Generatív modellek

Neurális háló

- Valós vektorokat valós vektorokba képző folytonos függvény
- Bemenetek lehetnek: kép, hang, videó, szöveg, ...
- Kimenet például egy klasszifikáció feladat esetén a valószínűségek vektora
- Veszteségfüggvény: például $L(\hat{y}, y)$, vagyis mennyire vagyunk elégedettek a kimenettel az elvárt kimenet ismeretében (pl. négyzetes hiba)
- tanítás (paraméterek hangolása)

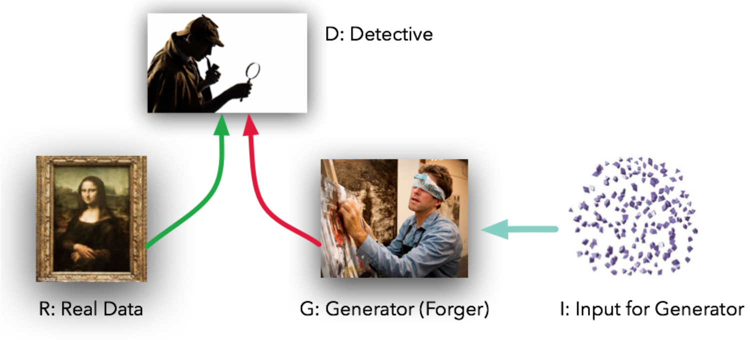
Felügyelet nélküli tanítás

- Nincsenek címkék, csak adatok
- Hasznos adatreprezentáció létrehozása
- A jó reprezentációk szükségesek a hatékony tanításhoz
- Például faktorok azonosítása, dimenzió redukció
- Lényeges információk kinyerése

Generatív modellek

- Cél: egy adott adathalmazhoz szeretnénk olyan modellt építeni, ami generálni tud új adatpontokat
- Ma a két legnépszerűbb generatív modell:
  - Variational Autoencoder (VAE)
  - Generative Adversarial Network (GAN)
- Látens változós modellek: egy megfigyelhetetlen z változóból generálunk
- Eredeti cikkek:
  - VAE: https://arxiv.org/abs/1312.6114
  - GAN: https://arxiv.org/abs/1406.2661

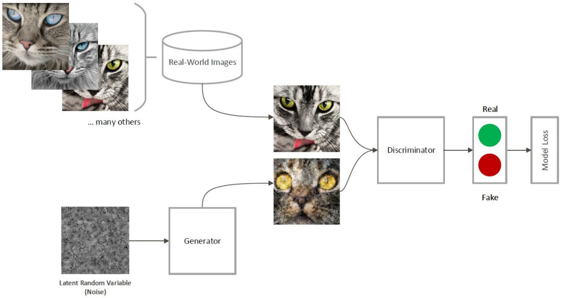
Generative Adversarial Network



Generative Adversarial Network

$$\min_G \max_D V(D, G)$$

$$V(D, G) = \mathbb{E}_{x \sim p_d}[\log D(x)] + \mathbb{E}_{z \sim p_z}[\log(1 - D(G(z)))]$$

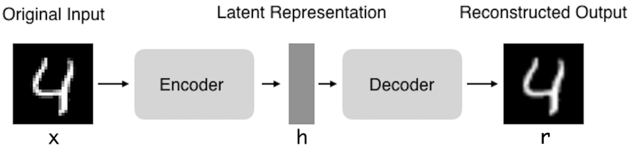


Tanítási nehézségek

- Nem konvergál
- Mode collapse: Nem elég változatosak a generátumok
- Nagyon érzékeny a hiperparaméterekre, véletlen perturbációkra

Autoencoder

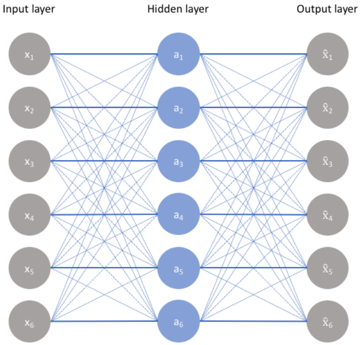
Cél: Az identitás függvény megtanulása



Encoder: $h = f(x)$
Decoder: Az eredeti input rekonsctrukciója; $r = g(h)$
Kellene, hogy $\hat{x} = g(f(x))$ közel legyen x-hez.

Korlátozások

Korlátozásokat kell bevezetnünk, hogy elkerüljük a memorizálást



Az adathalmaz rejtett összefüggéseit szeretnénk megtalálni

Autoencoder

Sok esetben a loss függvény két részből áll:

- Reconstruction loss $\mathcal{L}(x, g(f(x)))$
- Regularizer, ami megakadályozza a memorizálást.

