RL is TERMINAL

5/1/2024

COS 435: Reinforcement Learning

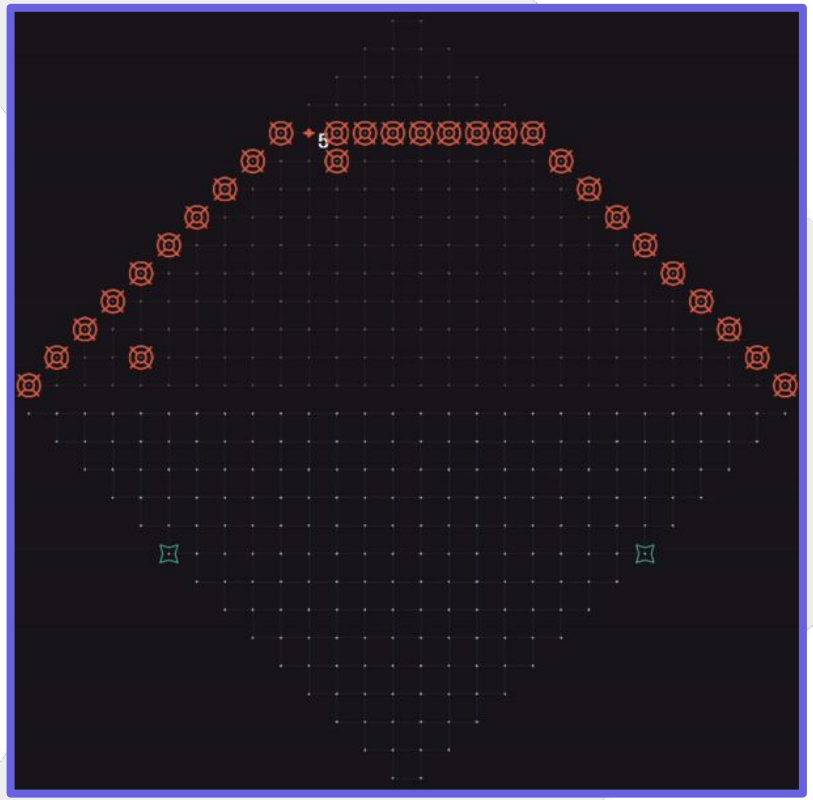
Professor Mengdi Wang, Professor Benjamin Eysenbach

Authors: Yagiz Devre, Ishaan Javali, Shivam Kak

What is Terminal

Terminal is a type of Tower Defense game

- **Environment:** Diamond-shaped board
- **Actions:** 3 Attack and 3 Defense units:
- **Budget:** Each player has a fixed budget
- **Objective:** Breach the opponent's edges and reduce the opponent's Health Points to zero.

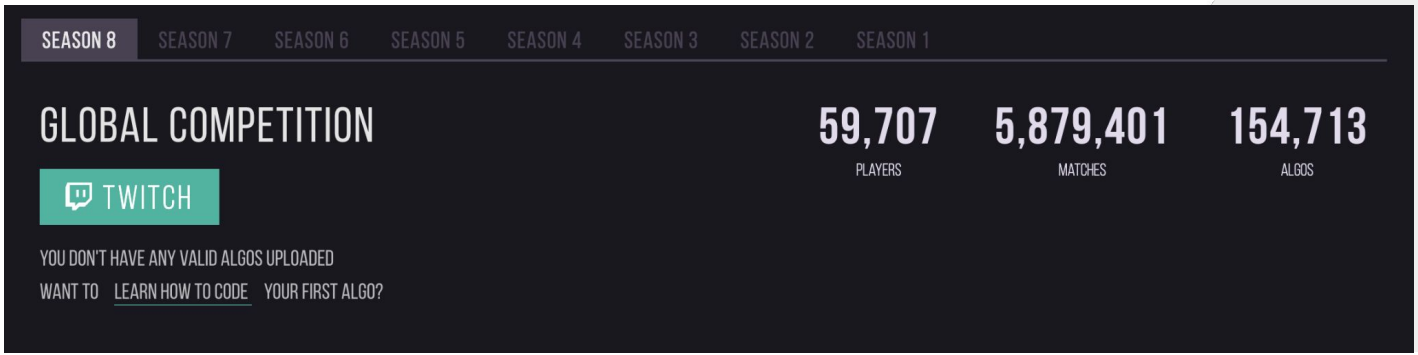


Problem Statement

Why is Terminal Challenging?

