

1. Onderzoeksresultaten

Hieronder zal worden beschreven wat de resultaten zijn van het beste eindmodel, dat een 3x3x3 yard indeelt met containers.

Verzamelde data

Om de resultaten van het model te analyseren, worden er twee aspecten geanalyseerd. Het eerste aspect zijn de voorspellingen en visualisatie van het model. Dit aspect geeft inzicht in hoe de containers in de yard zijn geplaatst. Het tweede aspect is de leerprestaties van het model, die worden gemeten en weergegeven in grafieken door middel van gegevens. Deze gegevens worden verzameld tijdens het leerproces van het model. Hierdoor kan worden geanalyseerd of het model het verwachte leerpatroon heeft gevolgd.

Reward-functie

Bij het maken van de reward-functie is er gekeken naar de volgende punten:

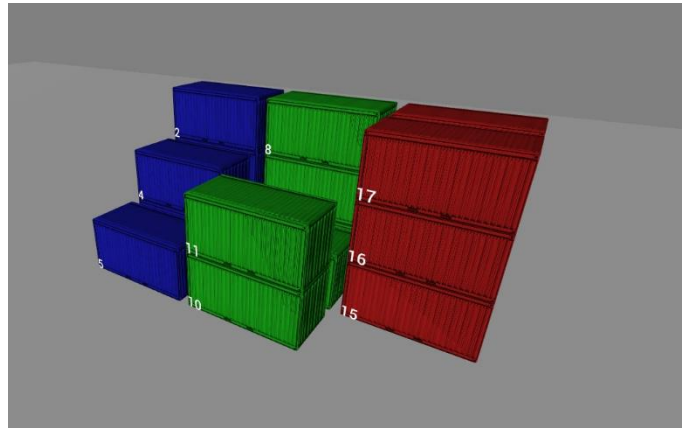
- De container mag niet boven de maximale hoogte komen.
- De container mag niet op een container worden gestapeld die voor een ander zeevaartschip bedoeld is.
- De container mag niet worden ingeboxed

Bij het kijken naar de maximale hoogte wordt alleen een penalty gegeven als de agent een container op een container wil plaatsen die al op de maximale hoogte staat. Hierbij wordt de bovenste container overschreven als dit het geval is. In het geval dat de maximale hoogte nog niet is bereikt wordt er gekeken naar voor welk schip de container eronder bestemd is. Als het voor hetzelfde schip bestemd is, wordt er een beloning toegekend. Als het voor een ander schip bestemd is zal er een penalty worden gegeven. Daarnaast wordt er een beloning gegeven als een container op een lege plek wordt neergezet. Vervolgens, bij het plaatsen van een container op een lege plek, is er een check voor het inboxen van de container. Deze check geeft een penalty als er wordt geconstateerd dat er een container is ingeboxed. Het neerzetten op een lege plek en checken voor inboxen wordt bij elkaar opgeteld en als één waarde teruggegeven als een beloning of penalty.

Visualisatie resultaten

Nadat de agent geleerd heeft, geeft het model een oplossing voor het neerzetten van de containers. De indeling van de containers, die het model maakt na trainen, wordt visueel weergegeven. Deze uitkomst kan vervolgens worden vergeleken met de aanvankelijke situatie en met andere oplossingen om te bepalen hoe efficiënt het model de containers heeft geplaatst. Bovendien kan de uitkomst worden gebruikt als referentie voor toekomstige verbeteringen en vergelijkingen met andere methoden.

In Figuur 1 is te zien dat de containers gesorteerd zijn op kleur, waarbij blauw, groen en rood worden gebruikt als kleuren voor de verschillende zeevaartschepen waarvoor de containers bestemd zijn. Hierdoor is het duidelijk te zien welke containers voor welk zeevaartschip zijn bedoeld. In het figuur is te zien dat de containers zijn gegroepeerd volgens kleurcode, en zo zijn opgesteld dat ze kunnen worden opgepakt door een stacker via de lange zijde. Dit geeft een duidelijke visuele weergave van hoe de containers zijn georganiseerd.



Figuur 1 - Visualisatie output model

Grafieken leerproces modellen

Om te beoordelen of een model effectief leert of vooruitgang boekt tijdens het leerproces, kan men gebruik maken van het visualiseren van de data dat verzameld wordt tijdens het leerproces van het model. In Figuur 2, 3 en 4 worden een aantal visualisaties weergegeven die informatie geven over de prestaties van het model.

