VertxAI - AI/ML Engineer Assignment

Assignment 1: Founder-Investor Matching AI Model

Objective

Develop an AI model that takes structured startup data and investor preferences as input and provides a **match score** between founders and investors.

Dataset (Sample Data Provided)

You will work with a sample dataset containing:

- **Founder Information:** Industry, startup stage, funding required, traction, business model.
- **Investor Preferences:** Preferred industry, investment range, etc.

Tasks

1. **Model Implementation Using Gemini API:**
    - **Utilize Gemini API to analyze and process investor-founder compatibility.**
    - **Ensure proper API integration and response handling.**
2. **Match Score Calculation:**
    - **Extract relevant insights using Gemini API and compute a compatibility score.**
    - **The score should reflect how well an investor's interest aligns with a startup's profile.**
3. **Output:**
    - **Return a list of all investors matching the founder's profile along with a match score for each.**
    - **Display ranked results in a structured format.**

Deliverables:

- A **Python notebook / VS code(.ipynb)** with code and explanations.
- A **README file** explaining your approach and improvements you'd suggest.

Assignment 2: AI Pitch Analysis Model

Objective

Develop an **LLM-powered pitch analysis model** that evaluates a startup's pitch deck and provides a **pitch score, strengths, and weaknesses**.

Dataset & Inputs

- A **set of sample pitch decks**
- AI model must extract and analyze key sections: **Problem, Solution, Market, Business Model, Financials, Team**.

Tasks

1. **Text Extraction & Preprocessing:**
    - Use **OCR or PDF parsing** to extract text.
    - Preprocess text by removing unnecessary elements and formatting.
2. **Feature Engineering:**
    - Identify key sections from a pitch deck.
    - Assign weights based on the importance of different sections.
3. **Scoring Model:**
    - Use **LLM-based evaluation (GPT/Gemini API or fine-tuned BERT model)** to analyze the quality of each section.
    - Generate a **pitch score (0-100)** based on predefined metrics.
4. **Strength & Weakness Analysis:**
    - Provide **personalized feedback** on which areas need improvement.
    - Suggest content improvements or additional data needed.
5. **Output:**
    - Display **pitch score + AI-generated feedback**.

Deliverables:

- Python code **Or VS code** in **.ipynb format**.
- Sample pitch analysis results for at least **3 different pitch decks**.
- A **report (PDF or Markdown)** explaining methodology and insights.

Submission Guidelines

- **Deadline:** Submit both assignments within **2-3 days**.
- **Format:**
    - Code: GitHub links with a brief explanation of the assignments.
    - Results: CSVs, JSON, or screenshots where applicable.
- **Evaluation Criteria:**
    - Model accuracy and relevance.
    - Code quality and documentation.
    - Creativity and problem-solving approach.

Bonus (Optional, Extra Points)

If you'd like to go beyond the basic requirements, consider:

- Creating a **basic UI** for interactive model testing.
- **Deploy this using Flask** (higher priority)

Contact & Support

If you have any questions, feel free to reach out. We're excited to see your approach to these real-world AI challenges!

Best of luck, **Surya Prakash** CTO, VertxAI