100k+ algos in a worldwide competition.

No previous work using RL or ML has been done!



Research Question

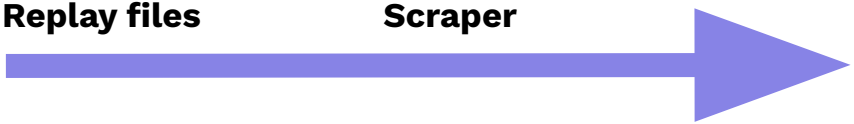
How can we applying Reinforcement Learning / Machine Learning Techniques toward Terminal in a way that has never been done before by the global community?

Data Collection Phase

Parsing

Public
Replay files

Selenium Web
Scraper



Data Pipeline

Data Collection Phase

Parsing

Public
Replay files

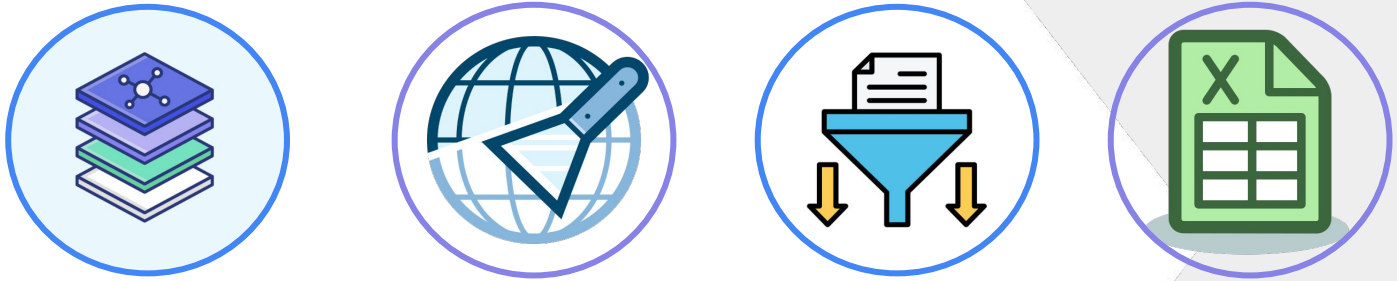
Selenium Web
Scraper

Extract
desired state

Data Pipeline

Data Collection Phase

Parsing



Public
Replay files

Selenium Web
Scraper

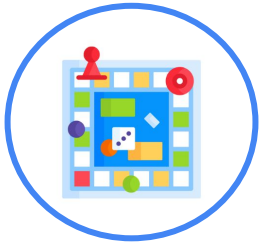
Extract
desired state

Export
to CSV

Data Pipeline

Data Collection Phase

Data Preparation

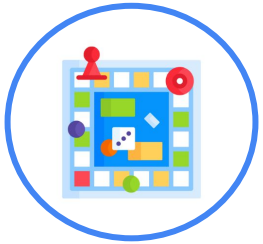


Prepare Gameboard State

Data Pipeline

Data Collection Phase

Data Preparation



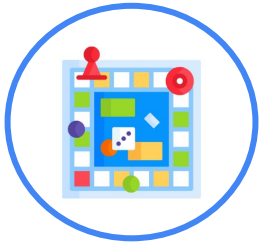
Prepare Gameboard State

Correlate Action Response

Data Pipeline

Data Collection Phase

Data Preparation



Prepare Gameboard State

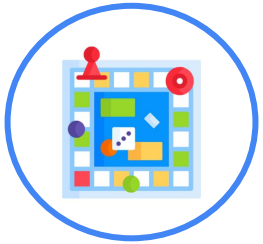
Correlate Action Response

Save as tensors By turn and trajectories

Data Pipeline

Data Collection Phase

Data Preparation



Prepare Gameboard State

Correlate Action Response

Save as tensors By turn and trajectories

Save as .pkl file

Ready for ML!

