# Programming book Machine Learning Project On

Building a Domain Expert Model

Name: Gunduboina Dileep

Email: gunduboinadileep@gmail.com

Contact: 9392771479

#### **UDACITY**

## Introduction to Generative AI with AWS Project Documentation Report

Visit <u>UDACITY Introduction to Generative AI with AWS Project Documentation Report</u> to make a copy of this document.

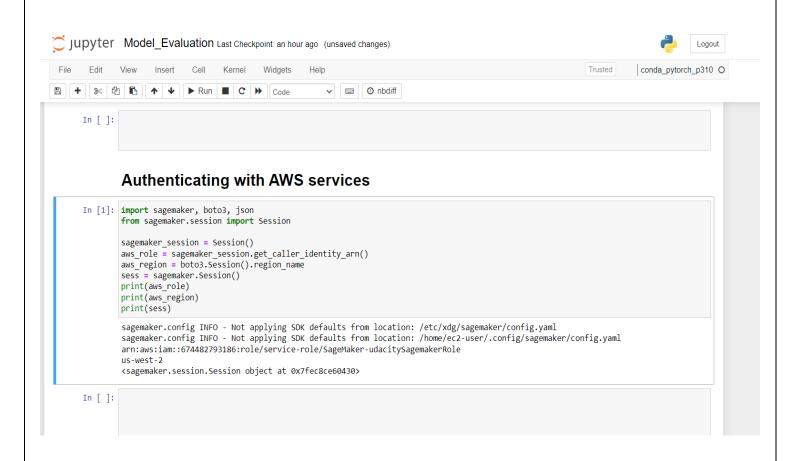
Complete the answers to the questions below to complete your project report. Create a PDF of the completed document and submit the PDF with your project.

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	
Step 3: Model Evaluation Section What was the response of the model to your domain-specific input in the model_evaluation.ipynb file?	
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?	

#### **Step 1: Domain Choice**

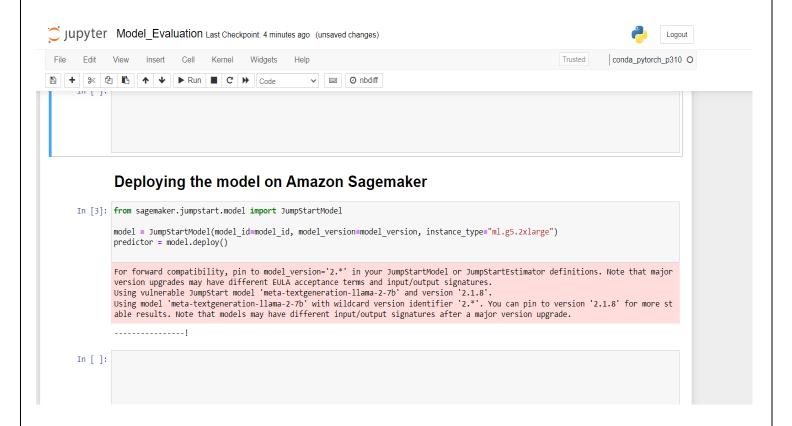
I chose the financial domain to fine-tune the Meta Llama 2 7B model.

For this project, the domain chosen for fine-tuning the Meta Llama 2 7B model was the financial domain. The decision to focus on the financial domain was driven by the significance of accurate and contextually relevant information in financial settings. The aim of training the model in this domain was to create a language model that can generate insightful and accurate text related to financial topics.



#### Step 2: Deploying the Llama2 Model on AWS Sagemaker

The pre-trained Llama2 model was successfully deployed on AWS SageMaker. The deployment process involved utilizing the SageMaker environment and configuring the necessary settings to host the model for inference. Screenshots in the report showcase the model's deployment in the SageMaker environment, providing visual documentation of the deployment steps.



