Intern Assignment – Product Recommendation/Analytics Web App

You are required to build a complete ML-driven web application over a span of **2 days**. This project will test your ability to combine multiple AI domains (ML, NLP, CV, GenAI) with full-stack development and analytics visualization.

Project Overview

Build a **Product Recommendation Web App** that uses AI to recommend furniture products and generate creative product descriptions. The application should be developed using **FastAPI (backend)** and **React (frontend)**, integrating with a **vector database**.

- The recommendation page will be a like back and forth conversation page.
- The analytics page will be on a different route with analytics of all the items in the database.

Dataset Information

A dataset will be provided to you with the following columns:
- title
- brand
- description
- price
- categories
- images
- manufacturer
- package dimensions
- country_of_origin
- material
- color
- uniq_id

You are free to tweak or augment the data as needed,
but do **not remove or drop** any important columns.

Dataset Link:
https://drive.google.com/file/d/1uD1UMXT2-13GQkb_H9NmEOyUVI-zKyl6/view?usp=sharing

Project Requirements

1. Machine Learning (ML): Build a product recommendation model based on the given dataset. You may use any techniques you may find fit for this.
2. Natural Language Processing (NLP): Implement text analysis features to process products and group similar or related products.
3. Computer Vision (CV): Develop an image classification model that can identify product categories/types from the provided product images.
4. Generative AI (GenAI): Use a lightweight GenAI model (your choice) to generate creative product design descriptions to accompany recommended products on the frontend.
5. Vector Database: Store embeddings (text/image-based) in a vector database for semantic search and retrieval. You can use any vector database of your choice (Pinecone preferred).
6. Frontend (React): Create a simple yet functional frontend to display product recommendations once someone sends a prompt. These recommendations should come with generated descriptions and product images..
7. Analytics Page: Add a page (using React routing) showing analytics on the current dataset. You may use a different database for analytics if you wish.

Tech Stack Requirements

- Backend: FastAPI
- Frontend: React (any library/framework of your choice for UI)
- Vector DB: Pinecone (preferred) or any other vector DB
- ML Models: Any framework (e.g., scikit-learn, PyTorch, TensorFlow)
- NLP: Any model/tool (e.g., spaCy, HuggingFace Transformers)
- CV: CNN/ResNet/Vision Transformer/GenAI model (your choice)
- GenAI: Any open-source or API-based lightweight model
- Integration Framework: LangChain (must be used for GenAI or embedding-based tasks)

Deliverables:

You must submit a GitHub repository containing:

1. Frontend (React App)
2. Backend (FastAPI App)
3. Data Analytics Notebook (.ipynb) — must include:
   - Clear and descriptive comments explaining your reasoning for each major step
4. Model Training Notebook (.ipynb) — must include:
   - Clear and descriptive comments explaining your reasoning for each major step
   - Model performance evaluation (if applicable)
5. Instructions (README.md) — detailing setup, usage, and environment requirements.

Duration

Total completion time: 2 days from the time of assignment.

Tips & Notes

- Focus on clarity, modularity, and reasoning.
- Use mock data if needed to demonstrate model flow.
- Creativity in GenAI output and analytics design will be rewarded.
- Comment your notebook well to explain your approach.

Submission

Submit the GitHub repository link with:
- React frontend
- FastAPI backend
- Model training notebook
- README instructions

- All components must be working together end-to-end.