Data Pipeline

Our Approach - Behavioral Cloning

Model Architectures: CNN & ResNet

Model Inputs

Game Board State: 28x28x6 grid

- *(attack, current health, max health, range, type, owner)*

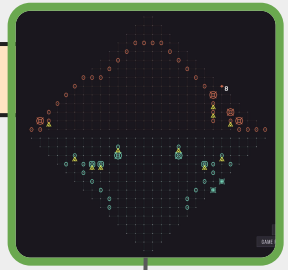
Player Stats:

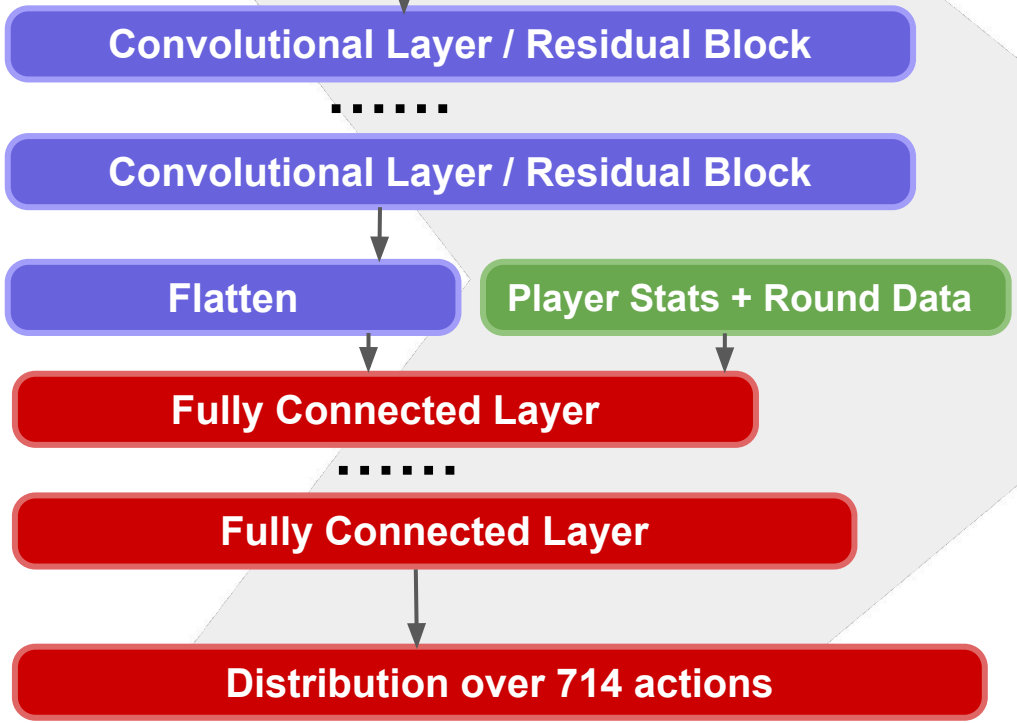
- *(health, previous health, mobile pnts, structure pnts)*

