

Ananas : Apprentissage de conduite automatique par Deep Q Learning

Deep Learning - Projet

Clément, Grégoire, Nathan

January 31, 2025

Introduction

But du projet

Implémenter une voiture qui apprend à conduire en utilisant le Deep Reinforcement Learning.

Pour des raisons d'efficacité temporelle, reimplémentation de notre propre environnement plutôt que d'en utiliser un déjà tout fait

Source principale : *Playing Atari with Deep Reinforcement Learning*

Reinforcement learning

A quoi ça sert

Reinforcement learning : Agent qui évolue dans un environnement et qui apprend au fil de ses erreurs

Formulation mathématiques

L'agent interagit avec un environnement stochastique \mathcal{E} dans lequel il joue des parties composées d'un ensemble d'état, de score et d'action

A chaque étape, il sélectionne une action a_t parmi un ensemble $A = \{1, \dots, K\}$ d'action légale

A chaque étape, l'agent a accès à un ensemble d'information $x_t \in \mathbb{R}^d$.

Objectif de l'agent

On définit le rendement futur attendu par

$$R_t = \sum_{t'=t}^T \gamma^{t'-t} r_{t'}$$

où T est l'instant auquel le jeu s'arrête.

But de l'agent : maximiser ce rendement.

Optimal action value function

On définit l'optimal action value function $Q^*(s', a')$ qui donne le bon choix à faire en fonction de la situation actuelle.

Cette fonction suit l'équation de Bellman:

$$Q^*(s, a) = \mathbb{E}_{s' \sim \mathcal{E}}[r + \gamma \cdot \max_{a'} Q^*(s', a') | s, a]$$

Deep Q Learning

Equations blabla comment ça marche

Environnement de simulation

Pourquoi recoder notre propre environnement

Inconvénient d'un émulateur : Temps d'accès potentiellement long, peu de flexibilité sur les informations qu'on peut passer à notre modèle donc moins d'experimentations possibles

En recodant notre simulation, on peut contourner tous ces problèmes

Notre simulation

La on met des images de différents exemples

Entraînement, ablation study et tout

Conclusion
