute/Infrastructure:** Azure Machine Learning, Azure Container Registry, Azure VMs (for hosting API/MLflow).
- **Development Tools:** Python, scikit-learn/PyTorch, FastAPI, Docker, MLflow, Git/GitHub.
- **Human Resources:** 4-5 cross-functional team members (roles defined below).

3. Task Assignment & Roles

To ensure accountability, project responsibilities are divided among the following roles:

Role	Responsibilities	Key Milestones
Project Manager / Lead	Oversees timeline, manages risks, coordinates documentation, and leads the final presentation.	All Milestones
Data Engineer	Aggregates data, builds data pipelines, cleans/normalizes features, and structures interaction matrices.	Milestone 1
Data Scientist / ML Engineer	Conducts EDA, trains recommendation models, performs hyperparameter tuning, and evaluates metrics.	Milestones 1, 2
DevOps / Cloud Engineer	Wraps the model in a REST API, writes Dockerfiles, and manages Azure ML deployments.	Milestone 3

MLOps Engineer	Sets up MLflow tracking, configures system logging, and builds the automated retraining pipelines.	Milestone 4

4. Risk Assessment & Mitigation Plan

Risk / Threat	Probability	Impact	Mitigation Strategy
Cold Start Problem (New Users/Items)	High	High	Implement hybrid models that fall back on content-based filtering or popular trends until enough data is gathered.
High API Latency	Medium	High	Optimize API using FastAPI. Implement caching (e.g., Redis) for frequently requested recommendations.
Concept Drift (Changing User Habits)	High	Medium	Rely on Milestone 4 deliverables: use automated retraining pipelines triggered by performance drops or scheduled weekly.
Cloud Budget Overruns	Medium	Medium	Set strict Azure billing alerts. Automatically spin down compute clusters when not

			actively training models.
Data Sparsity Issues	Medium	High	Use Matrix Factorization techniques or incorporate implicit feedback (browsing time, clicks) alongside explicit ratings.

---

5. KPIs (Key Performance Indicators)

To measure the success of the project, we will track metrics across three main categories:

Model Performance KPIs:

- **Precision@K / Recall@K:** To measure the relevance of the top *K* recommended items.
- **NDCG (Normalized Discounted Cumulative Gain):** To evaluate the ranking quality of the recommendations.
- **MAP (Mean Average Precision):** To measure the overall accuracy across all users.

System Operational KPIs:

- **API Response Time (Latency):** Target < 100ms at the 95th percentile to ensure a seamless user experience.
- **System Uptime/Availability:** Target 99.9% uptime for the recommendation API.
- **Pipeline Execution Success Rate:** Ensure the automated data preprocessing and retraining pipelines run without failure > 95% of the time.

Business / User Interaction KPIs:

- **User Adoption/Click-Through Rate (CTR):** The percentage of recommended items that users actually click on.
- **Conversion Rate:** The percentage of clicked recommendations that lead to a tangible business event (e.g., a purchase or long-term viewing).