

Rapport Pratique IFT6135-A2022

Assignment 2

Hamed Nazim MAMACHE

Décembre 2022

1 Question 1 : VAE

3. Après avoir entraîné le modèle VAE, voici les résultats obtenus de nos expériences :
- (a) On peut observer ici quelques reconstructions de l'ensemble de test et des échantillons de modèle VAE en fin d'entraînement.



Figure 1: Quelques reconstructions après l'entraînement du modèle VAE (gauche) et de l'ensemble de test (droite)

- (b) Sur les échantillons on arrive difficilement parfois à observer les modèles de chiffres : les images sont floues.
Le fait que le modèle génère des images floues pourrait être dû à son hypothèse de gaussianité et de la loss L2, qui est basé sur l'hypothèse que les données suivent une seule distribution gaussienne.
En effet, la loss L2 entraîne à l'obtention d'images floues lorsque les données sont tirées de distributions multimodales.
- (c) Voici les images issues de l'interpolation :

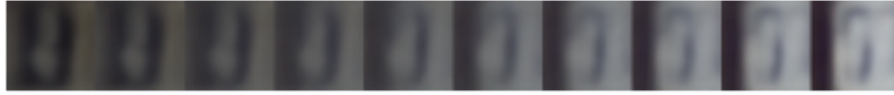


Figure 2: Les images de l'interpolation pour VAE

L'interpolation entre deux points ne semble pas très fluide. En effet, les images semblent évoluer sans gagner en douceur : elles restent toujours floues et ne gagnent ni en netteté ni en finesse.

2 Question 2 : GAN

2. Après avoir entraîné le modèle GAN, voici les résultats obtenus de nos expériences :

(a) Quelques échantillons du modèle GAN à la fin de l'entraînement :



Figure 3: Quelques reconstructions en fin d'entraînement pour GAN

- (b) On arrive cette fois-ci à bien observer les chiffres dans les échantillons : les images sont beaucoup moins floues que ceux du VAE.
On remarque de plus une grande distinction entre le "vrai" et le "faux".
- (c) Voici les images pour le résultat de l'interpolation entre deux points de la distribution du bruit :



Figure 4: Les images de l'interpolation pour GAN

L'interpolation paraît plus fluide que celle observée pour le modèle VAE. Les images sont de plus en plus précises : on arrive de mieux en mieux à distinguer ce qu'il y est représenté. Les bords sont de plus en plus nets, et les détails de plus en plus fins.

3. Dans la boucle d'entraînement du modèle GAN, utiliser `.detach()` dans l'entraînement du discriminateur pour les fausses images générées est indispensable. En effet, on cherche à ce que la perte du discriminateur calcule uniquement les gradients pour le discriminateur et non pour le générateur. Donc on ajoute `.detach()` pour rompre la "gradient connection" avec les données "fake", seulement lorsqu'on calcule la loss du discriminateur.

4. Le modèle GAN, tel qu'il est actuellement, peut être utilisé pour reconstruire des images d'entrée et l'apprentissage de la représentation.

Cependant, il ne peut pas être utilisé pour calculer (exactement ou approximativement) la log-likelihood.

En effet, il ne considère pas de modèle de probabilité explicite pour les données. Le modèle GAN cherche à obtenir un modèle génératif en minimisant la distance entre les distributions observées et générées sans tenir compte d'un modèle de probabilité explicite pour les données. Il revient donc impossible de calculer la log-likelihood.

3 Question 3 : Modèle de diffusion (DDPM)

2. Après avoir entraîné le modèle de Diffusion, voici les résultats obtenus de nos expériences :

- (a) Les échantillons du modèle Diffusion à la fin de l'entraînement :



Figure 5: Quelques reconstructions en fin d'entraînement pour DDPM

- (b) Les échantillons semblent réalistes et de bonne qualité : ils sont moins flous que ceux du VAE, et on n'observe pas la forte distinction entre le "vrai" et le "faux" propre au modèle GAN.

4 Question 4 : Modèle génératif ; avantages et inconvénients

1. Comparatif entre les trois modèles génératifs populaires : VAE, GAN et DDPM.
On évoque ici les principaux points forts et points faibles de ces différents modèles :

(a) **modèle VAE**

Avantages:

- Le modèle VAE semble plus facile à entraîner et à "optimiser" : il est relativement facile à mettre en œuvre et est robuste aux choix d'hyperparamètres.
- On peut suivre l'évolution de la log-likelihood, grâce à l'approche probabiliste du modèle.
- Il existe une manière claire et reconnue d'évaluer la qualité du modèle.
- Le temps de calcul n'est pas important.

Inconvénients:

- les images générées par le modèle VAE sont floues et peu réalistes. Cela est dû à la perte L2, qui est basée sur l'hypothèse que les données suivent une seule distribution gaussienne. Lorsque les échantillons de l'ensemble de données ont une distribution multimodale (ce qui est généralement le cas), VAE ne peut pas générer d'images avec des bords nets et des détails fins.

(b) **modèle GAN**

Avantages:

- Le modèle GAN est très bon pour générer des caractéristiques visuelles. Les échantillons générés présentent des images réalistes et nettes.
- Le temps de calcul n'est pas important.

Inconvénients:

- Il n'y a pas de manière généralement convenue d'évaluer un modèle de générateur GAN donné, sauf en visualisant des échantillons. En effet, il n'y a pas de fonction objective ou de mesure objective d'évaluation pour un meilleur entraînement du modèle et la génération de sorties complexes.

Par exemple, cela peut poser problème dans les situations suivantes :

- Choix d'un modèle de générateur GAN final lors d'un entraînement.
 - Comparaison des architectures de modèles GAN.
 - Comparaison des configurations de modèles GAN.
- La stabilité des GAN dépend fortement de l'architecture du réseau : dans certains cas, l'entraînement devient instable et il devient plus difficile de former et de générer des résultats. Une mise en œuvre correcte du modèle GAN nécessite une sélection minutieuse des hyper-paramètres et une initialisation appropriée des paramètres. Autrement, une défaillance entraîne une non-convergence ou des mauvais résultats.

(c) **modèle DDPM**

Avantages:

- Le modèle de diffusion est traitable analytiquement. En effet, il peut être évalué analytiquement et peut ajuster les données.
- Le modèle DDPM est également flexible, c'est-à-dire qu'il sait s'adapter à des structures arbitraires dans les données.
- Les résultats obtenus sont très bons : pas d'effet de flou, et les images semblent réalistes.

Inconvénients:

- Le plus grand inconvénient du modèle DDPM est incontestablement le temps. En effet, le modèle de diffusion repose sur une longue chaîne de Markov d'étapes de diffusion pour générer des échantillons, ce qui peut être assez coûteux en termes de temps et de calcul.