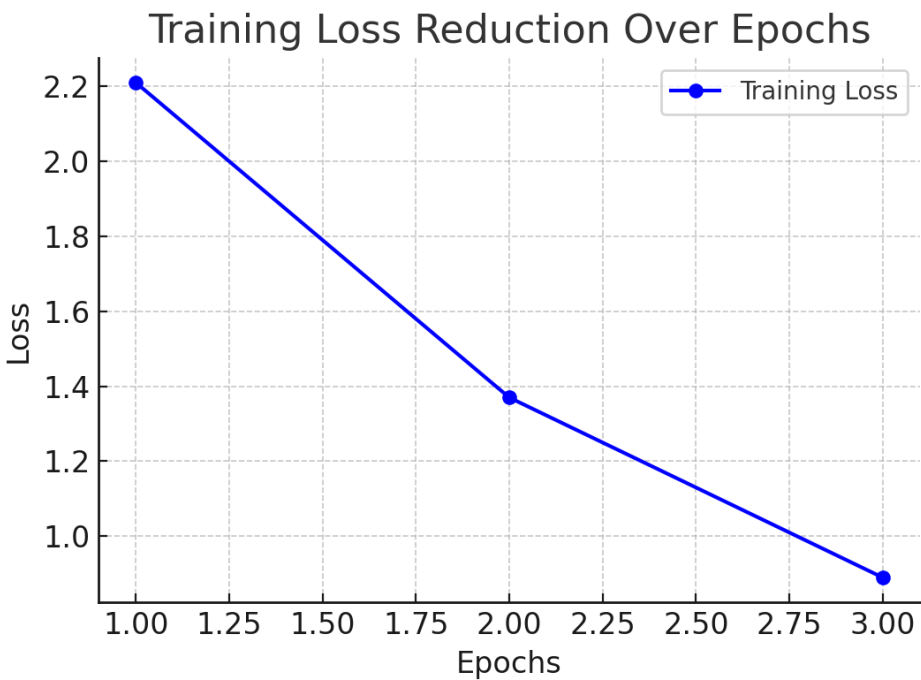


Model Insights - Gemma-2B Fine-Tuned on Alpaca-GPT4

This document presents insights from fine-tuning the Gemma-2B model on the Vicgalle/Alpaca-GPT4 dataset. The fine-tuning focused on enhancing instruction-following capabilities.

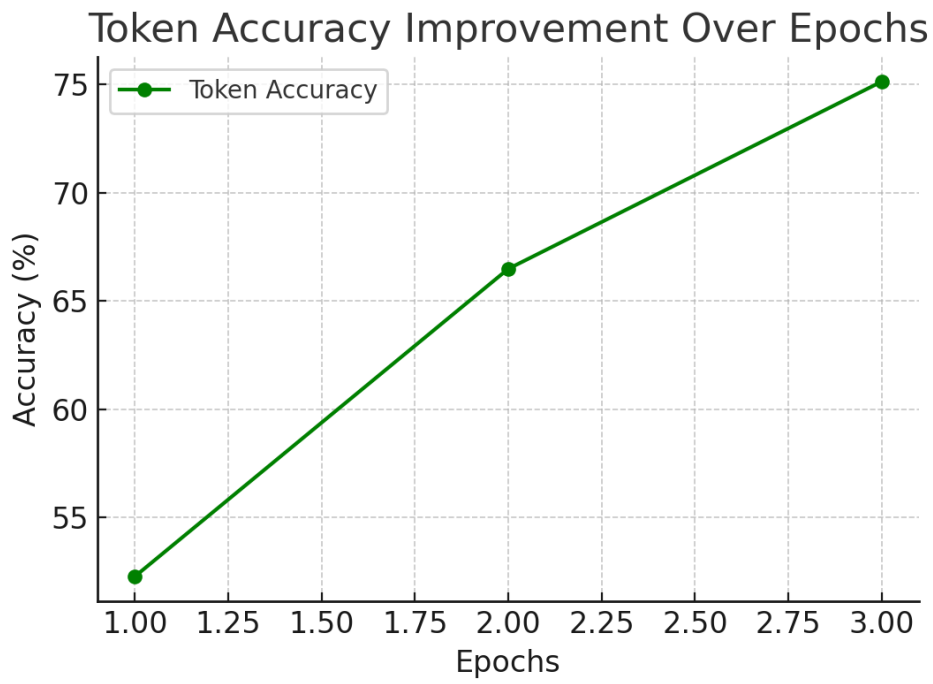
Training Metrics

- Initial Loss: 2.21
- Final Loss: 0.89
- Token Accuracy (Final): 75.13%
- Training Stability: No instability detected
- Dataset: Vicgalle/Alpaca-GPT4



Token Accuracy Improvement

The model showed steady improvement in token accuracy, demonstrating its ability to follow instructions more effectively after fine-tuning.



Observations & Next Steps

The model successfully learned structured instruction-following. However, some anomalies were noted with complex, multi-step instructions. Further tuning with datasets like MMLU or TruthfulQA may improve robustness.