Uma imagem contendo Interface gráfica do usuário

Descrição gerada automaticamente

**Correlação entre o valor dos hiperparâmetros e a progressão de recompensas por episódios.**

Nesse gráfico fizemos os ajustes dos 3 principais hiperparâmetros para Q-Learning. Com base nele, podemos observar que:

- Alpha: Quanto maior o alpha, foi-se observado mais oscilações e uma demorada convergência para zero, o que indica que o algoritmo demorou mais para aprender.

- Gamma: Um maior gamma representa uma maior priorização das recompensas de longo prazo, o que resulta em um crescimento gradual das recompensas por episódio. Assim, se tivermos um gamma menor, haverá menos flutuações grandes e aumentará o foco nas recompensas imediatas.

- Epsilon: O Epsilon diz respeito a exploração da árvore de possibilidades, então, caso o hiperparâmetro seja maior, significa que haverá uma exploração inicial maior e uma curva mais irregular. Consequentemente, caso o épsilon seja menor, a curva será mais suave.