

Générer des images réalistes

grâce aux « Variational Autoencoders »

Contacts

- Natacha NJONGWA YEPNGA:
<https://www.linkedin.com/in/natacha-njongwa-yepnga/?originalSubdomain=fr>
- LeCoinStat:
<https://www.youtube.com/c/LeCoinStat>
- Newsletter: <https://natacha-njongwa-yepnga.ck.page/inscriptionnewsletter>



PROGRAMME

1

Qu'est-ce qu'un
VAE (Variational
Autoencoder) ?

2

Architecture
d'un VAE

3

Entraînement
d'un VAE

4

Échantillonnage
à partir de
l'espace latent

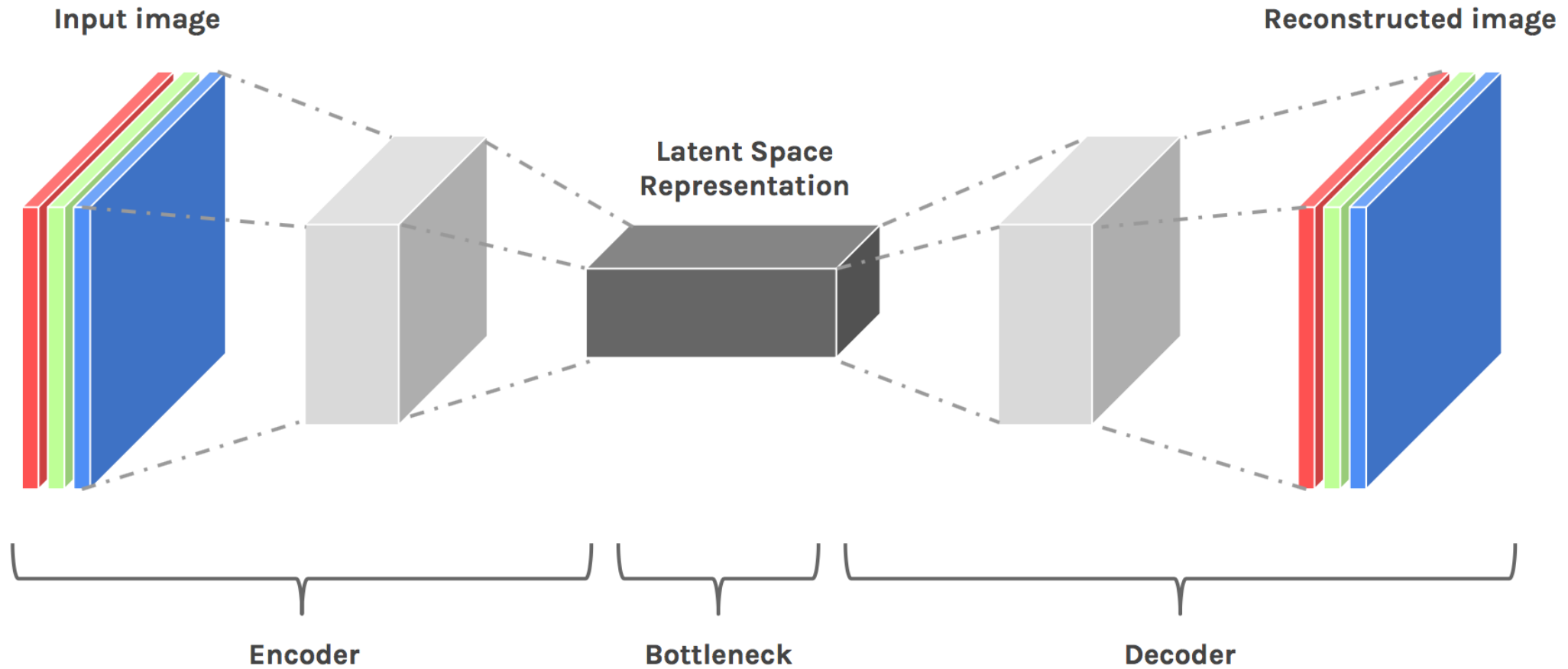
5

Ressources

A thick yellow horizontal bar spans the width of the slide, with a vertical yellow bar extending downwards from its right end.

