

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 2857

# **Automatizirani razvoj agenata za različita okruženja**

Paulo Sanković

Zagreb, lipanj 2022.

Zagreb, 11. ožujka 2022.

## DIPLOMSKI ZADATAK br. 2857

Pristupnik: **Paulo Sanković (0036509900)**

Studij: Računarstvo

Profil: Računarska znanost

Mentor: doc. dr. sc. Marko Đurasević

Zadatak: **Automatizirani razvoj agenata za različita okruženja**

### Opis zadatka:

Proučiti različita okruženja poput jednostavnih igara ili fizikalnih simulatora za koje je moguće razviti agente. Proučiti različite metode i modele poput umjetnih neuronskih mreža koje se mogu iskoristiti za upravljanje agentom u takvim okruženjima. Proučiti postojeća programska rješenja koja pružaju simulaciju pojedinih okruženja te istražiti mogućnosti povezivanja s takvim okruženjima. Razviti programski okvir koji omogućuje automatski razvoj agenata za odabrana okruženja te grafički prikazuje ponašanje razvijenih agenata. Ocijeniti efikasnost razvijenih agenata za odabrana okruženja. Predložiti moguća poboljšanja predloženog postupka s ciljem razvoja efikasnijih agenata. Radu priložiti izvorne programske kodove, dobivene rezultate uz potrebna objašnjenja i korištenu literaturu.

Rok za predaju rada: 27. lipnja 2022.

*Zahvaljujem se mentoru doc. dr. sc. Marko Đurasević na pomoći i razumijevanju tijekom izrade ovog diplomskog rada. Također posebno bih se zahvalio svojoj obitelji na pruženoj podršci tijekom svih godina studiranja.*

# SADRŽAJ

<b>1. Uvod</b>	<b>1</b>
<b>2. Podržano učenje</b>	<b>2</b>
2.1. Neuronske mreže . . . . .	4
2.1.1. Potpuno povezane neuronske mreže . . . . .	4
2.1.2. Konvolucijske neuronske mreže . . . . .	4
2.2. Algoritmi podržanog učenja . . . . .	4
2.2.1. Deep Q Learning . . . . .	4
2.2.2. Double Deep Q Learning . . . . .	4
2.2.3. Double Deep Q Learning . . . . .	4
<b>3. OpenAI Gym</b>	<b>5</b>
3.1. Okruženja . . . . .	5
<b>4. Implementacija</b>	<b>6</b>
<b>5. Moguća poboljšanja</b>	<b>7</b>
<b>6. Zaključak</b>	<b>8</b>
<b>Literatura</b>	<b>9</b>

# 1. Uvod

Inteligentni sustavi sposobni su analizirati, razumjeti i učiti iz dostupnih podataka putem posebno dizajniranih algoritama umjetne inteligencije. Zahvaljujući dostupnosti velikih skupova podataka, razvoja tehnologije i napretka u algoritmima došli smo do razine gdje nam razvijeni modeli mogu uvelike poslužiti u svakodnevnom životu.

Posebno je zanimljiva problematika u kojoj je bez znanja o pravilima i funkcioniranju specifične okoline, potrebno konstruirati specijaliziranog agenta koji se nalazi u određenom stanju okoline i ponavlja korake izvršavanja optimalne akcije i prijelaza u novo stanje okoline. Za svaku akciju agent prima određenu nagradu - mjeru koja označava koliko su akcije agenta ispravne za tu okolinu i koliko je napredak agenta ispravan. Agent izvršava akcije i prelazi u nova stanja sve dok se ne nađe u terminalnom (završnom) stanju. Dakle, cilj agenta u okolini jest pronaći optimalnu strategiju koja će maksimizirati očekivanu dobit (nagradu) u određenom vremenskom okviru.

