

# LLM Training Report - CP

Date: 2025-02-07 09:28:03 (IST)

Pretrained Model: meta-llama/Meta-Llama-3.1-8B

Dataset: Pamzyy/merged\_final1\_cleaned

## Introduction

This report provides a comprehensive overview of the training process, highlighting key performance metrics and trends observed during model training. The goal of this analysis is to evaluate the model's learning behavior, assess the effectiveness of the optimization process, and ensure the training progress aligns with the desired objectives.

The report includes visualizations of important metrics, such as loss and learning rate, recorded continuously throughout the training. These curves offer valuable insights into how the model's performance evolves over time and help identify potential issues, such as underfitting, overfitting, or convergence problems.

By reviewing this report, we aim to gain a deeper understanding of the model's training dynamics and identify opportunities for further optimization and refinement.

## Hyperparameters:

This section provides an overview of the hyperparameters used during training. Hyperparameters are crucial as they control the learning process, including aspects such as batch size, learning rate, optimizer type, and model architecture settings. Carefully selecting these parameters can significantly influence the model's performance and stability.

# LLM Training Report - CP

Parameter	Value
model	meta-llama/Meta-Llama-3.1-8B
dataset	Pamzyy/merged_final1_cleaned
pretrained_dataset	JeanKaddour/minipile
learning_rate	4.5e-05
lr_scheduler_type	cosine
per_device_train_batch_size	8
per_device_eval_batch_size	8
max_steps	10
eval_steps	10
save_steps	10
hub_model_id	Udith-Sandaruwan/Llama-test-2
finetune_tokenizer	False
output_dir	./results
resume	False

## Training Details

This section highlights key details about the training process, including configurations, hardware used, and any specific parameters that shaped the fine-tuning process. Understanding these details can help replicate or refine the training process in future experiments.

- Epochs: 1
- Training Steps: 10
- Final Loss: 1.4226
- Final Learning Rate: 0.0000
- Total Training Time (s): 858.62

## Training Metrics Summary

The training metrics summary offers a snapshot of the model's progress during fine-tuning. Metrics such as training loss, total training time, and the learning rate provide valuable insights into how well the model adapted to the dataset. This data is essential for diagnosing overfitting or underfitting.

# LLM Training Report - CP

Total Training Time (s): 858.62

Epoch/Step Markers: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

Final Loss: 1.4226

Final Learning Rate: 0.000016

## Evaluation Results

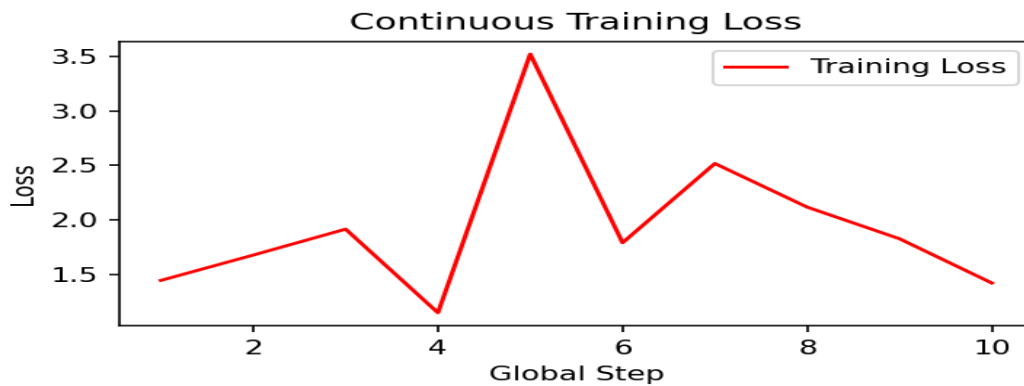
The evaluation results section contains the performance metrics of the fine-tuned model on the evaluation dataset. These metrics demonstrate how well the model generalizes to unseen data and include key measures such as accuracy, F1 score, and perplexity. A detailed evaluation ensures the model meets desired performance benchmarks.

- eval\_loss: 1.4747
- eval\_runtime: 96.5138
- eval\_samples\_per\_second: 8.5480
- eval\_steps\_per\_second: 1.0780
- epoch: 0.0001
  
- meteor\_scores: 0.007878151260504201
- rouge\_scores: 0.0
- bleu\_scores: 0.0
- perplexity: 80118250000.0

## Loss Curve

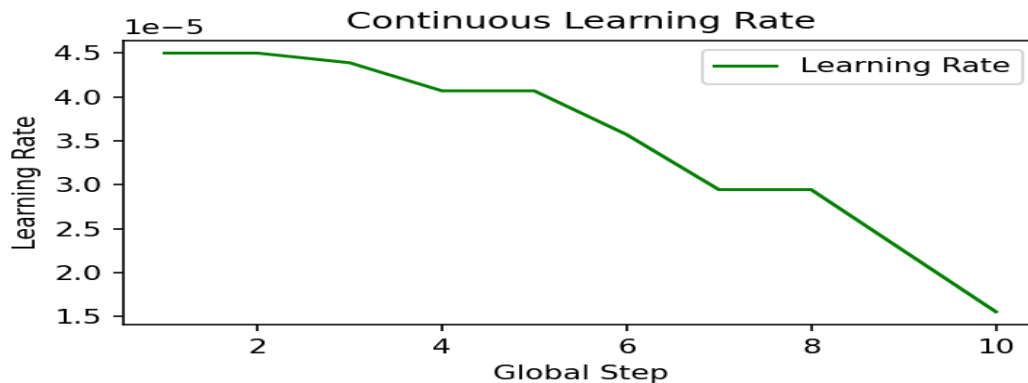
The training loss curve illustrates how the model's loss decreased over time during fine-tuning. A smoothly decreasing curve indicates that the model was learning effectively, while any irregular patterns may suggest optimization challenges.

# LLM Training Report - CP



## Learning Rate Curve

The learning rate curve shows how the learning rate was adjusted during the training process. Gradual or scheduled changes in learning rate help the model converge to an optimal solution without overshooting or stagnating.



## Inference Examples

Below are several examples of model inference results on sample prompts. These examples illustrate the model's performance at different stages (e.g., before and after tokenizer fine-tuning, and after continued pre-training).



# LLM Training Report - CP

Fine-Tuned Model Repository:

# LLM Training Report - CP

## Fine-Tuned Model Repository:

This section provides the location of the fine-tuned model repository. The repository serves as a central place to access the trained model, documentation, and configuration files. It ensures the model is reusable for further tasks and evaluations.

Udith-Sandaruwan/Llama-test-2