

Module de Deep avancé, 2025-2026

Modalités pour les TPs (Modèles génératifs)

Julien Rabin

10 décembre 2025

Modalités générales

Le travail de TP peut être évalué selon deux modalités au choix :

- un **compte rendu approfondi** portant sur l'un des modèles génératifs étudiés en TP ;
- un **projet exploratoire** sur un sujet plus avancé.

1 TP approfondi

Le compte rendu peut porter sur un ou plusieurs modèles parmi ceux étudiés en TP : AE+GMM, VAE, GAN, DDPM, GPT. L'objectif est d'entraîner et d'évaluer un modèle de manière approfondie.

Jeux de données réelles

Il est d'abord fortement recommandé de travailler sur un jeu de données plus complexes (que les nuages de points 2D vus en TP), par exemple issus de `torchvision` : <https://docs.pytorch.org/vision/stable/datasets.html>

- **Jeux simples** : MNIST, Fashion-MNIST, CIFAR-10/100 ...
 - **Jeux plus complexes de visage** : CelebA, CelebA-HQ, FFHQ, Animal Faces-HQ (AFHQ)...
- Une réduction de la résolution (par exemple à 64×64 px) est recommandé pour limiter le coût de calcul.

Pour ces jeux de données, prendre soin de préparer des sous-ensembles pour l'entraînement et les tests/validations.

Pistes d'approfondissement

Voici quelques exemples d'aspect qui peuvent être approfondis :

- Pour un même jeu de données : comparer les performances
 - de plusieurs modèles à nombre de paramètres égaux
 - pour un même modèle en fonction du nombre de paramètres
 - pour un même modèle, en fonction de la dimension de l'espace latent (si approprié)
 - etc.
- Évaluation de la qualité des échantillons aléatoires générés :

- détection de l'originalité par comparaison au plus proche voisin dans le jeu d'entraînement
- pour les images spécifiquement : utilisation de critère de qualité perceptuel comme FID, IS, LPIPS ;
- pour le texte : BLEU, perplexité, ROUGE...
- Exploration d'architectures plus adaptées que les MLP : ResNet, normalisation (Batch-Norm, LayerNorm...), skip-connections, etc. En particulier pour les images (convolutions ...) ou le texte (tokenisation adaptée au texte).
- Exploration de techniques d'entraînement avancées : utilisation de LPIPS comme critère d'optimisation (AE, GAN, DDPM), l'effet de la régularization (WGAN-GP), etc.

2 Projet exploratoire

Il s'agit d'étendre les méthodes vues en cours ou d'étudier des modèles non abordés en TP. Quelques exemples (non exhaustifs) :

- modèles génératifs : le rôle de la quantification dans l'auto-encodeur VQGAN, apprentissage d'un transformer latent (VQ-GAN), l'apprentissage de normalizing flows pour des images, modèles génératifs de type flow-matching ;
- modèles génératifs conditionnels (par classe, par texte, etc.) ;
- stylisation ou transformation d'images ;
- génération audio ou musicale (voir par exemple https://geoffroypeeters.github.io/deeplearning-101-audiomir_book/task_musicgeneration.html) ;
- génération multimodale : text-to-image, image-to-text.
- ...

Les étudiant·e·s peuvent proposer librement un sujet à leur encadrant de TP.

3 Format du compte rendu

Le rendu prendra la forme d'une archive **.zip** contenant :

1. un **README** (texte ou Markdown) décrivant :
 - les dépendances utilisées,
 - la création d'un environnement virtuel,
 - les instructions d'exécution ;
2. un **notebook Jupyter** contenant :
 - le code,
 - les visualisations,
 - les résultats expérimentaux ;
 - (eventuellement) les paramètres du/des réseaux
3. un **rapport** (PDF ou Markdown intégré dans le notebook).

Le travail en binôme est autorisé.

4 Date de rendu

Vendredi 10 janvier à minuit.

5 Évaluation

La performance brute du modèle n'est pas le critère principal (mais sera appréciée)! L'évaluation portera principalement sur :

- la qualité de la rédaction et des explications ;
- la clarté de la démarche expérimentale ;
- la documentation du code personnel et la citation des sources externes.