Példák:

- Undercomplete Autoencoder: a látens dimenziót korlátozzuk (bottle neck)
- Sparse autoencoder: az aktivációkra bevezetünk korlátozást.

Tömörítés

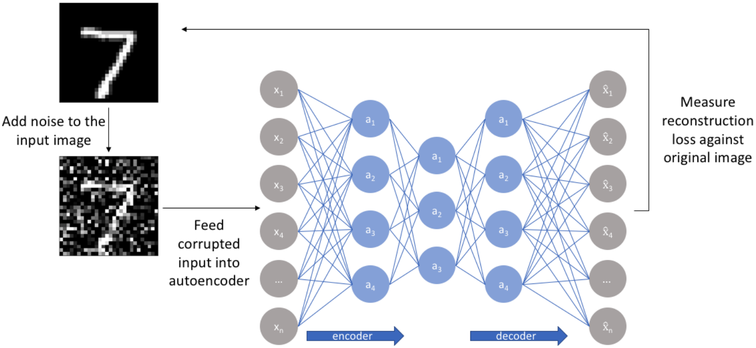
A be- meg kitömörítő függvények

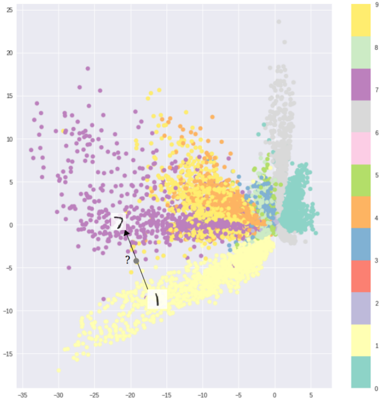
- adatspecifikusak
- veszteségesek
- példák alapján automatikusan tanultak

Nem jó tömörítésre:

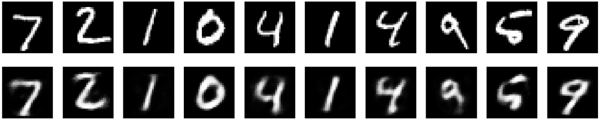
- adatspecifikus
- Nagyon nehéz olyat tanítani, ami jobb mint a JPEG

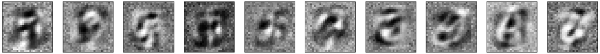
Denoising autoencoder





Training an autoencoder on the MNIST dataset, and visualizing the encodings from a 2D latent space reveals the formation of distinct clusters. This makes sense, as distinct encodings for each image type makes it far easier for the decoder to decode them.





Variational autoencoder

Van egy adathalmazunk: $\mathcal{D} = \{\mathbf{x}^{(i)}\}_{i=1}^{n}$

Feltesszük, hogy az adathalmazunk a következőképpen van generálva:

1. A $\mathbf{z}^{(i)}$ minta egy $p(\mathbf{z})$ prior eloszlásból
2. Az $\mathbf{x}^{(i)}$ pedig egy $p(\mathbf{x} \mid \mathbf{z})$ feltételes eloszlásból.

$$p(x) = \int p(x|z)p(z)\ dz$$

A célunk, hogy minden $x^{(i)}$ adatpont valószínűségét maximalizáljuk.

Variational Autoencoder

Olyan algoritmust szeretnénk, ami akkor is működik, ha

- a $p(x)$ marginális és a $p(z \mid x)$ posterior is nehezen számolható
- nagy az adathalmazunk

$$p(x) \approx \frac{1}{N} \sum_{i=1}^{N} p(x|z_i)$$

Feltesszük, hogy

- $p(z) = \mathcal{N}(z|0, I)$
- $p(x|z) = \mathcal{N}(x|f(z, \theta), \sigma^2 \cdot I)$

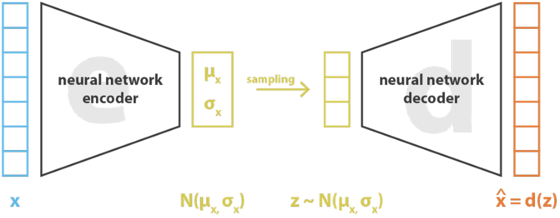
Variational Autoencoder

A gyakorlatban a legtöbb z-re a $p_\theta(x^{(i)} \mid z)$ majdnem 0.

$$p(x^{(i)}) \approx \frac{1}{N} \sum_{i=1}^{N} p(x^{(i)}|z_i)$$

A lényeg: olyan z értékeket szeretnénk mintavételezni, amire valószínű, hogy az $x^{(i)}$-t hozta létre.

