Homework 3

Haoli Yin, CS 8395

December 15, 2023

Problem 1:

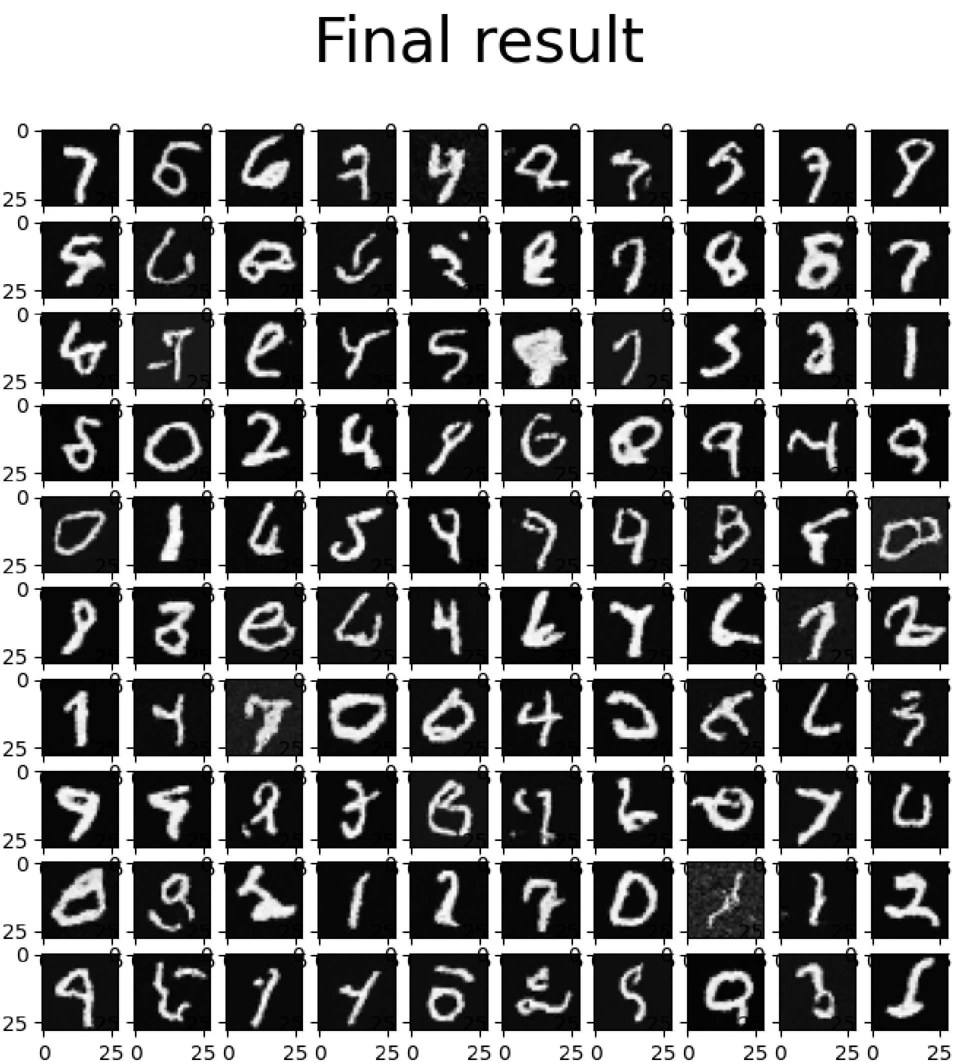

Figure 1: MNIST results generated by DDPM

(a) Compare and Contrast GANs and Diffusion models

Generative Adversarial Networks (GANs), including their variant Wasserstein GAN with Gradient Penalty (WGAN-GP), and Denoising Diffusion Probabilistic Models (DDPM) represent distinct methodologies in the field of generative models. Vanilla GANs operate on the principle of adversarial training, involving a dual network system of a generator and a discriminator. The generator creates synthetic data, while the discriminator evaluates its authenticity, leading to a competitive training dynamic. This approach, however, can suffer from challenges like mode collapse, where the generator may fail to produce diverse outputs. WGAN-GP addresses these limitations by incorporating the Wasserstein loss and a gradient penalty. These modifications provide a more stable training environment and help in producing more diverse and realistic outputs by mitigating common issues like mode collapse and training instability in traditional GANs.