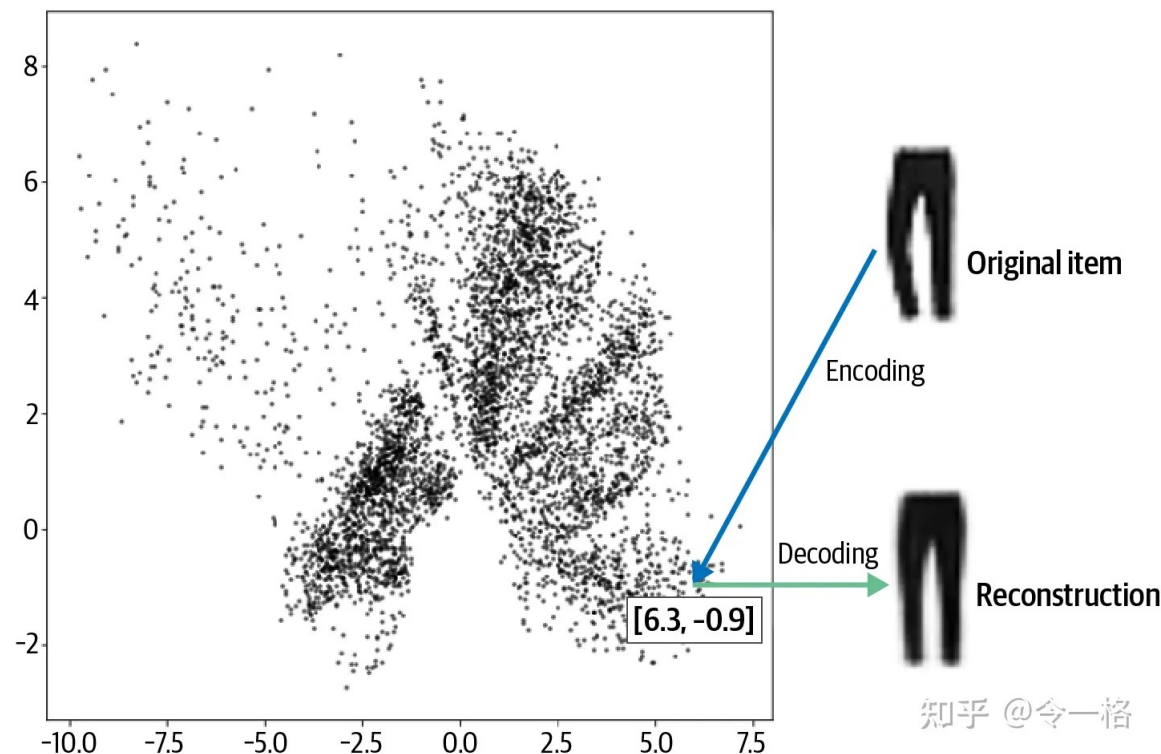
Qu'est-ce qu'un VAE (Variational Autoencoder) ?

