**FINETUNING LLMS**

1. Llama 3.1 8b instruct model:-

* Used Unsloth for Finetuning the model.
* Training dataset having 185 samples in it.
* Used Qlora for finetuning.
* Quantization- 4bit
* Model, tokenizer- unsloth/Meta-Llama-3.1-8B-Instruct-bnb-4bit
* Used SFTTrainer from hugging face for finetuning process.
* Training Loss=0.7298920561249057
* Finally Storing the model and tokenizer into 16 bit.
* Parameter Changes made in parameters of Lora:-
* r: Increased it to 16. (higher the value of r, more the number of lora training parameters in lora tracking weights)
* lora\_alfa: increased it to 16. (higher the value of lora\_alfa, more is the impact of Lora tracking weights on the fine tune model weight.
* Parameter Changes made in parameters of SFTTrainer-
* per\_device\_train\_batch\_size: Kept same, as by increasing it Training loss increased.
* gradient\_accumulation\_step: Kept same, as by increasing it Training loss increased.
* Num\_epochs: increased to 6.

1. Phi 2 model:-

* Used Huggingface for Finetuning the model.
* Used Qlora for finetuning.
* Quantization- 4bit
* Model, tokenizer- phi-2
* Used SFTTrainer from hugging face for finetuning process.
* Training Loss=1.6
* Finally Storing the model and tokenizer into 4bit.
* Comments:

Loss 2.6 and reduced to 1.6 by :-

* per\_device\_train\_batch\_size: Increased so that more data can be accumulated by it
* gradient\_accumulation\_steps: decreased to maintain the memory constrained as batch size inceased.
* Num\_epochs: increased
* Learning rate -lower