Round #

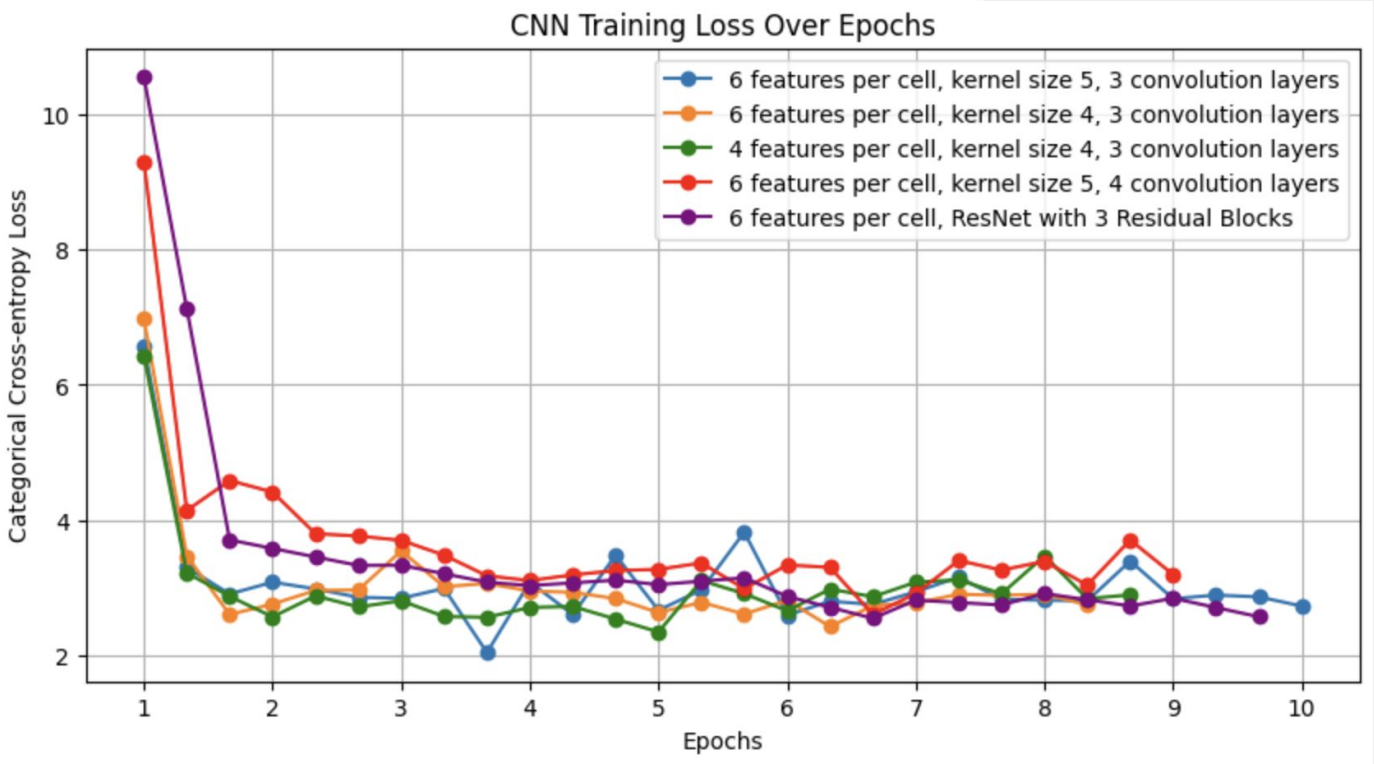
Model Output

The probability it should play the *714* possible actions with