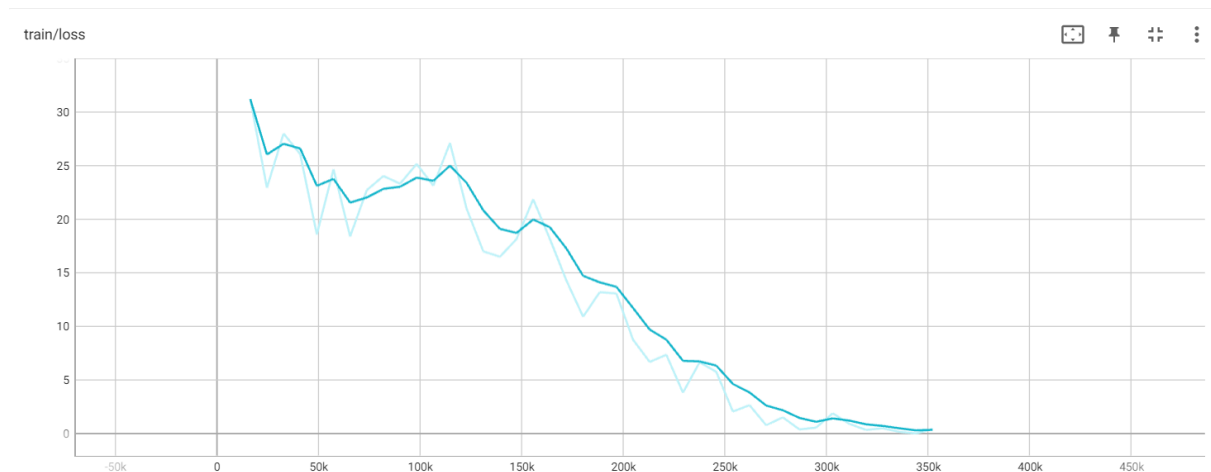
In Figuur 2 worden de scores tijdens het trainen van het model per iteratie getoond, deze resultaten geven aan dat het model effectief is getraind en goed presteert wat betreft de gekozen inputvariabelen.

time/	
fps	1054
iterations	43
time_elapsed	334
total_timesteps	352256
train/	
approx_kl	0.020521311
clip_fraction	0.173
clip_range	0.2
entropy_loss	-0.469
explained_variance	0.985
learning_rate	0.0005
loss	0.456
n_updates	420
policy_gradient_loss	-0.0126
value_loss	1.11

Figuur 2 - Prestaties model

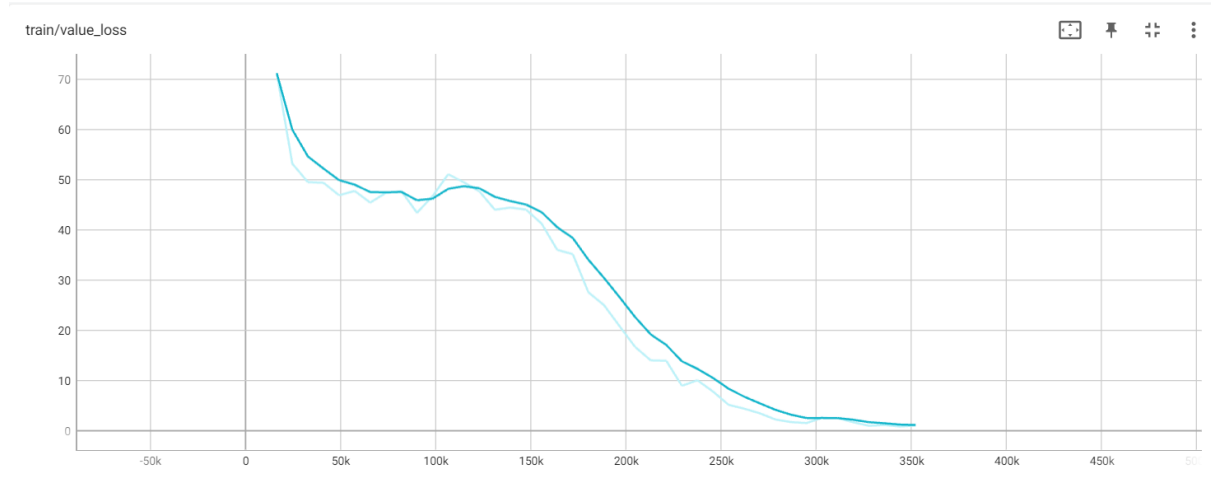
In de Figuren 3 en 4 zijn de waarde van de loss functie en de value loss functie geplot over het aantal timesteps. De doorzichtige lijn is de werkelijke waarde aan, terwijl de dikgedrukte lijn een gladgetrokken versie van de doorzichtige lijn is.

In Figuur 3 is de loss van het model te zien. Door te kijken naar hoe de loss waardes staan tegenover de timesteps die worden genomen, is te zien of het model leert. In dit geval zien we dat het model goed leert.



Figuur 3 De loss van het model

In Figuur 4 is de value loss van het model te zien, deze beschrijft de stabiliteit van het model over het aantal genomen timesteps.



Figuur 4 De value loss van het model

[Fout! Verwijzingsbron niet gevonden.](#)

Analyse

De prestaties van het PPO-model zijn grondig geanalyseerd om te bepalen hoe effectief het is in het indelen van de containers. Dit is gedaan door middel van een evaluatie van de visualisatie van de voorspellingen en de verzamelde gegevens van het leerproces. In de visualisatie van de geplaatste containers (zie [Figuur 2](#)), is te zien dat iedere kleurcode achter elkaar is geplaatst via de lange zijde. In de afgebeelde grafieken wordt opgemerkt dat bij een uiteindelijke timestep van circa 350.000, beide grafieken de nul benaderen en de curve afvlakt.