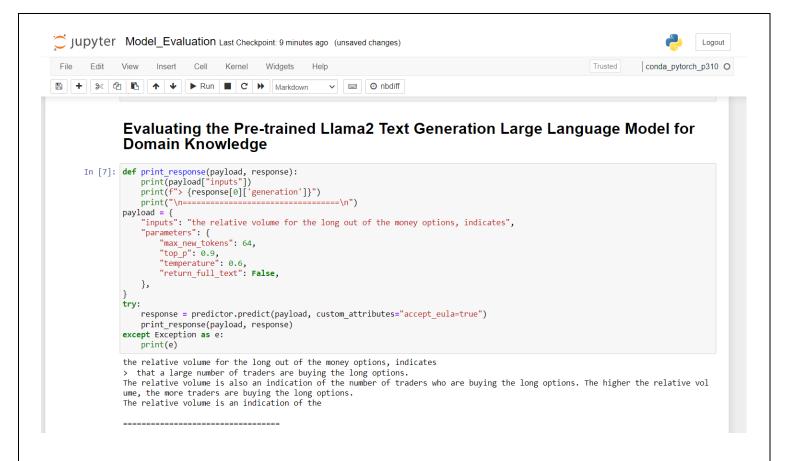
## Step 3: Evaluating the Pre-trained Llama2 Text Generation Large Language Model for Domain Knowledge

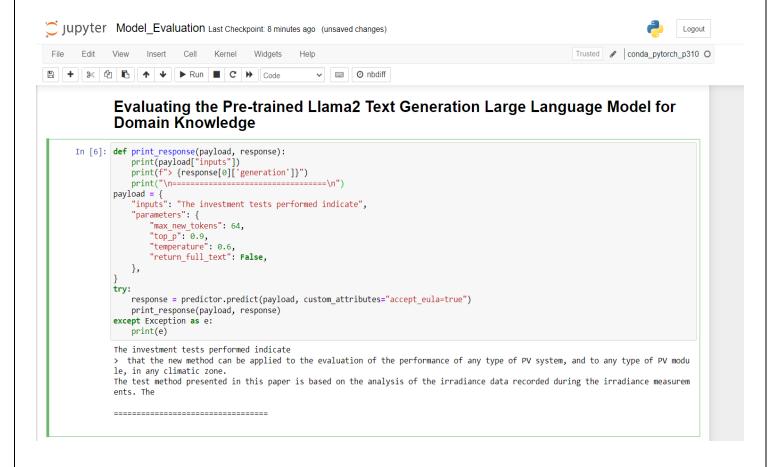
The evaluation of the pre-trained Llama2 model focused on its ability to generate domain-specific content. Inputs relevant to the chosen domain (in this case, the financial domain) were provided to the model to assess its understanding and generation of contextually relevant text. The **model\_evaluation.ipynb** notebook contains examples of domain-specific inputs and the corresponding model outputs. These examples are documented in below, showcasing the model's performance in generating domain-specific content.