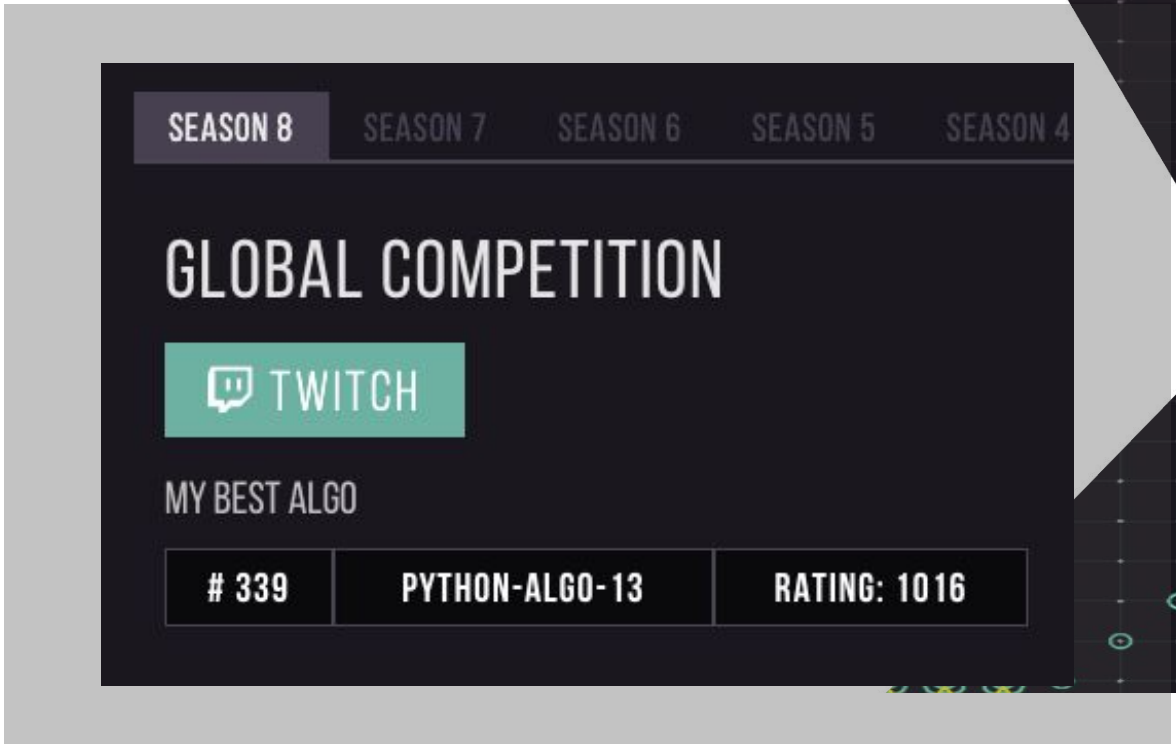
CNN Results

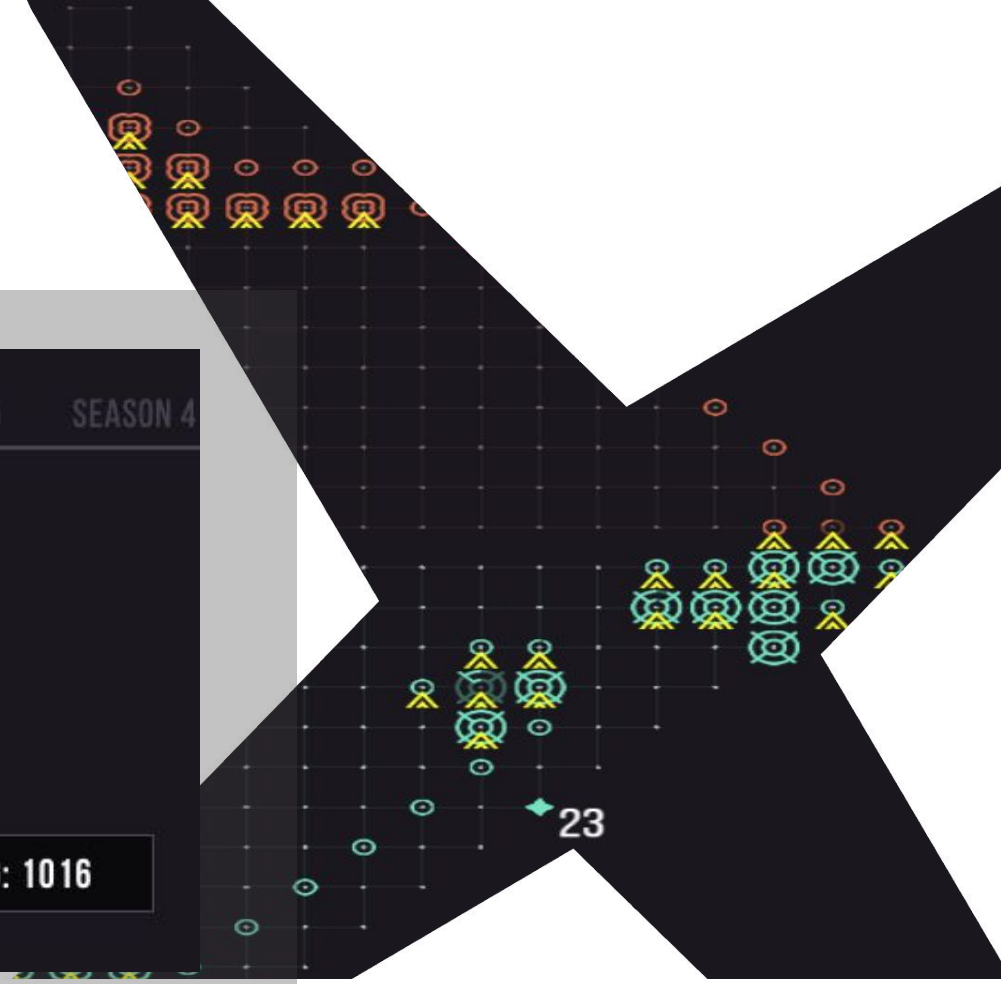