Područje strojnog učenja koje se bavi prethodno navedenom problematikom naziva se podržano učenje (engl. *Reinforcement Learning*). Agenti koji se prilično dobro ponašaju u takvim okolinama možemo implementirati pomoću različitih algoritama podržanog učenja koji se temelje na umjetnim neuronskim mrežama (engl. *Artificial Neural Networks*). Umjetne neuronske mreže su dobri aproksimatori funkcija i najbolje se ponašaju u okolini koja ima kompozitnu strukturu gdje vrlo kvalitetno duboki model predstave kao slijed naučenih nelinearnih transformacija.

U sklopu ovog rada bilo je potrebno proučiti i razumjeti metode i algoritme podržanog učenja i funkcioniranje umjetnih neuronskih mreža. Nadalje, bilo je potrebno istražiti, proučiti i naposljetku implementirati neke od algoritama podržanog učenja koji se zasnivaju na umjetnim neuronskim mrežama i koje je trebalo uklopiti u okoline koje su prikladne za simulaciju i testiranje ponašanja naučenih agenta.

Programsko rješenje implementirano je u programskom jeziku *Python*, primarno koristeći *PyTorch* biblioteku (engl. *library*) zajedno s ostalim korisnim bibliotekama poput *numpy*, *stable-baselines3*... Za simulaciju i testiranje ponašanja agenata (razvijenih modela) u posebnom okruženju korištena je biblioteka *OpenAI Gym*.

## 2. Podržano učenje

Strojno učenje (engl. *Machine Learning*) jest grana umjetne inteligencije (engl. *Artificial Intelligence*) koja se može definirati kao skup metoda koje u podacima mogu automatski otkrivati obrasce, i potom te otkrivene obrasce iskoristavati pri budućem predviđanju podataka, ili obavljati druge zadatke odlučivanja u prisustvu nesigurnosti Čupić (2022). Drugim riječima, bez eksplicitnog programiranja moguće je napraviti sustave koji funkcioniraju kao ljudski mozak - imaju pristup podacima, koriste ih za učenje i samim time bolje razumiju entitete, domene i veze između podataka.

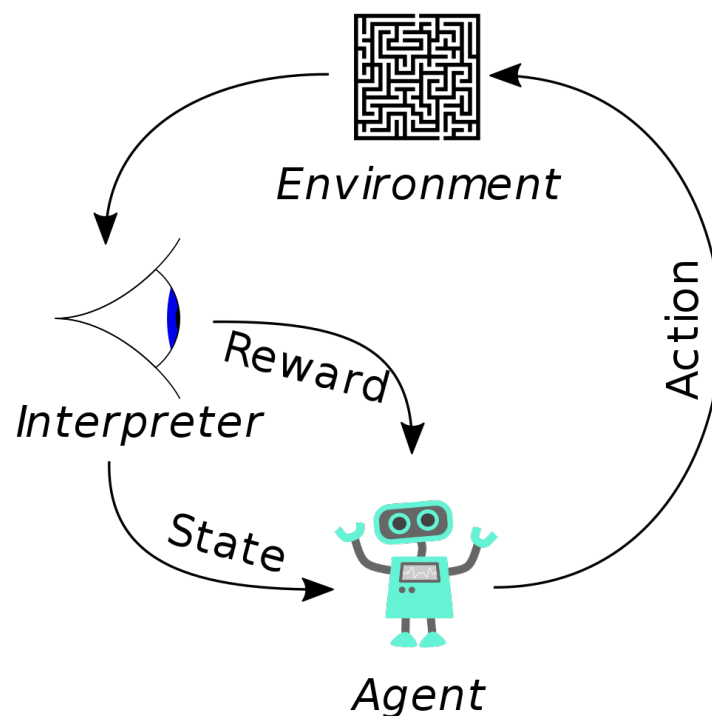
Strojno učenje dijeli se na 3 podvrste: nadzirano učenje, nenadzirano učenje i podržano (ojačano) učenje. Nadzirano učenje (engl. *supervised learning*) karakterizira učenje modela nad testnim podacima koji su označeni. Model točno zna da za određeni ulaz mora vratiti izlaz koji je istovjetan unaprijed pridruženoj oznaci. Algoritam mjeri točnost kroz funkciju gubitka, prilagođavajući se sve dok se izračunata razlika izlaza modela i stvarnog izlaza (pogreška) ne smanji u određenoj mjeri. U nenadziranom učenju (engl. *unsupervised learning*) za razliku od nadziranog, posjedujemo podatke bez zadanog izlaza - podatci su dani bez ciljane vrijednosti i u tim situacijama treba pronaći određenu pravilnost. Postupci poput grupiranja, smanjenja dimenzionalnosti, otkrivanja veza između primjeraka... pripadaju nenadziranom učenju.

