

Dit is een screenshot na 10000 rondes, we zien dat de total rewards ongeveer hetzelfde blijft. Dit kan 1 van 2 dingen betekenen. 1. De agent leert niet of 2. de agent leert slecht dit kan onder andere door parameters komen of door mijn implementatie van de klassen.

Het zou verbetert kunnen worden door de hyperparameters te veranderen en een idee zou ook kunnen zijn om short en long term memory toe te voegen.