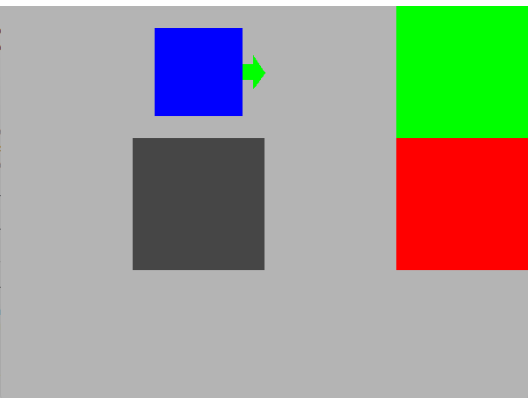


Explainable AI

Reinforcement learning explained

Finn Alberts, Laurent Dassen en Noud Wijngaards, studenten te Zuyd Hogeschool, Heerlen



Dit project is uitgevoerd binnen de minor Artificial Intelligence van Zuyd Hogeschool, Heerlen.

Betrokken organisaties:
Lectoraat Data Intelligence, Heerlen

Opdrachtgever: docent-onderzoeker
Shannen Dolls

**Lectoraat Data Intelligence,
Faculteit ICT
Zuyd Hogeschool**
Nieuw Eyckholt 300, 6419 DJ
Heerlen
T +31 (0)45 400 6400
E lectoraat-di@zuyd.nl
www.zuyd.nl

AANLEIDING

Artificial Intelligence (AI) en reinforcement learning (RL) worden steeds vaker en breder ingezet. Bij het grotere publiek ontbreekt echter kennis over de werking van deze technieken. Vanuit het lectoraat Data Intelligence is de wens om de kennis bij men hierover de verbreden, met behulp van explainable AI.

DOELSTELLING

Het doel is om met explainable AI en reinforcement learning op een laagdrempelige manier uit te leggen.

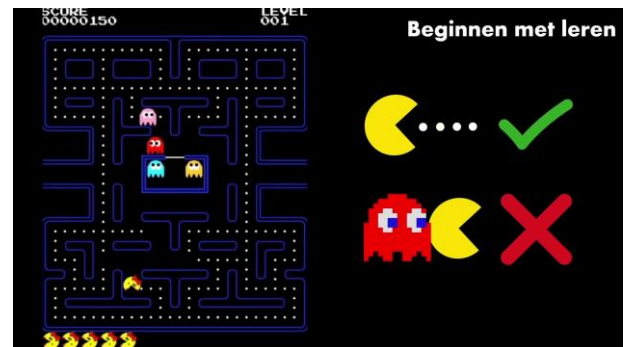
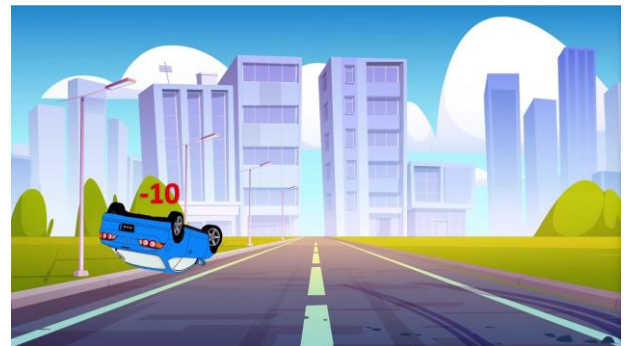
AANPAK

Dit project is op een iteratieve wijze uitgevoerd met invloed van de agile werkmethode. De producten die zijn gerealiseerd in de project zijn een demonstrator en explainer video. De explainer video geeft duidelijkheid over de werking van reinforcement learning aan de hand van Pacman. De demonstrator laat mensen hands-on aan de slag gaan met reinforcement learning. Door de combinatie van de explainer video met deze demonstrator wordt ook vanuit educatief perspectief een goede uitleg gegeven over reinforcement learning.

RESULTATEN

In de gerealiseerde explainer video wordt op een laagdrempelige manier een algemene uitleg gegeven over de werking van reinforcement learning. Deze uitleg wordt aangevuld met de mogelijkheden en beperkingen van reinforcement learning, aan de hand van een gerealiseerde Pacman demonstrator en andere demonstrators uit de literatuur.

Deze explainer video (te zien in onderstaande afbeeldingen) wordt aangevuld door de custom rewards demonstrator. In deze demonstrator heeft iemand zelf inspraak op de rewardfunctie voor een simpele simulatie (de demonstrator is te zien in de onderste afbeelding links). De demonstrator simuleert vervolgens het effect van de gekozen reward functie.



CONCLUSIE

Door middel van de gerealiseerde explainable AI kan de kennis over reinforcement learning bij het grotere publiek worden verbreed. Deze explainable AI bestaat uit een explainer video en een demonstrator, wat ook volgens educatief perspectief voor een goede uitleg zorgt. Hiermee wordt het oorspronkelijke doel bereikt.

**Zuyd
Research**

**ZU
YD**