Take home - AI engineer

Goal: To assess your ability as an AI engineer to code and evaluate machine learning pipelines.

Data:

https://www.nrma.com.au/sites/nrma/files/nrma/policy_booklets/nrma-car-pds-1023-east.pdf

https://www.allianz.com.au/openCurrentPolicyDocument/POL011BA/$File/POL011BA.pdf

Task: Create at least 2 different RAG pipelines for the provided data, and then compare and evaluate the performance between the pipelines.

1. Create the retrieval frameworks. i.e. could be hybrid search vs vector search. vector search using openai embeddings vs cohere embeddings
2. Create the generation frameworks. i.e. could be out of the box GPT vs fine tuned Mistral.
3. Create an evaluation framework to compare and evaluate the performance between the pipelines.

Ideal frameworks to utilise:

- huggingface, langchain, llamaindex, numpy, pytorch, etc.

We are not assessing the accuracy of these pipelines, but instead your ability to:

- To write good code
- To learn new concepts
- To research and bring scientific evaluation to models and processes