

Liebe Miriam, lieber Nico 🤗

Hier unsere Feedback zu eurer tollen Zwischenpräsentation von heute. Wir hoffen das Feedback gibt weitere Ideen und wünschen euch weiterhin viel Erfolg und Ausdauer beim Fliegen. 🕊️

Liebe Gruess

Lukas & Damaris 👨‍👩‍👧

1. Da Flappy Bird ein Spiel ist, das unendlich gespielt werden kann muss der Agent ziemlich weit in die Zukunft schauen / zukünftige Rewards bedenken. In eurem Code habe ihr einen Discount Faktor von 0.99 implementiert. Damit würde der Agent eigentlich sehr weitsichtig handeln. Vielleicht könntet ihr trotzdem noch verschiedene Discount Faktoren (Gamma) ausprobieren? 🕒
2. Ihr könntet das Epsilon (ist bei euch im linearen Q-Learning aktuell auf 0.5 eingestellt) anfangs zuerst auf einen höheren Wert einstellen und nah bis nah den Wert gegen null gehen lassen (gemäss e-greedy policy, die ihr auch im File keras_deepq_nn genutzt habt). So probiert eurer Agent anfangs mehr aus, aber ab einem bestimmten Zeitpunkt exploriert er dann mehrheitlich. 📈
3. Im Sport kehrt ein Spieler ja jeweils in die Mitte zurück (z.B. Squash) um optimal für den nächsten Ball des Gegners positioniert zu sein. Vielleicht könntet ihr den Agenten belohnen, wenn er relativ mittig fliegt? (in einem gewissen Bereich um die Mitte herum) 🏸
4. Bezüglich des Belohnungssystems: ihr habt im versucht den Agenten durch bestrafen zum Optimum hinzuführen. Wie wäre es auszuprobieren mehr Belohnungen zu geben statt zu strafen? 🍰
5. Im keras_deepq_nn könntet ihr noch weitere Layer und auch andere Aktivierungsfunktionen (z.B. relu, leaky relu, softmax, tanh...) ausprobieren. 📈

Kleines letztes Feedback (in Anbetracht dessen, dass wir nochmals eine Präsi haben):
Es war ein bisschen ein hin und her zwischen euch zwei. Vielleicht könntet ihr die Slides so legen, dass es etwas weniger Wechsel beim Präsentieren gibt? 😊