**ML Engineer – Analytics Assignment**

You are tasked with building a customer service chatbot for a fictional insurance company, **Nexora Pty Ltd**, using a **Retrieval-Augmented Generation (RAG)** approach. This assignment aims to assess your ability to design and implement an AI application with MLOps considerations. You are expected to analyze the provided datasets, demonstrate the chatbot via a simple web app, and outline your implementation and deployment strategy.

You must prepare a short presentation outlining the implementation, key features, and deployment strategy. You may use any AI model or tools for the development and present findings in Word, PowerPoint, Notebooks, or any other preferred format.  
  
**Datasets**

You have been provided with three datasets:

* **products\_occupation.json**  
  Contains details of six insurance product types and occupation-based risk data.
* **faqs.csv**  
  Frequently asked questions and answers about Nexora’s products and services.
* **chat\_conversations.json**  
  Real-world customer service interactions with annotated intents, entities, and outcomes.

Please use the FAQs and Insurance Product data as the knowledge base for answering user queries, while the chat conversations dataset will serve as a reference for outlining the fine-tuning process of an LLM.

**Assignment Task**

1. **Build a Customer Service Chatbot** **(Practical)**

* **Develop a question-answering chatbot** using a **Retrieval-Augmented Generation (RAG)** approach with the provided datasets
* Create a simple web interface to demonstrate the chatbot functionality
* Implement basic intent recognition and entity extraction to support relevant query handling.

1. **Conceptual Tasks – No Practical Implementation Required**

* **MLOps Pipeline**: Present your conceptual approach for an end-to-end MLOps pipeline including application lifecycle management, monitoring, and deployment. Optional local deployment is allowed but will not be assessed.
* **Fine-Tuning Strategy**: Outline a high-level plan to fine-tune a language model using the conversation data and highlight its potential benefits over a base LLM using RAG.