Concept d'auto-encodeur

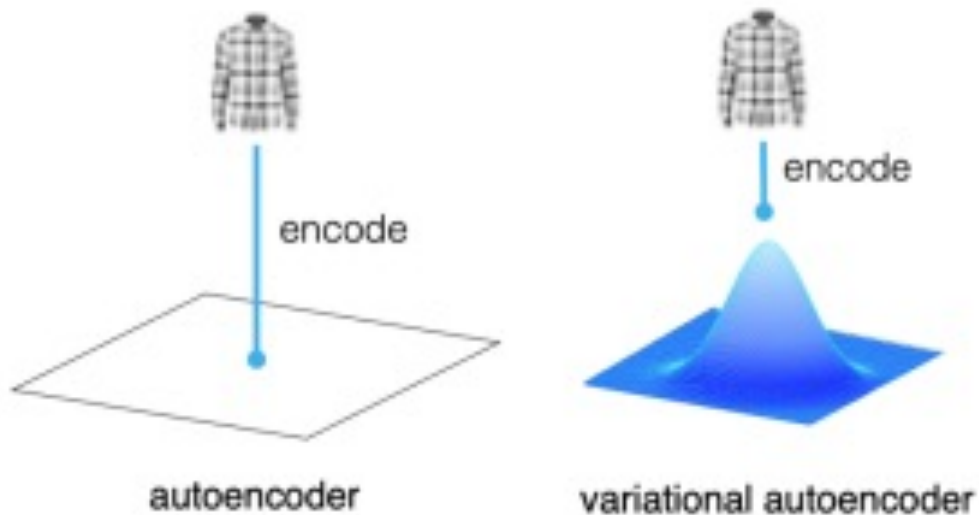


Problème des auto-encodeurs

L'auto-encodeur est un réseau de neurones qui est entraîné pour encoder et décoder les éléments de telle sorte que **le résultat final soit le plus proche possible de l'élément initial**.

