

# Final Part Report

---

## Section 5: LLaMA Training

### 5.8 Analysis

#### LoRA settings

`lora_rank = 16 lora_alpha = 32, lora_dropout = 0.05`

#### Trainable parameter count

Total parameters: 6747639808 Trainable parameters: 8388608 Percentage of trainable parameters: 0.12%

#### GPU memory usage

40402MiB / 40960MiB

#### Gradient checkpoints

A checkpoint is placed during the forwarding process of the TransformerBlock as follows:

```
...
def forward(
    self,
    x: torch.Tensor,
    start_pos: int,
    freqs_cis: torch.Tensor,
    mask: Optional[torch.Tensor],
):
    h = x + self.attention.forward(
        self.attention_norm(x), start_pos, freqs_cis, mask
    )
    # Checkpoint
    def check_forward(x):
        return self.feed_forward.forward(self.ffn_norm(x))
    out = h + checkpoint(check_forward, h, use_reentrant=True)
    return out
...
```

#### Training loss for every epoch:

Epoch 0: Average Loss = 1.325083038618483 Epoch 1: Average Loss = 1.0072221773462813 Epoch 2: Average Loss = 0.7484890972584303 Epoch 3: Average Loss = 0.5021964082850953 Epoch 4: Average Loss = 0.3443939249542737

#### Comprehensive analysis

For each of the techniques applied (low precision, gradient accumulation, gradient checkpointing, LoRA), analyze whether or not it can contribute to memory reduction (in terms of model weight, model activation, model gradient, and optimizer state), computational cost, and potential impact on model accuracy.

| Technique              | Memory Reduction                          | Computational Cost Reduction                | Impact on Accuracy                                       |
|------------------------|---|---|--|
| Low Precision          | Reduce gradient and optimizer             | Reduce loss calculations                    | Decrease accuracy  |
| Gradient Accumulation  | Reduce gradient storage only              | Reduce gradient calculations                | No impact on accuracy                                    |
| Gradient Checkpointing | Reduce gradient storage                   | Increase optimizer and loss calculations    | No impact on accuracy                                    |
| LoRA                   | Reduce model gradient and optimizer state | Only calculate gradient for LoRA parameters | Decrease accuracy compared to training on all parameters |

Example prompts and results (fine-tuned model)

Explain why a given topic is important. Environmental protection

Environmental protection is essential for maintaining the health of our planet for future generations. Human activities such as emissions, farming, and deforestation are contributing to climate change, air and water pollution, reduction in biodiversity, and other environmental threats. By placing a priority on environmental protection, we can reduce the speed of environmental damage and conserve natural resources for our own and future generations' use.

=====

Rewrite the following sentence without changing the meaning. The sales report was not accurate.

The sales report was inaccurate.

=====

Explain why the given definition is wrong. A mole is an animal that lives underground.

This definition is wrong because moles are small, furry mammals that live above ground and are known for their distinctive long snouts and wide, black eyes. They are most commonly found in grassy areas and wooded areas.

=====

Generate a password of 8 characters.

1q2w3e4r s5t6u7v8

=====

Below are 4 random questions beyond the range of [0:200].

Name two different ways to reduce poverty in a community.

1. Improve education opportunities and access to healthcare, 2. Increase access to financing and investment opportunities, 3. Provide subsidies and incentives, 4. Foster entrepreneurship and creativity, 5. Create jobs and

=====

Convert the following number from base 10 to hexadecimal: 278

11 07 def 11 07 def def 11 07 def 11 07 def 11 07 def 11 07 def 11 07 def 11

=====

Suggest a suitable card game for a birthday party with 10 year old kids.

- Casino Royale
- War
- Crazy Eights
- Uno
- Candy Land
- Chutes and Ladders
- Tic-Tac-Toe
- Mancala
- Memory
- Happy Birthday

=====

Explain the difference between a personal computer and a server

A personal computer (PC) is a general-purpose computing device designed for individual or home use. It is typically a small, portable device with a limited amount of memory and processing power. A server, on the other hand, is a more specialized and powerful computing device designed for use in business,

=====

Example prompts and results (original model)

Explain why a given topic is important. Environmental protection

is a major issue that is faced by the world today. Many people are not aware of the negative impacts that they have on the environment. Environmental protection is important because it will help to conserve natural resources, improve the quality of the environment and also improve human health. Environmental protection is the process of protecting the environment from pollution and degradation. It involves the use of technology to reduce pollution and improve the quality of the environment. The importance of environmental protection is that it will help to conserve natural resources. Natural resources are the resources that are found in nature. These resources are used by humans for their survival. Natural resources include water, air, land, forests, minerals and animals. Natural resources are important because they are used by humans for their survival. They are also important because they are used by animals for their survival. The importance of environmental protection is that it will help to



Name two different ways to reduce poverty in a community.

Name two different ways to reduce poverty in a community.

- 1. How does the government help reduce poverty?
- 2. What are some things that people can do to help reduce poverty?
- 3. What are some ways to help reduce poverty?
- 4. What are some ways that people

=====

Convert the following number from base 10 to hexadecimal: 278

Convert the following number from base 10 to hexadecimal: 278. Convert the following number from base 10 to hexadecimal: 278.

Solution:

```
import math
def convert_to_hexadec
```

=====

Suggest a suitable card game for a birthday party with 10 year old kids.

I am looking for a suitable card game for a birthday party with 10 year old kids. I have heard of Uno and Apples to Apples but they are not the best choices. I would like to know of a suitable card game. card-games party I've

=====

Explain the difference between a personal computer and a server

. How do you explain the difference between a personal computer and a server? A personal computer is a computer that is used by a single user, whereas a server is a computer that is used to provide services to multiple users. How do you explain the difference between a personal computer and a server? A personal

=====