On the other hand, DDPM adopts a fundamentally different approach. It involves a diffusion process where noise is incrementally added to and then removed from data, effectively learning the data distribution through a denoising process. This method tends to be computationally more intensive due to its iterative nature but is known for generating high-quality and diverse samples. Unlike GANs, DDPMs are less prone to issues like mode collapse and offer a more controlled generation process, allowing for nuanced manipulation during the generation. In essence, while GANs and WGAN-GP excel in efficiency and quality of generation, they can be susceptible to certain training instabilities. DDPMs, conversely, offer robustness in training and diversity in output, albeit with higher computational requirements, highlighting the trade-offs between these generative approaches.

Paper 1 Review

Title: $\mathbb{D}^2$ Pruning: Message Passing for Balancing Diversity & Difficulty in Data Pruning

(a) Paper Summary

The paper presents a novel coreset selection algorithm, PRUNING (also referred to as D2 pruning for Diversity-Difficulty pruning), aimed at improving data selection for deep learning model training. It uses a graph-based approach combined with a message-passing algorithm to balance data diversity and sample difficulty. This method is particularly effective in image classification and NLP tasks, especially at low-to-medium data pruning rates.

(b) Strengths and Weaknesses

Strengths:

1. PRUNING offers a novel approach that effectively balances data diversity and difficulty, applicable to both supervised and self-supervised learning.

2. The paper provides clear and detailed experiments, including ablation studies to support its claims and the introduction of new hyper-parameters.

Weaknesses:

1. PRUNING introduces additional hyper-parameters, potentially increasing the complexity and cost of coreset selection.

2. The performance improvements offered by PRUNING over state-of-the-art methods are relatively incremental

(c) Clarity

The paper is well-written, presenting its methodology and experiments with clarity. However, there are inconsistencies in section numbering and references to the appendix, which could be fixed for better readability.

(d) Quality, Novelty, and Reproducibility

PRUNING is a quality contribution to the field, offering a novel approach to coreset selection. Its applicability to NLP datasets fills a gap in existing literature. However, the novelty is somewhat dampened by the marginal performance improvements over existing methods. The paper's detailed presentation suggests that the results could be reproducible, although the introduction of new hyper-parameters might complicate this process.

(e) Recommendation

Score of 6, marginally accept

Paper 2 Review

Title: DeCUR: decoupling common & unique representations for multimodal self-supervision

(a) Paper Summary

The paper introduces a new multimodal self-supervised learning method called DeCUR, which addresses a key limitation of existing methods: the inability to adequately train modality-unique representations. This method focuses on decoupling common and unique representations by differentiating between inter-modal and intra-modal embeddings. This allows for the integration of complementary information from various modalities. The effectiveness of DeCUR is demonstrated through experiments in three multimodal scenarios and two downstream tasks, complemented by an interpretability analysis.

(b) Strengths and Weaknesses

Strengths:

1. Insightful Ablation Study and Analysis: The ablation study and the analytical section offer valuable insights into the method's functionality.
2. Interpretability Analysis: The paper includes an in-depth interpretability analysis, providing insights into the workings and effectiveness of the proposed method.

Weaknesses:

1. Missing Comparative Results: The paper does not include results from training with 100% labels using methods like Barlow Twins and VICReg.
2. Limited Modality Application: The research primarily focuses on image modalities, raising questions about its applicability to other modalities like audio and text.

(c) Clarity

The paper is well-structured and presents its concepts clearly, making it accessible to readers. The method and experiments are described in a way that is easy to comprehend.

(d) Quality, Novelty, and Reproducibility

The paper is of good quality, with comprehensive experiments supporting its claims. While the approach shows potential, its novelty is somewhat limited as it primarily extends existing concepts to new settings. The inclusion of supplementary materials and a detailed method description suggests that the results are reproducible.

(e) Recommendation

Rating: 7