A $p_\theta(z \mid x^{(i)})$ posterior nehezen számolható, egy másik eloszlással közelítjük: $q_\phi(z \mid x^{(i)})$.



$$\text{loss} = \| x - \hat{x} \|^2 + \mathrm{KL}[\, N(\mu_x, \sigma_x), N(0, I) \,] = \| x - d(z) \|^2 + \mathrm{KL}[\, N(\mu_x, \sigma_x), N(0, I) \,]$$

Variational lower bound

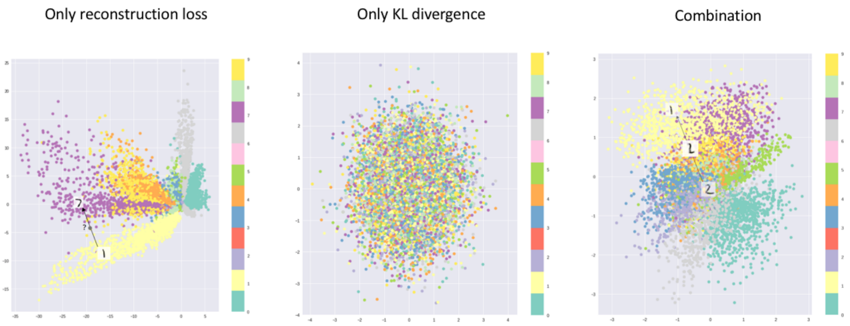
$$\log p_\theta(x^{(i)}) - D_{KL}(q_\phi(z|x^{(i)})\|p_\theta(z|x^{(i)})) =$$

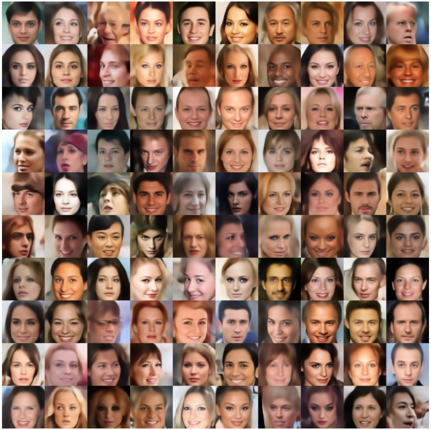
$$= E_{q_\phi(z|x^{(i)})} \log p_\theta(x^{(i)}|z) - D_{KL}(q_\phi(z|x^{(i)})\|p(z)).$$

Jobb oldal a Variational lower bound (ELBO)

$$\mathcal{L}(\theta, \phi, x^{(i)}) = E_{q_\phi(z|x^{(i)})} \log p_\theta(x^{(i)}|z) - D_{KL}(q_\phi(z|x^{(i)})\|p(z)).$$

Látens tér

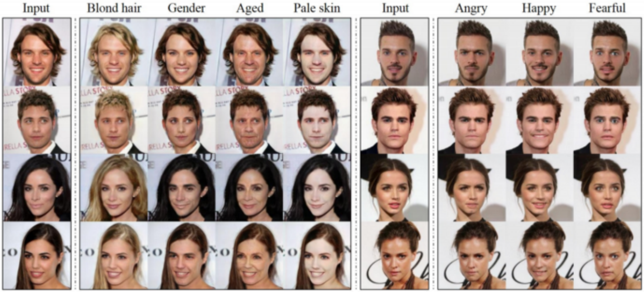




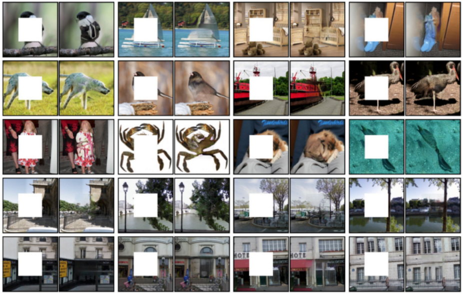
Applications

Figure 7: Generated samples

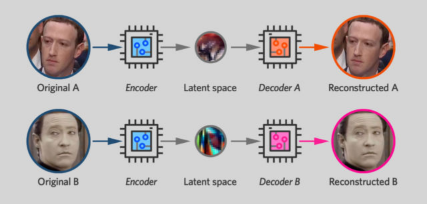
Applications

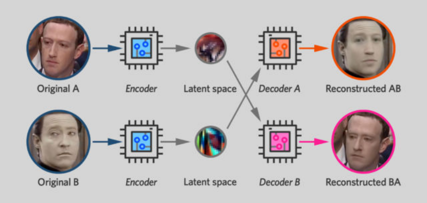


Applications



Deepfake


Original A
Encoder
Latent space
Decoder A
Reconstructed A
Original B
Encoder
Latent space
Decoder B
Reconstructed B


Original A
Encoder
Latent space
Decoder A
Reconstructed AB
Original B
Encoder
Latent space
Decoder B
Reconstructed BA