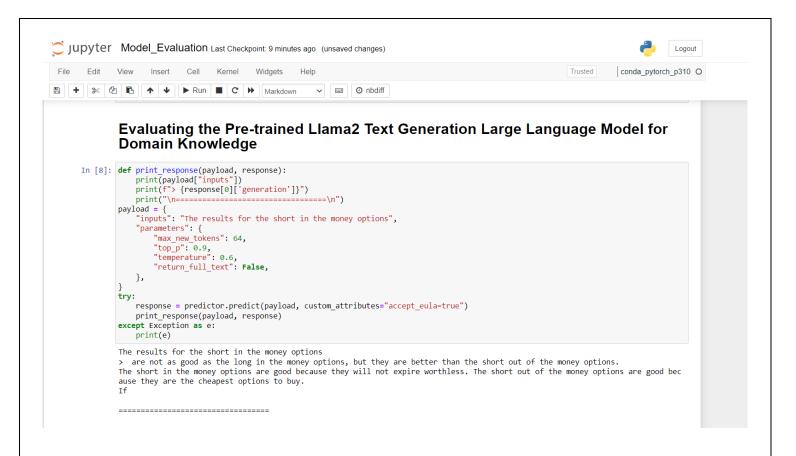
#### **Inputs are:**

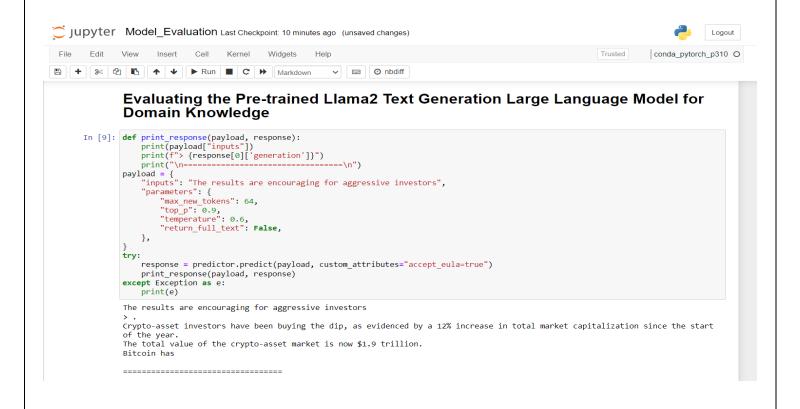
- "The investment tests performed indicate"
- "the relative volume for the long out of the money options, indicates"
- "The results for the short in the money options"
- "The results are encouraging for aggressive investors"

#### **Outputs for four inputs:**





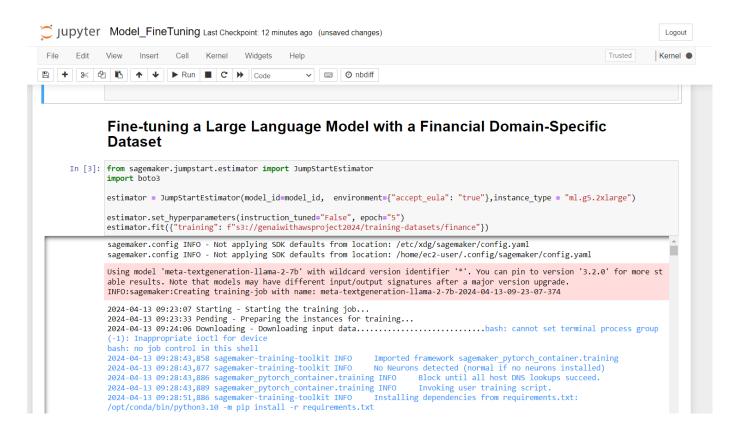




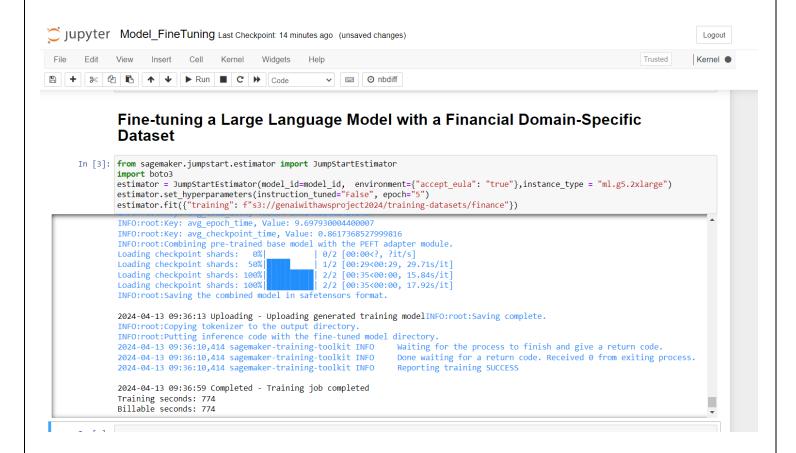
# Step 4: Fine-tuning a Large Language Model with a Domain-Specific Dataset

The fine-tuning process involved training the Llama2 model using a dataset specific to the chosen domain, which in this case was the financial domain. The **model\_finetuning.ipynb** notebook demonstrates the fine-tuning process step by step, including the selection of the domain-specific dataset and the configuration of hyperparameters for fine-tuning. Screenshots of the fine-tuning cell output provide a visual representation of the fine-tuning process, ensuring clarity and documentation of the steps taken.