Roughly similar performance across models

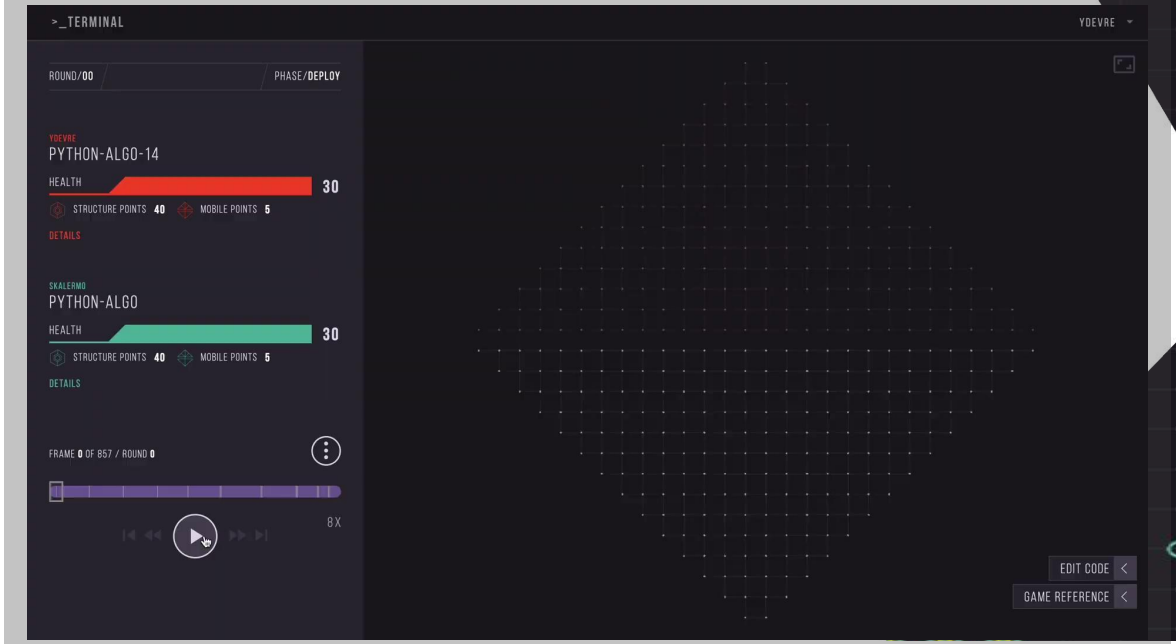
In practice: they exhibit different learned strategies

