


Stochastic Inc. Competency Assessment

: Junior AI Engineer

Task Overview

Develop a lightweight AI agent based on open LLM APIs and demonstrate integration with enterprise performance features.

 You are free to utilize AI tools, but DO NOT simply copy and paste the AI generated contents without your thoughts and insights. It will be automatically disqualified.

Task 1: Building an Enterprise-Ready AI Agent Prototype

Objective: To evaluate the candidate's ability to build and integrate AI agents using Large Language Model (LLM) APIs while ensuring alignment with enterprise-grade optimizations and user needs.

Task Details

Develop a Document Q&A AI Agent

- Set up the development environment using Python and necessary libraries (e.g., OpenAI or Gemini API SDKs).
- Develop a document ingestion pipeline with multi-modal LLM:
 - Handle multiple PDF documents.
 - Extract meaningful text and structure (titles, abstracts, sections, tables, and references).
 - Ensure high accuracy in content extraction, preserving equations, figures, and tables where possible.
- Implement an NLP-powered interface that allows users to query extracted content. The interface should support the following functionalities:
 - Direct content lookup (e.g., "What is the conclusion of Paper X?")
 - Summarization of key insights (e.g., "Summarize the methodology of Paper C. ")
 - Extraction of specific evaluation results (e.g., "What are the accuracy and F1-score reported in Paper D?")
- (Bonus Point) Add a functional calling capability to the AI agent, for it to leverage the Arxiv API to look up a paper based on the user's description.

Submission Format:

- Provide a **GitHub repository link** containing a README file with setup instructions, the developed Python code, and a video demo of your implementation.
- **Video recording File** sharing your screen and presenting ins and outs of the

deliverables.



Evaluation Criteria:

- Technical proficiency in working with LLM APIs and Python development.
- Integration of enterprise-grade features such as context handling and response optimization.
- Clear and well-documented code following basic programming and security standards.