

Assignment 2: Personalized Course Recommendation Engine

1. Background & Context

Online learning platforms host thousands of courses across domains—learners often feel overwhelmed by choices. A personalized recommender that understands both course content and individual learner profiles can boost engagement and completion rates by suggesting the most relevant next steps.

2. Problem Statement

“Design and implement a Course Recommendation Engine that—given a user query (completed courses + a short interests blurb)—returns the top-5 most relevant courses from a catalog of course offerings, using embedding models and a vector database for semantic matching.”

3. Objectives & Learning Outcomes

- **Embeddings & Semantic Search:** Convert course descriptions into high-dimensional vectors.
- **Vector Database:** Index and query vectors for fast similarity retrieval.
- **Recommendation Logic:** Rank courses by cosine similarity to the user query vector.
- **Basic UI/CLI (optional):** Demonstrate the engine end-to-end with sample queries.

4. Technical Requirements

Data Ingestion & Indexing: Read the course catalog (title + description), compute embeddings, and upsert into a vector DB.

Recommendation Logic & API:

```
def recommend_courses(profile: str, completed_ids: List[str]) ->
List[Tuple[str, float]]:
    """
    Returns a list of (course_id, similarity_score) for the top-5
    recommendations.
    """
```

5. Deliverables

- Code: Jupyter notebook file in PDF format OR .py files in zip (keep requirements .txt if needed)
- Evaluation Report: Jupyter Notebook should include test results for each of 5 test profiles, list recommendations and comment on relevance.

6. Dataset

Dataset file: assignment2data.csv

<https://raw.githubusercontent.com/Bluedata-Consulting/GAAPB01-training-code-base/refs/heads/main/Assignments/assignment2dataset.csv>

7. Sample Input Queries

- “I’ve completed the ‘Python Programming for Data Science’ course and enjoy data visualization. What should I take next?”
- “I know Azure basics and want to manage containers and build CI/CD pipelines. Recommend courses.”
- “My background is in ML fundamentals; I’d like to specialize in neural networks and production workflows.”
- “I want to learn to build and deploy microservices with Kubernetes—what courses fit best?”
- “I’m interested in blockchain and smart contracts but have no prior experience. Which courses do you suggest?”