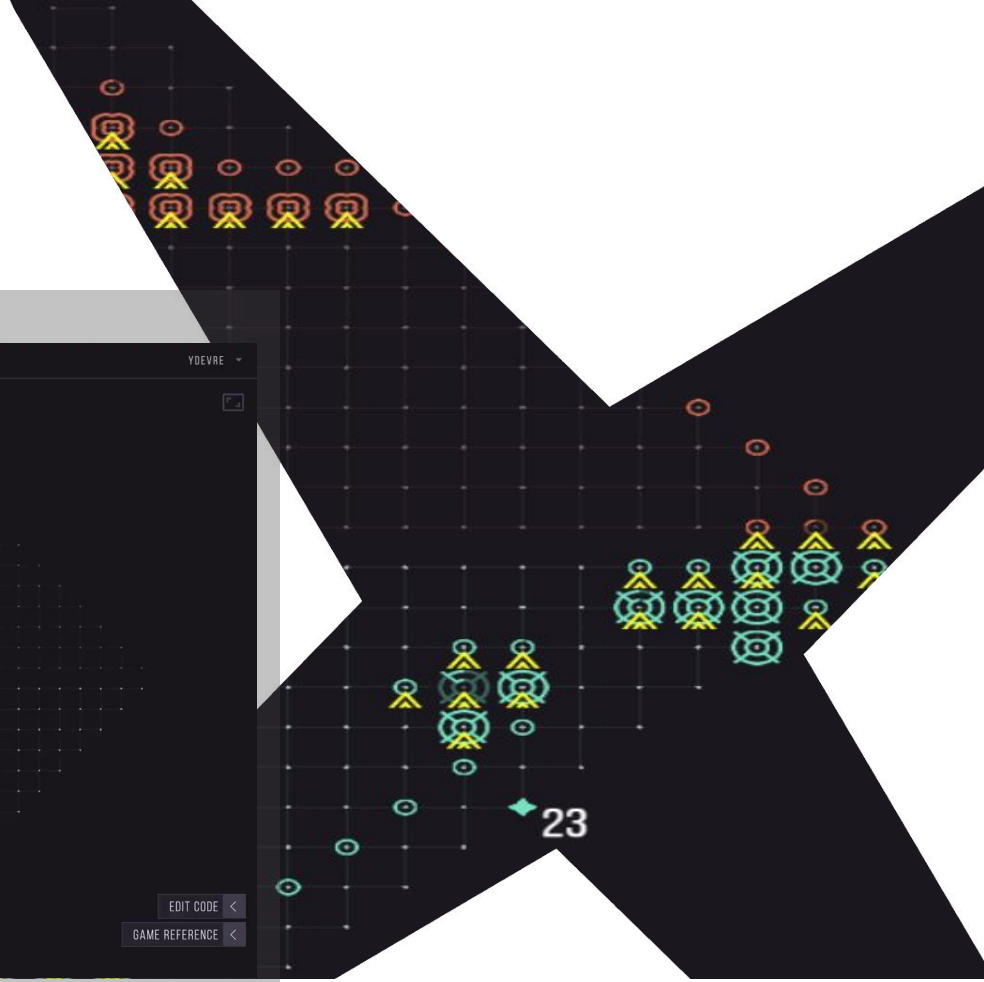
Live Demo!


SEASON 8
SEASON 7
SEASON 6
SEASON 5
SEASON 4
GLOBAL COMPETITION
TWITCH
MY BEST ALGO
339
PYTHON-ALGO-13
RATING: 1016


23

Live Demo!


>_TERMINAL
YDEVRE
ROUND/00
PHASE/DEPLOY
YDEVRE
PYTHON-ALGO-14
HEALTH
30
STRUCTURE POINTS 40
MOBILE POINTS 5
DETAILS
SKALERMO
PYTHON-ALGO
HEALTH
30
STRUCTURE POINTS 40
MOBILE POINTS 5
DETAILS
FRAME 0 OF 857 / ROUND 0
8X
EDIT CODE
GAME REFERENCE


23

RL is TERMINAL

5/1/2024

COS 435: Reinforcement Learning

Professor Mengdi Wang, Professor Benjamin Eysenbach

Authors: Yagiz Devre, Ishaan Javali, Shivam Kak