

Réseaux de neurones adversaires

- Comprendre un nouveau paradigme d'apprentissage
- Découvrir ce que Yann LeCun a appelé la plus grande avancée en ML de cette décennie

- Un réseau de neurones apprend des fonctions complexes
- Toutefois sa métrique d'évaluation est très simple
- Elle dépend aussi du domaine
- Elle est « hardcodée » par l'humain

→ Un apprentissage est toujours limité par une fonction de perte imparfaite.

Évaluer le réseau de neurones avec un autre réseau de neurones et non une fonction de perte. [?]

- Un réseau générateur
- Un réseau discriminateur

Machine Learning

Generative Adversarial Network (**GAN**)

Minimax : entraîner les deux réseaux en maximisant des objectifs opposés.

1. Maximisation de l'efficacité du discriminateur sur les données
2. Maximisation des erreurs du discriminateur sur les données générées

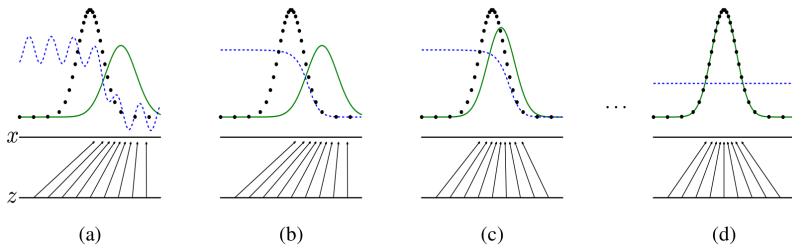
$$\max_D \mathbb{E}_{x \sim \mathcal{P}_r} [\log(D(x))] + \mathbb{E}_{\tilde{x} \sim \mathcal{P}_g} [\log(1 - D(\tilde{x}))]$$

$$\max_D \mathbb{E}_{x \sim \mathcal{P}_r} [\log(D(x))] + \mathbb{E}_{\tilde{x} \sim \mathcal{P}_g} [\log(1 - D(\tilde{x}))]$$

$$\max_D \mathbb{E}_{x \sim \mathcal{P}_r} [\log(D(x))] + \mathbb{E}_{\tilde{x} \sim \mathcal{P}_g} [\log(1 - D(\tilde{x}))]$$

$$\min_G \mathbb{E}_{\tilde{x} \sim \mathcal{P}_g} [\log(1 - D(\tilde{x}))]$$

Convergence



Version adverse du paradigme supervisé classique [?]

- Discriminateur prédit si un décodeur a produit un couple (\tilde{x}, y) ou si un encodeur a produit un couple (x, \tilde{y})

Traductions d'image à image supervisée [?].

Traductions d'image à image **non supervisée** [?].

