Project Report - Latent Diffusion for Language

Author Name

October 22, 2024

1 Brief Overview

This project aims to implement the techniques proposed in the paper titled **"Latent Diffusion for Language Generation"**. The paper introduces a novel approach to language generation by leveraging latent diffusion models in the continuous latent space of a pretrained encoder-decoder network, such as BART or T5. The project demonstrates significant improvements in language generation tasks, especially in reducing the sampling steps and maintaining high output quality. The project implementation is available on GitHub: `https://github.com/justinlovelace/latent-diffusion-for-language`.

2 Training Details

The implementation of the paper was already available but there were some challenges in running the codebase. Some challenges also arose due to hardware limitations.
Due to the limited computational resources on my personal system, I used Kaggle's P100 GPU, which provides 16GB of VRAM. This forced me to reduce the batch size to half of the original paper's specification. Additionally, training time on Kaggle's platform was limited, hence the results mentioned in this paper are by running the model for only 1/5th of the total training steps described in the paper.
The latent model was trained for approximately 7 hours on Kaggle's P100 GPU, while the diffusion model was trained for around 9 hours. Thus the results are lagging compared to those mentioned in the paper but the results shown demonstrate that model is able to learn some features and gives hope that running for the full time will be able to re-create nearby results as those mentioned in the paper.

3 Dataset Description

The model was evaluated on the ROCStories dataset, which consists of five-sentence stories that require the model to generate coherent and logically consistent narratives.
The dataset is publicly available and the link is also provided in the submission form.

4 Results

Given the limited training time and reduced batch size, the results obtained reflect the model's early-stage performance. We can compare the performance based on various metrics used in the paper which includes perplexity, unique word count, n-gram repetition, Diversity, Memorization and MAUVE score.

Metric	**Achieved Value**	**Reference Value**
Unique Word Count	1173	1236
2-gram Repetition	0.263	0.244
3-gram Repetition	0.048	0.042
4-gram Repetition	0.009	0.008
Diversity	0.695	0.718
Perplexity	39.22	20.37
Memorization	0.328	0.370
MAUVE Score	0.768	0.901

Table 1: Comparison of Achieved Metrics with Reference Values

5 References

- Paper Link: https://arxiv.org/pdf/2212.09462
- GitHub link for implementation:
https://github.com/justinlovelace/latent-diffusion-for-language