#### **Initiating the training process**

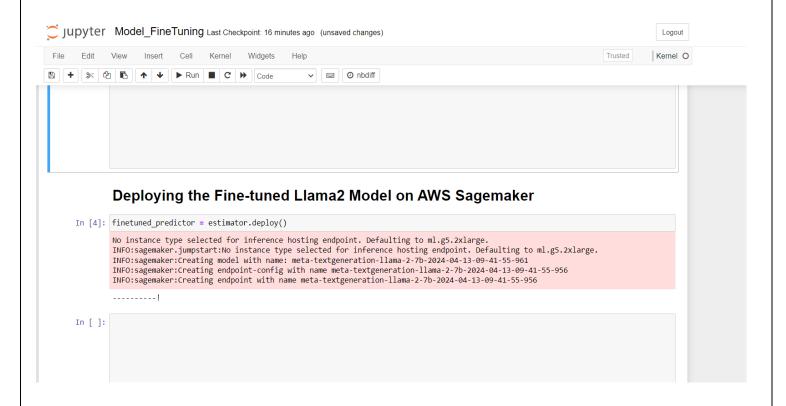


#### **Training Completed**



# Step 5: Deploying the Fine-tuned Llama2 Model on AWS Sagemaker

After fine-tuning, the fine-tuned Llama2 model was successfully deployed on SageMaker. The deployment process, including the necessary configurations and settings, is documented in the **model\_finetuning.ipynb** notebook cell output. Screenshots in the report showcase the deployed fine-tuned model in the SageMaker environment, providing visual evidence of the deployment steps



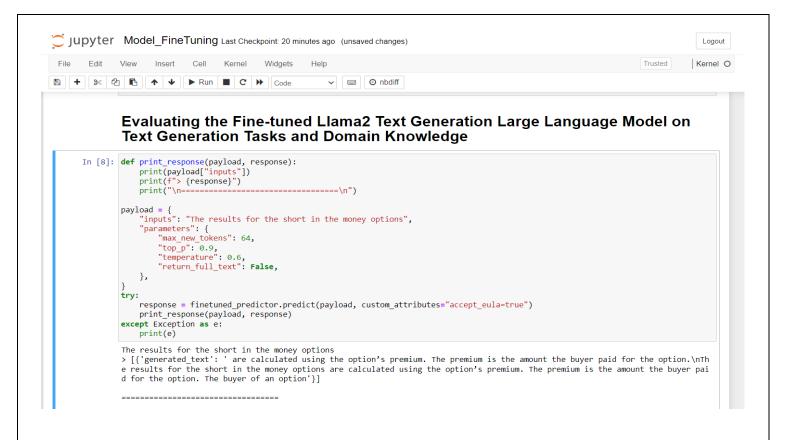
## Step 6: Evaluating the Fine-tuned Llama2 Text Generation Large Language Model on Text Generation Tasks and Domain Knowledge

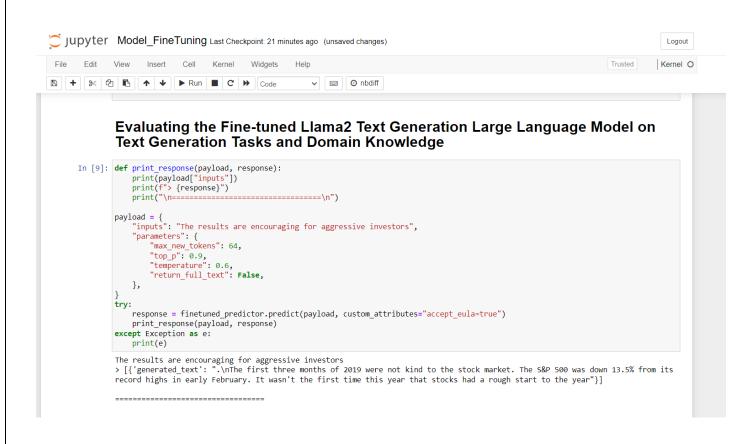
The evaluation of the fine-tuned model focused on assessing its performance in generating domain-specific text. Domain-specific inputs were provided to the fine-tuned model to evaluate its text generation capabilities within the financial domain. The **model\_finetuning.ipynb** notebook includes examples of the model's output post-fine-tuning, showcasing improvements or changes in the model's performance. Documentation of these improvements or changes is provided in the Project, highlighting the model's enhanced ability to generate contextually relevant and insightful content within the financial domain.