Posebna i nama najzanimljivija podvrsta strojnog učenja jest podržano učenje (engl. *reinforcement learning*). Podržano učenje bavi se optimizacijom ponašanja agenta koji je u interakciji s okolinom (u kojoj se nalazi) i koji na temelju informacija koje dobiva iz okoline izvršava akcije, i kao odgovor na svaku akciju dobiva nagradu ili kaznu. Za razliku od prethodno dvije navedene podvrste koje mapiraju ulazne podatke na određeni format izlaza, u podržanom učenju je naizraženije učenje iz iskustva koje je čovjeku kao biću ključan način na koji se razvija. Od najranije dobi, bića nastoje shvatiti i razumjeti okolinu u kojoj se nalaze na temelju niza aktivnosti kojima utječu na okolinu i opažanja kako okolina pri toj interakciji utječe na nas.

Za potpuno razumijevanje podržanog učenja, bitno je u potpunosti razumjeti glavne pojmove. Okolina (engl. *environment*) označava svijet u kojem se agent nalazi i s ko-

jim interaktira. Stanje (engl. *state*) reprezentira presjek okoline u određenom trenutku. Agentu korisna informacija jest nagrada (engl. *reward*) koja predstavlja povratnu informaciju okoline. Način na koji agent bira akciju (engl. *action*) koju želi izvesti naziva se politika (engl. *policy*).

Cilj podržanog učenja jest naći optimalnu strategiju (niz optimalnih akcija) koje maksimiziraju ukupnu (kumulativnu) nagradu. U svakom koraku interakcije agenta s okolinom, agent prima opis stanja okoline u kojoj se nalazi. S obzirom na to stanje, izvršava akciju koja vrši neku promjenu nad okolinom i prebacuje ju u novo stanje. Agent prima povratnu informaciju od okoline koja reprezentira koliko je odabrana akcija u skladu sa stanjem okoline. Prethodno opisana interakcija agenta s okolinom vizualizirana je na slici 2.1.



**Slika 2.1:** Prikaz ciklusa i interakcije agenta s okolinom

Onaj dio s towards science oko exploration...

## **2.1. Neuronske mreže**

### **2.1.1. Potpuno povezane neuronske mreže**

### **2.1.2. Konvolucijske neuronske mreže**

## **2.2. Algoritmi podržanog učenja**

### **2.2.1. Deep Q Learning**

### **2.2.2. Double Deep Q Learning**

### **2.2.3. Double Deep Q Learning**



## **3. OpenAI Gym**

### **3.1. Okruženja**

## **4. Implementacija**

## **5. Moguća poboljšanja**

## **6. Zaključak**

Bože pomози hehe.

# LITERATURA

Marko Čupić. *Umjetna Inteligencija - Uvod u strojno učenje*. 2022.  
<http://java.zemris.fer.hr/nastava/ui/ml/ml-20220430.pdf>.

## **Automatizirani razvoj agenata za različita okruženja**

### **Sažetak**

Sažetak na hrvatskom jeziku.

**Ključne riječi:** Ključne riječi, odvojene zarezima.

### **Title**

### **Abstract**

Abstract.

**Keywords:** Keywords.