

Primena Reinforcement Learning-a za AI protiv AI fudbalsku simulaciju u 2D okruženju (4v4)

...

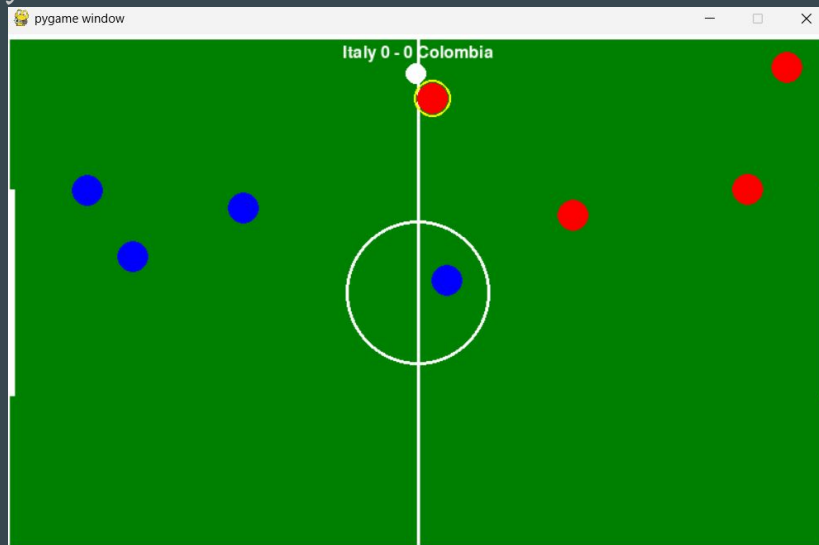
Patrik Barši
SV7/2022

Sadržaj

- Okruženje
- Agenti
- Treniranje
- Reward funkcije
- Simulacija turnira
- Moguća poboljšanja

Okruženje

- Arena od 800x500 piksela
- 4 igrača - 3 field igrača - 1 golman
- Moguće akcije: kretanja u 8 direkcija, pas, šut, dribling
- Interakcija između igrača zavisi od njihovih statistika
- Fizika



Agenti

- Svaki igrač ima svog agenta
- Golmani su limitirani u kretanju
- Agenti koriste Deep Q-Learning sa aproksimacijom sa neuronskim mrežama

Treniranje

- Za treniranje je korišćeno 1000 epizoda / 400 koraka
- Epsilon - od 0.99 do 0.05
- Epsilon decay - 250000
- Learning rate - 0.001
- Gamma - 0.99
- Trajanje - Zavisi od broja epizoda i koraka

Reward funkcije

- Inicijalno jednostavan
 - Gol, distanca od lopte
- Kasnije kompleksniji
 - Gol, posedovanje lopte, distanca od lopte, pozicija na terenu, ograničenje nagrade

Simulacija turnira

- Simulacija 10 puta - Random seeding
- Rezultati:
 - Argentina 7
 - Brazilija 1
 - Francuska 1
 - Italija 1
- Na ove rezultate utiču tehnički izazovi

Moguća poboljšanja

- Još kompleksnija reward funkcija
- Kompleksniji algoritam za RL
- Fine tuning parametara
- Treniranja na više epizoda/koraka