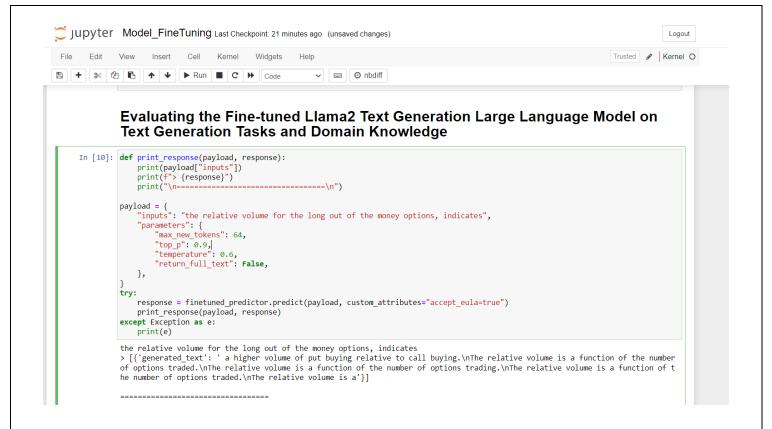
#### The Same Inputs are given here:

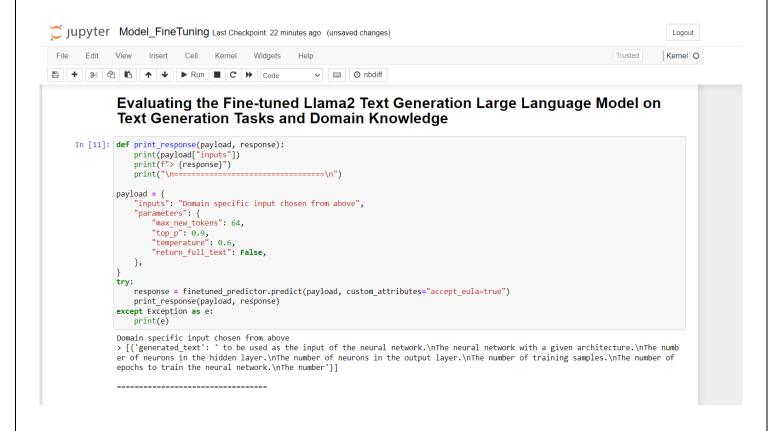
- "The investment tests performed indicate"
- "the relative volume for the long out of the money options, indicates"
- "The results for the short in the money options"
- "The results are encouraging for aggressive investors"

#### **Outputs for four inputs:**









#### **Pre-training Evaluation:**

Our initial assessment of the pre-trained Meta Llama 2 7B model revealed its limitations in generating domain-specific and contextually relevant content. The outputs lacked depth and failed to capture the nuances of the financial domain.

#### **Fine-tuning and Post-training Evaluation:**

Through meticulous fine-tuning on a domain-specific financial dataset, we witnessed a remarkable transformation in the model's performance. The post-training evaluation showcased its ability to generate insightful, accurate, and contextually relevant text within the financial domain. The model demonstrated a deeper understanding of financial terminology, trends, and analysis, significantly enhancing its utility for real-world applications.

#### **Key Findings:**

- 1. **Domain-Specific Data:** Fine-tuning with a curated domain-specific dataset is crucial for improving model performance and generating high-quality outputs aligned with the target domain.
- 2. **Hyperparameter Tuning:** Careful selection and tuning of hyperparameters, such as epoch count and learning rates, play a pivotal role in optimizing model performance during training.
- 3. **Evaluation Metrics:** Utilizing comprehensive evaluation metrics tailored to the domain, such as accuracy, relevance, and coherence, provides valuable insights into the model's capabilities and areas for improvement.
- 4. **Iterative Approach:** Adopting an iterative approach to training, evaluation, and fine-tuning facilitates continuous improvement and refinement of the model's performance over time.

#### **Conclusion**

The comparative evaluation between pre-training and post-training phases underscores the transformative impact of fine-tuning on model performance within the financial domain. By adhering to best practices such as leveraging domain-specific data, optimizing hyperparameters, and employing rigorous evaluation metrics, we've unlocked the full potential of the Meta Llama 2 7B model, paving the way for advanced applications in financial analysis, decision-making, and information generation.