

UDACITY

Introduction to Generative AI with AWS Project Documentation Report

Visit [UDACITY Introduction to Generative AI with AWS Project Documentation Report](#) to make a copy of this document.

Complete the answers to the questions below to complete your project report. Create a PDF of the completed document and submit the PDF with your project.

Question	Your answer:
Step 2: Domain Choice What domain did you choose to fine-tune the Meta Llama 2 7B model on? Choices: 1. Financial 2. Healthcare 3. IT	IT
Step 3: Model Evaluation Section What was the response of the model to your domain-specific input in the model_evaluation.ipynb file?	<pre>payload = { "inputs": "Traditional approaches to data management such as", "parameters": { "max_new_tokens": 64, "top_p": 0.5, "temperature": 0.5, "return_full_text": False, }, } try: response = predictor.predict(payload, custom_attributes={"prompt_id": "eval"}) print(response(payload, response)) except Exception as e: print(e)</pre> <p>Traditional approaches to data management such as relational databases have been unable to scale to the size and speed of modern data processing needs. A new type of database, called a NoSQL database, has emerged to meet the demands of modern data management. NoSQL databases are designed to handle large amounts of data and to operate at high speeds.</p> <p>.....</p> <p>The prompt is related to the domain you want to fine-tune your model on. You will see the outputs from the model without fine-tuning are limited in providing insightful or relevant content.</p> <p>Use the output from this notebook to fill out the "model evaluation" section of the project documentation report</p>
Step 4: Fine-Tuning Section After fine-tuning the model, what was the response of the model to your domain-specific input in the model_finertuning.ipynb file?	<pre>payload = { "inputs": "Traditional approaches to data management such as", "parameters": { "max_new_tokens": 64, "top_p": 0.5, "temperature": 0.5, "return_full_text": False, }, } try: response = finetuned_predictor.predict(payload, custom_attributes={"prompt_id": "eval"}) print(response(payload, response)) except Exception as e: print(e)</pre> <p>Traditional approaches to data management such as relational databases have been unable to scale to the size and speed of modern data processing needs. A new type of database, called a NoSQL database, has emerged to meet the demands of modern data management. NoSQL databases are designed to handle large amounts of data and to operate at high speeds.</p> <p>.....</p> <p>Do the outputs from the fine-tuned model provide domain-specific insightful and relevant content? You can continue experimenting with the inputs of the model to test it's domain knowledge.</p>