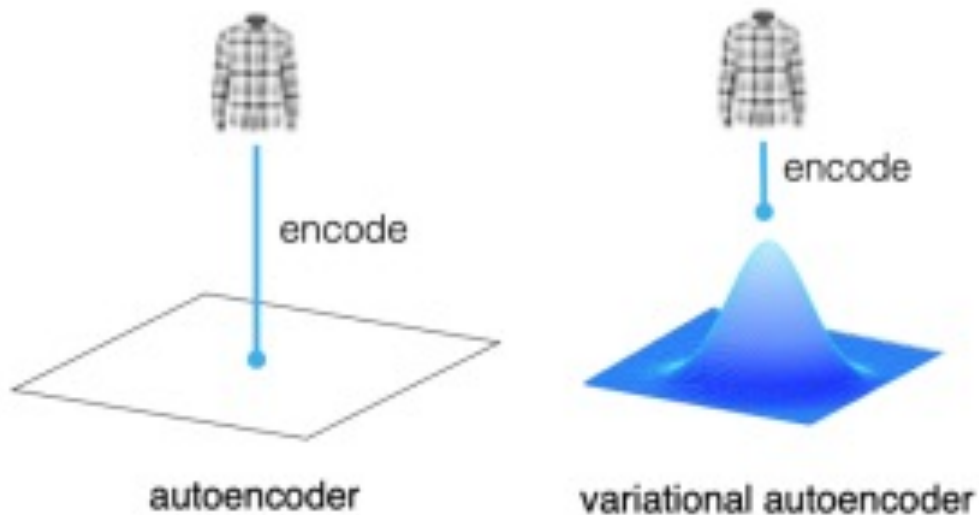


VAE: Variational Autoencoders



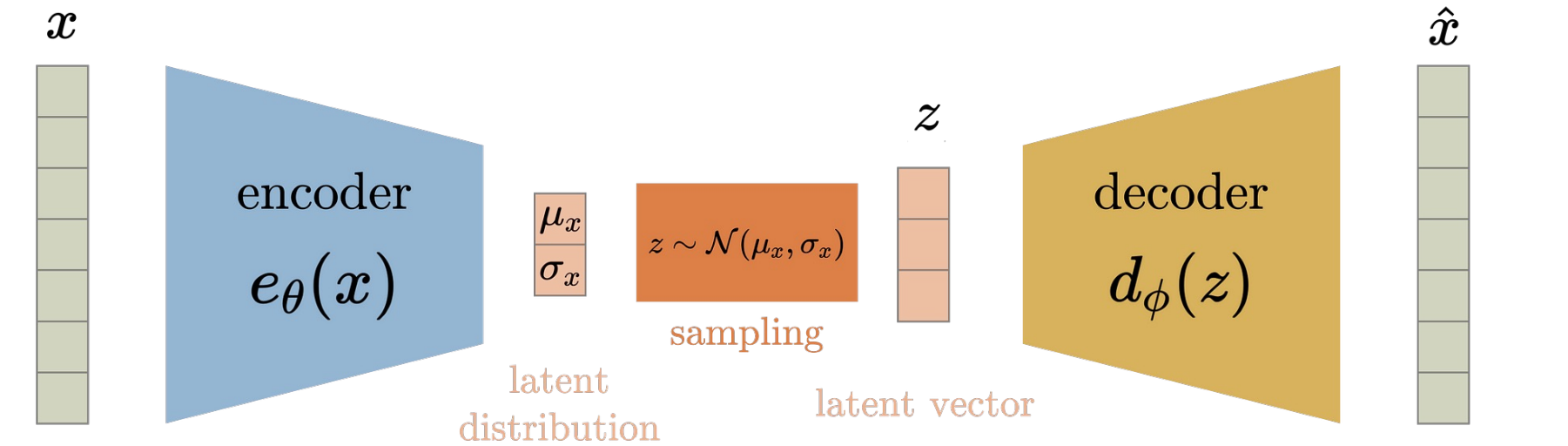
- Dans un encodeur, toutes les images sont associées à un point dans un espace latent.
- Dans un VAE, toutes les images sont associées à une loi normale multivariée autour du point.
- Dans un VAE, on attribue à chaque point un vecteur de moyenne et de variance.
- Hypothèses : il n'y a pas de corrélation entre les dimensions dans l'espace latent.

VAE: Variational Autoencoders



- Dans un encodeur, toutes les images sont associées à un point dans un espace latent.
- Dans un VAE, toutes les images sont associées à une loi normale multivariée autour du point.
- Dans un VAE, on attribue à chaque point un vecteur de moyenne et de variance.
- Hypothèses : il n'y a pas de corrélation entre les dimensions dans l'espace latent.

VAE: Architecture



reconstruction loss $= \|x - \hat{x}\|_2 = \|x - d_{\phi}(z)\|_2 = \|x - d_{\phi}(\mu_x + \sigma_x \epsilon)\|_2$

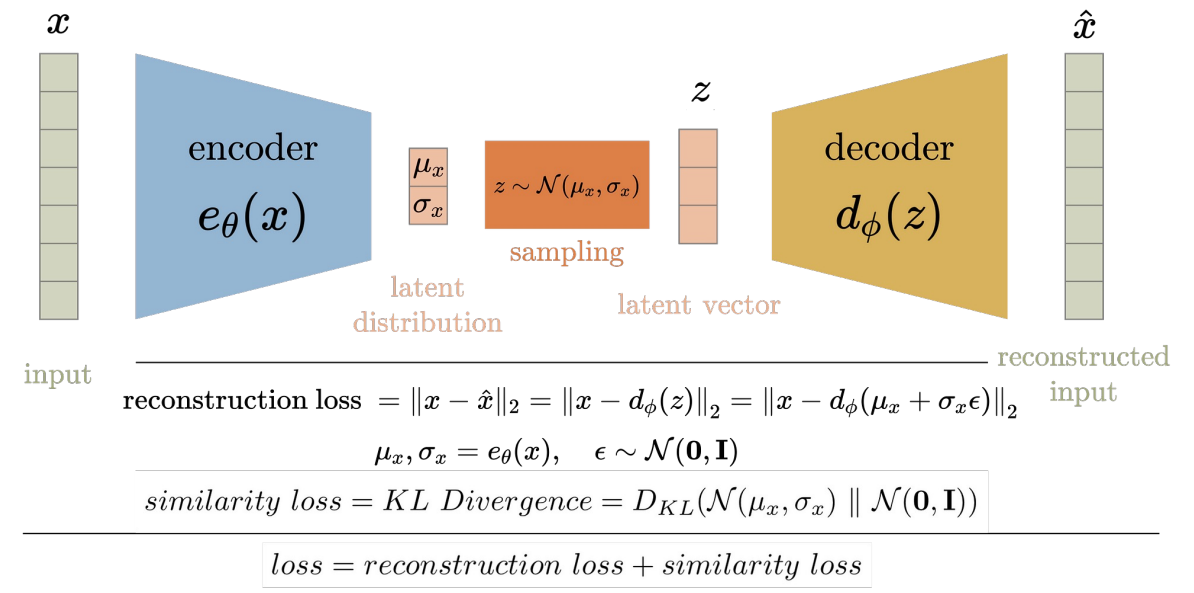
$$\mu_x, \sigma_x = e_{\theta}(x), \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

$$\text{similarity loss} = KL \text{ Divergence} = D_{KL}(\mathcal{N}(\mu_x, \sigma_x) \parallel \mathcal{N}(\mathbf{0}, \mathbf{I}))$$

