UDACITY

**Introduction to Generative AI with AWS**

**Project Documentation Report**

Visit [UDACITY Introduction to Generative AI with AWS Project Documentation Report](https://docs.google.com/document/d/1kqRy-gVGZjwl9r03hqMeWSm-D6hEY8KWuxz4GO0vdOw/copy) 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 | I have chosen IT domain to fine-tune Meta Llama 2 7B model.  I uploaded required dataset on S3 bucket for same and passed respective S3 URI in the notebook code to achieve same. |
| **Step 3: Model Evaluation Section**  What was the response of the model to your domain-specific input in the **model\_evaluation.ipynb file**? | Before fine-tuning, the model response was vague.  My input string was “A second important aspect of ubiquitous computing environments is”  And model response was  > the ability to communicate with other devices and systems. The mobile device of the present invention can communicate with other mobile devices and/or with stationary devices. The communication may occur over a wired or wireless connection.  In addition, the mobile device of the present invention may be used to communicate with other systems  Refer to Model\_Evaluation\_UdacityGenAIAWS HTML output |
| **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\_finetuning.ipynb file**? | After fine-tuning, the model response utilized the passed IT dataset information and gave more precise response.  My input string was “A second important aspect of ubiquitous computing environments is”  And model response was  > [{'generated\_text': ' the need to support multiple users in a single location. The goal of this research is to develop an architecture for multi-user ubiquitous computing environments. This architecture must support a rich set of applications that are suitable for multiple users in a single location. In addition, the architecture must support the dynamic addition and removal'}]  Refer to Model\_FineTuning HTML output |