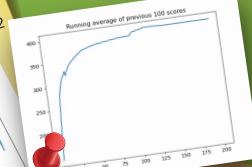
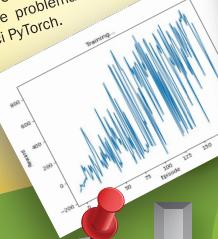


## Rezultati i Diskusija

Naša analiza rezultata pokazuje da su DQN i PPO algoritmi sposobni za uspešno rešavanje problema Box2D CarRacing-v2 koristeći PyTorch.



## Evaluacija Performansi

Evaluacija performansi se vrši kroz više epizoda simulacije Box2D CarRacing-v2. Koristimo metriku ukupne nagrade koju agent ostvaruje tokom epizoda kao meru uspešnosti. Poredujemo performanse DQN i PPO algoritama na osnovu prosećne ukupne nagrade i njihove stabilnosti tokom treninga koristeći PyTorch.

## ZAKLJUČAK

U ovom istraživanju smo poučavali primenu DQN i PPO algoritama koristeći PyTorch u rešavanju problema Box2D CarRacing-v2. Naši rezultati ukazuju na uspešnost oba algoritma u postizanju zadovoljavajućih performansi. Ovo istraživanje pruža osnovu za dalje istraživanje i primenu ovih algoritama u rešavanju složenijih problema autonomne vožnje koristeći PyTorch.



Dušan Rožić

SV 80/2020

## Podešavanje Parametara

Da bismo postigli optimalne rezultate, pažljivo podešavamo parametre za oba algoritma koristeći PyTorch. Ovi parametri uključuju stopu učenja, faktor smanjenja, veličinu memoriskog bufera, broj iteracija učenja i druge relevantne parametre. Izvršavamo niz eksperimenata kako bismo pronašli najbolje parametre za svaki algoritam.

BATCH\_SIZE = 64  
GAMMA = 0.99  
EPS\_START=0.9  
EPS\_END=0.05  
EPS\_DECAY=7000  
TAU = 0.005  
LR = 1e-4

## Proporcionalno-proporcionalne optimizacije

PPO je algoritam optimizacije politike koji pripada klasi algoritama poznatih kao "actor-critic" metode. Implementiramo PPO model koji koristi neuronsku mrežu za parametrizaciju politike i vrednosne funkcije. Koristimo PPO algoritam za optimizaciju politike na osnovu sakupljenih uzoraka koristeći PyTorch.

## DUBOKE Q-MREZE

DQN je reinforcement learning algoritam koji koristi neuronske mreže za aproksimaciju Q-funkcije. U našem istraživanju, implementiramo DQN model sa konvolucijskim slojevima koji prima slikevni prikaz stanja automobila kao ulaz. Koristimo epsilon-pohlepnu strategiju za istraživanje i eksperimentalno određujemo optimalne mreže koristeći PyTorch.

## UVOD

Box2D CarRacing-v2 predstavlja izazovno okruženje u kojem autonomni agent uči kako da efikasno upravlja automobilom i uspešno završi datu stazu. DQN i PPO su popularni reinforcement learning algoritmi koji su se pokazali uspešnim u rešavanju sličnih problema. U ovom istraživanju proučavamo njihovu primenu koristeći PyTorch kao okvir za duboko učenje.

## METODOLOGIJA

Koristimo programski jezik Python i biblioteku PyTorch za implementaciju DQN i PPO modela. Koristimo takođe biblioteku OpenAI Gym koja pruža simulacijsko okruženje Box2D CarRacing-v2. Implementacija modela uključuje definisanje neuronskih mreža i odgovarajućih funkcija za učenje, istraživanje i donošenje odluka.