Unsupervised Deep Learning - DSB

Aymeric Dieuleveut

December 6, 2024

Project

Rules:

1. Groups of 3 max.
2. Submit, by email, a single notebook, before next friday. Use [DL-Unsup] as email object, and name the file Name1Name2Name3, before next friday 6pm.
3. All code should be prerunned and runnable.
4. **Describe your approach - what you could have tried / have tried. What you observe, etc.**
5. **I encourage you to try to think on your own. If ideas or parts of the code are obtained from internet or else, mention it (you won't get penalized IF and ONLY IF you do). Reusing blocks from our labs is ok, as well as existing block of code, IF you explain how you obtained them.**
6. The goal is not to maximize the performance but to describe your chain of thought. Originality will be appreciated

1. Anomaly detection on corrupted EMNIST

1. **Altered EMNIST**
   - I altered the **EMNIST** dataset (in one or several ways).
   - I have posted the dataset on slack.
2. **Unsupervised anomaly detection.** Propose techniques to obtain anomaly detection.
   - Try to explore at least 2 techniques
   - Identify how the dataset has been corrupted
3. **Recovering the dataset.** For the corruption that you have identified and can be addressed, propose ways to "rebuild" the dataset.
4. **Evaluation** Propose metrics to quantify the improvement of your solution.

2. Explore Mode collapse for training a GAN on MNIST

For this task, you **are not allowed to use the labels on MNIST**, even for point 3.

- Train a GAN on Mnist, use multiple seeds or architectures
- Explore mode collapse. Try to obtain cases in which you have mode collapse.
- Propose and implement solutions to limit mode collapse:
  - Other methods
  - Adding conditionality
  - etc.
- Propose metrics to evaluate mode collapse.

- Extend the results to other datasets (CIFAR, EMNIST).
- Compare generation to other methods (Diffusions, VAE, etc.)