

Rapport Projet Reinforcement Learning : Dueling Network Architecture on Racing Car environnement

DIBE SANA, MERIENNE ALEXIS Polytech Nice Sophia - IAID

November 19, 2022

Abstract

Ce projet à pour objectif d'implémenter une architecture de réseau de neurones à double sortie pour développer un agent capable d'interagir avec l'environnement CarRace proposé par la librairie python Gym. Cette environnement propose un ensemble d'états continue générés aléatoirement et qui correspondent à un circuit de course et un ensemble d'actions également continue. Après avoir discrétiser l'espace d'actions selon différentes finesses, nous utilisons l'algorithme Duel Deep Q-Learning pour déterminer la règle (policy) de notre agent. Cette algorithme, selon les travaux présentés dans l'article Dueling Network Architecture for Reinforcement Learning, présente de meilleures résultats lorsque l'espace d'actions est grand. À ce jour, nous avons pas réussi à développer un agent performant.

I. INTRODUCTION

L'environnement Car Racing de la librairie Gym modélise une voiture de course évoluant sur un circuit. L'espace d'état est générés aléatoirement à chaque épisode. Il représente un circuit de course. Il est composé de frame de taille 96x96. À l'état initial, la voiture de course de situe au milieu de la piste et est à l'arrêt.

L'espace d'action est également continue. Les actions permettent de déterminer la direction de la voiture, la vitesse et le freinage. La direction est implémentée selon une valeur comprise entre -1 et 1. Si cette valeur est à -1, la voiture braque totalement à gauche, si elle est à 1, la voiture braque à droite. La vitesse et le freinage sont des scalaires.

L'objectif d'une voiture de course est de finir le circuit le plus rapidement possible. Le système de récompense modélise cette idée en pondérant les récompenses distribuées à l'agent par le temps de parcours. En effet, à

chaque frame, la voiture reçoit une récompense égale à -0.1. En revanche, pour favoriser la voiture à avancer, la piste est divisée en tuile et à chaque fois que la voiture roule sur une de ces tuiles, elle reçoit une récompense égale à $1000/N$, avec N le nombre totale de tuiles parcourues.

II. METHODS

Maecenas sed ultricies felis. Sed imperdiet dictum arcu a egestas.

- Donec dolor arcu, rutrum id molestie in, viverra sed diam
- Curabitur feugiat
- turpis sed auctor facilisis
- arcu eros accumsan lorem, at posuere mi diam sit amet tortor
- Fusce fermentum, mi sit amet euismod rutrum
- sem lorem molestie diam, iaculis aliquet sapien tortor non nisi
- Pellentesque bibendum pretium aliquet

Table 1: *Example table*

Name		
First name	Last Name	Grade
John	Doe	7.5
Richard	Miles	2

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Text requiring further explanation¹.

III. RESULTS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum.

¹Example footnote

Nunc quis urna dictum turpis accumsan semper.

$$e = mc^2 \quad (1)$$

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

IV. DISCUSSION

i. Subsection One

A statement requiring citation [Figueredo and Wolf, 2009]. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

ii. Subsection Two

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

REFERENCES

- [Figueredo and Wolf, 2009] Figueredo, A. J. and Wolf, P. S. A. (2009). Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330.