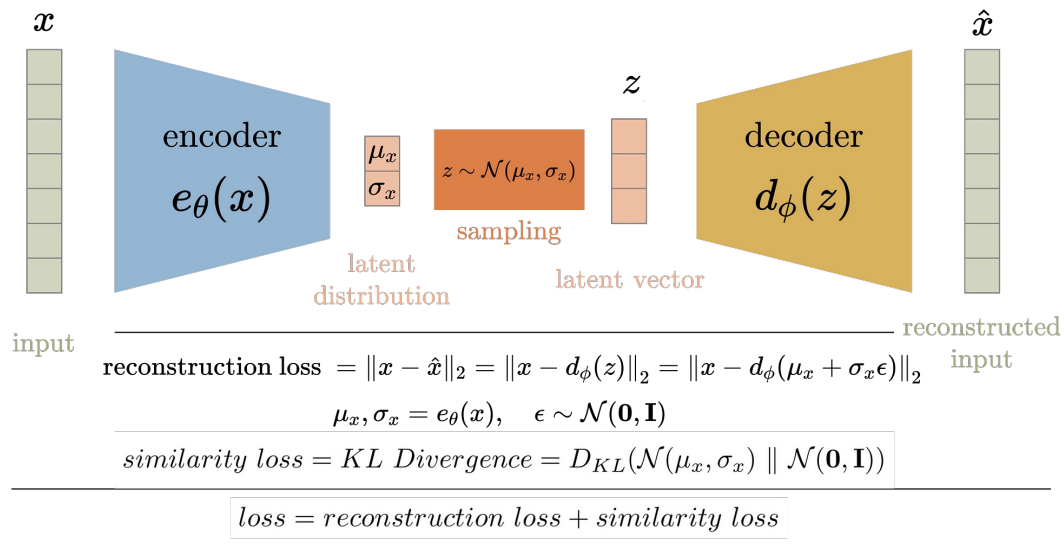
$$\text{loss} = \text{reconstruction loss} + \text{similarity loss}$$

Architecture VAE

- **L'encodeur** : Transforme les données d'entrée en une représentation latente.
- Comprend des couches de neurones qui réduisent progressivement la dimensionnalité des données d'entrée.
- Fournit en sortie un vecteur de moyennes et un vecteur d'écart-types (pour la distribution normale dans l'espace latent).



Architecture VAE



- **Le décodeur** : Génère des données reconstruites à partir de la représentation latente.
- Comprend également des couches de neurones, mais elles augmentent progressivement la dimensionnalité pour obtenir la sortie.
- La sortie est une estimation des données d'entrée originales.

Processus d'entraînement

- 1. Alimentation des données d'entraînement :** Les données d'entraînement (par exemple, des images) sont fournies au modèle.
- 2. Encodage :** L'encodeur réduit la dimensionnalité des données et génère deux vecteurs, la moyenne et l'écart-type, de la distribution normale dans l'espace latent.
- 3. Échantillonnage stochastique :** À partir de la distribution générée par l'encodeur, un échantillon stochastique est prélevé pour obtenir un point dans l'espace latent.



