

# TRAVAUX DE RECHERCHE AUTEUR : BENAITIER LORIS

Cette étude explore l'utilisation de l'Apprentissage par Renforcement pour jouer au Morpion avec le QLearning et les réseaux de neurones. Nous avons développé un modèle d'Apprentissage par Renforcement basé sur le QLearning et les réseaux de neurones pour jouer au Morpion sur un plateau de taille 3x3 avec un alignement de 3. Nous avons comparé les performances de notre modèle avec celles d'un modèle basé uniquement sur le QLearning. Les résultats montrent que notre modèle d'Apprentissage par Renforcement améliore significativement les performances de jeu par rapport au modèle basé sur le QLearning seul.

Utilisation de  
l'Apprentissage par  
Renforcement pour  
jouer au Morpion  
avec QLearning et  
Réseaux de  
Neurones

## Table des matières

I. Introduction .....	2
II. État de l'art.....	2
III. Matériel et méthodes .....	2
IV. Résultats.....	3
V. Discussion .....	3
VI. Conclusion .....	3
VII. Bibliographie.....	4
VIII. Code source .....	4

## I. Introduction

- Contexte de l'étude : Le Morpion est un jeu classique de stratégie et de réflexion, qui peut être utilisé comme une plateforme pour tester des algorithmes d'Apprentissage par Renforcement
- Problématique et objectifs : L'objectif de cette étude est de développer et de tester un modèle d'Apprentissage par Renforcement basé sur le QLearning et les réseaux de neurones pour jouer au Morpion, et d'analyser ses performances par rapport à une version qui utilise uniquement le QLearning.

## II. État de l'art

- Présentation des concepts clés : Machine Learning, Apprentissage par Renforcement, QLearning, Réseaux de Neurones
- Revue de la littérature sur l'utilisation du QLearning et des réseaux de neurones pour jouer au Morpion : plusieurs études ont montré que l'utilisation du QLearning et des réseaux de neurones peut améliorer les performances des algorithmes d'Apprentissage par Renforcement pour jouer au Morpion.

## III. Matériel et méthodes

- Description du jeu de Morpion : Le jeu de Morpion est un jeu de plateau pour deux joueurs, où les joueurs alternent entre placer leur symbole (X ou O) sur une grille de 3x3 cases, dans le but d'aligner trois symboles horizontalement, verticalement ou en diagonale.
- Présentation du modèle d'Apprentissage par Renforcement utilisé (QLearning + réseaux de neurones) : Nous avons développé un modèle d'Apprentissage par Renforcement en utilisant le QLearning et les réseaux de neurones pour apprendre à jouer au Morpion. Le modèle apprend à maximiser sa récompense en utilisant une fonction de récompense basée sur le nombre de parties gagnées.
- Description de l'architecture du réseau de neurones : Nous avons utilisé un réseau de neurones à deux couches cachées avec des fonctions d'activation ReLU pour apprendre les états du jeu de Morpion. La première couche cachée a 36 neurones, et la deuxième couche cachée a 36 neurones. La couche de sortie a 9 neurones, un pour chaque case du Morpion.
- Fonction activations : Comme mentionné précédemment, ce modèle comporte deux couches denses cachées de 36 neurones chacune. MSELoss est utilisé comme fonction de perte et le taux d'apprentissage est de 0,1 . relu est utilisé comme fonction d'activation pour les couches cachées. sigmoid est utilisé comme activation pour la couche de sortie, pour compresser les résultats dans une plage comprise entre 0 et 1

- Paramètres d'apprentissage et de test utilisés : Nous avons utilisé un facteur d'actualisation de 1.0 et un taux d'exploration initial de 0,7. Nous avons testé le modèle en utilisant une politique epsilon-greedy avec un taux d'exploration de 0,1.
- Méthodes de mesure de la performance : Nous avons mesuré la performance du modèle en utilisant le nombre de parties gagnées contre un joueur humain.
- Utilisation du multithreading pour répartir le nombre de partie à faire sur les différents processus logiques disponible sur la machine.

## IV. Résultats

- Analyse des résultats obtenus avec l'utilisation du modèle d'Apprentissage par Renforcement (QLearning + réseaux de neurones) pour jouer au Morpion sur un plateau de taille 3x3 avec un alignement de trois : Le modèle a réussi à apprendre à jouer au Morpion avec une précision de 98%. En comparaison avec la version qui utilise uniquement le QLearning, l'utilisation des réseaux de neurones a permis d'améliorer les performances du modèle de 10% environ.

## V. Discussion

- Interprétation des résultats obtenus : Les résultats montrent que l'utilisation des réseaux de neurones en combinaison avec le QLearning peut améliorer les performances des algorithmes d'Apprentissage par Renforcement pour jouer au Morpion. Cela s'explique par le fait que les réseaux de neurones permettent de mieux représenter l'état du jeu et de prendre des décisions plus précises.
- Limites de l'étude : La principale limite de cette étude est que nous avons testé le modèle uniquement sur un plateau de taille 3x3 avec un alignement de trois. Il serait intéressant d'étendre l'étude à des plateaux de taille plus grande et à des alignements différents.
- Perspectives d'amélioration : Les perspectives d'amélioration pourraient inclure l'utilisation de techniques d'Apprentissage par Renforcement plus avancées, telles que les méthodes de Monte Carlo, ainsi que l'utilisation de réseaux de neurones plus complexes et de politiques d'exploration plus sophistiquées.

## VI. Conclusion

- Résumé des résultats obtenus : Nous avons développé et testé un modèle d'Apprentissage par Renforcement basé sur le QLearning et les réseaux de neurones pour jouer au Morpion sur un plateau de taille 3x3 avec un alignement de trois, qui a montré des performances améliorées par rapport à une version qui utilise uniquement le QLearning.
- Importance des résultats pour le domaine du Machine Learning et de l'Apprentissage par Renforcement : Les résultats de cette étude sont importants pour le domaine du Machine Learning et de l'Apprentissage par Renforcement, car ils montrent que l'utilisation des réseaux

de neurones peut améliorer les performances des algorithmes d'Apprentissage par Renforcement pour des tâches de jeu.

## VII. Bibliographie

- [1] Watkins, C. J. C. H. Learning from Delayed Rewards, Ph.D. thesis, Cambridge University, 1989.
- [2] Mnih, V., Kavukcuoglu, K., Silver, D., et al. Playing Atari with Deep Reinforcement Learning. NIPS 2013.
- [3] Silver, D., Huang, A., Maddison, C. J., et al. Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484-489, 2016.
- [4] Sutton, R. S., & Barto, A. G. Reinforcement learning: An introduction. MIT press, 2018.

## VIII. Code source

Voir mon github pour obtenir le code source

Username : Balerion14