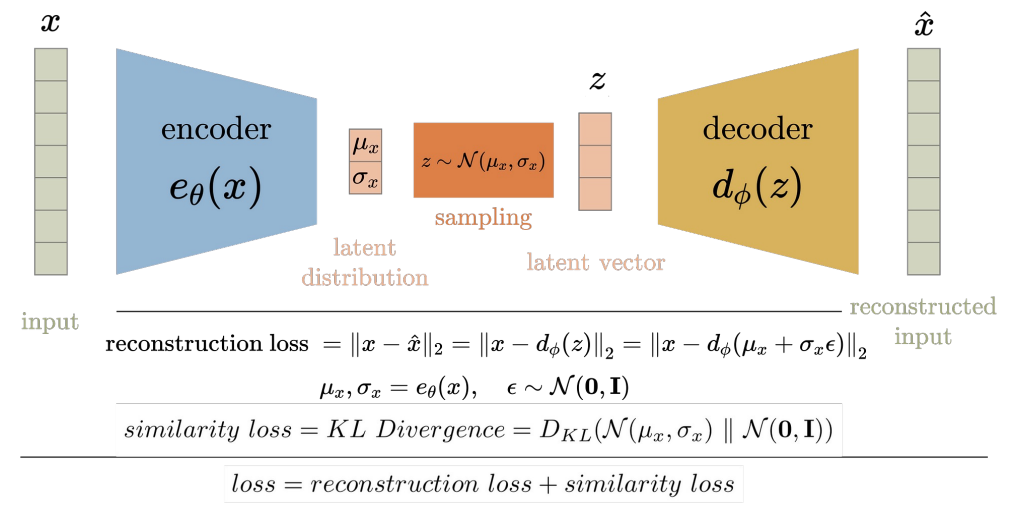
Processus d'entraînement

4-Décodeur : Le point échantillonné est passé au décodeur, qui génère une sortie qui est une reconstruction des données d'entrée.

5-Fonction de perte : Une fonction de perte, souvent constituée d'une composante de reconstruction et d'une composante de régularisation KL (Kullback-Leibler), est calculée pour mesurer à quel point la reconstruction est fidèle aux données d'entrée tout en régulant la distribution latente.

6-Rétropropagation : Les gradients sont propagés à travers le modèle, permettant aux poids du réseau d'être ajustés.

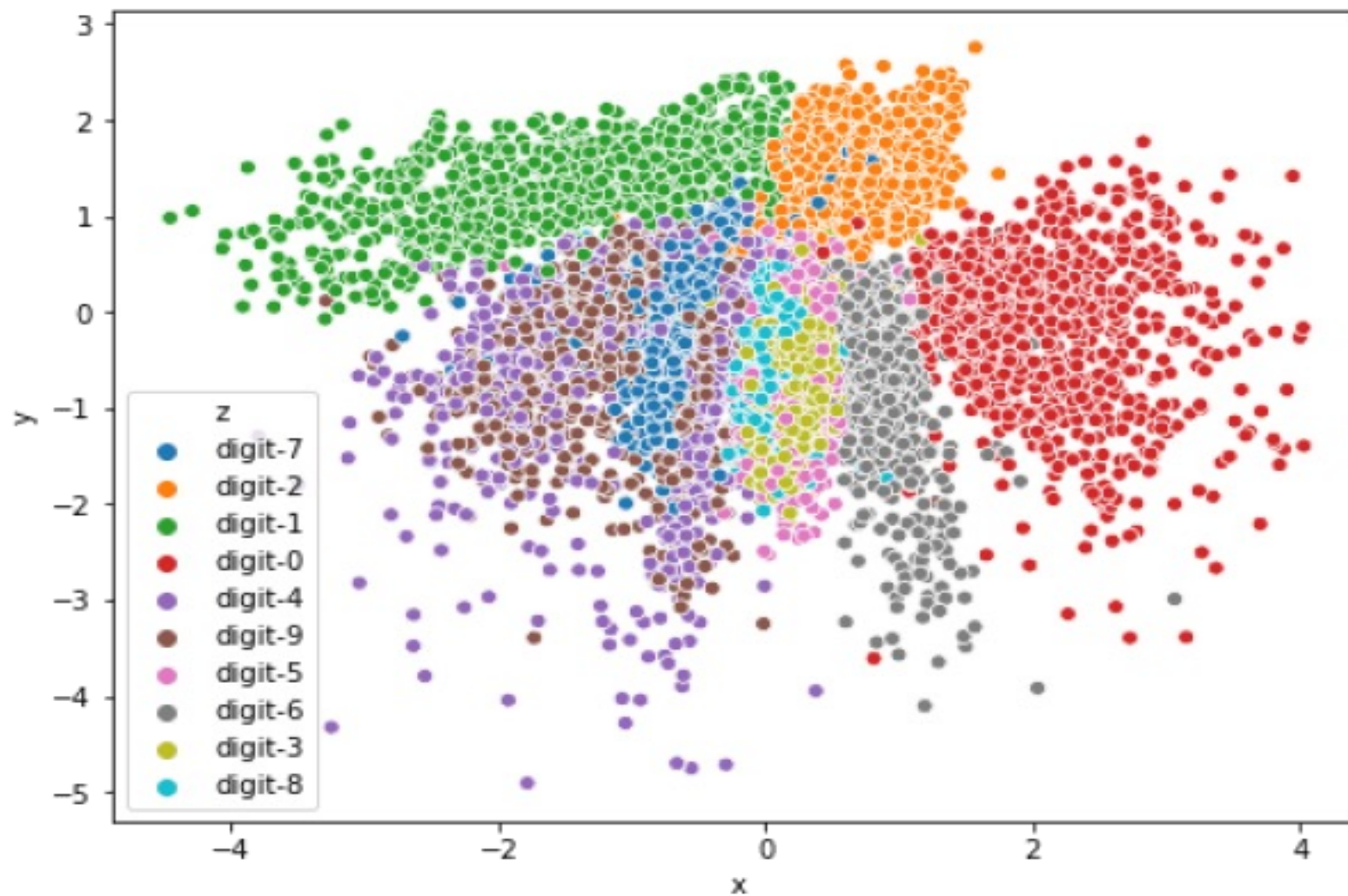
7-Optimisation : Un optimiseur (par exemple, la descente de gradient stochastique) est utilisé pour ajuster les poids du réseau afin de minimiser la fonction de perte.



Génération de nouvelles images

- L'échantillonnage à partir de l'espace latent est une caractéristique essentielle des VAE, permettant la génération de nouvelles données.
- Processus d'échantillonnage :
 - **Distribution latente** : Lors de l'entraînement, l'encodeur génère une distribution normale multivariée (souvent une gaussienne) dans l'espace latent.
 - **Échantillonnage stochastique** : Un point est échantillonné de cette distribution en utilisant la méthode de la reparamétrisation. Cela introduit de la stochasticité dans le processus.
 - **Génération** : Le point échantillonné est passé au décodeur, qui génère une nouvelle image ou exemple de données.
- Avantages de l'échantillonnage :
 - Il permet au VAE de créer une variété d'échantillons réalistes à partir de la même représentation latente.
 - En échantillonnant à différents endroits de la distribution latente, le modèle peut générer des données diverses et créatives.

Utilisation



Quiz

- **Question 1** : Quelle est la principale différence entre un VAE et un autoencodeur standard ?
- a) Un VAE n'a pas d'encodeur.
- b) Un VAE génère des données aléatoires.
- c) Un VAE modélise une distribution probabiliste dans l'espace latent.
- d) Un VAE ne peut pas générer de nouvelles données.



Quiz

Quelle est l'une des hypothèses fondamentales du Variational Autoencoder (VAE) concernant l'espace latent ?

- a) Les dimensions de l'espace latent doivent être corrélées pour une meilleure performance.
- b) Les dimensions de l'espace latent doivent être indépendantes et non corrélées.
- c) L'espace latent ne doit pas être utilisé dans un VAE.
- d) L'espace latent doit être discrétisé pour simplifier le modèle.

Ressources

Generative Deep
Learning par David
Foster

Natural Language
Processing with
Transformers par
Lewis Tunstall et al.

Generative AI With
python and
TensorFlow 2 Joseph
Babcock, Raghav Bali