

RL is **TERMINAL**

5/1/2024

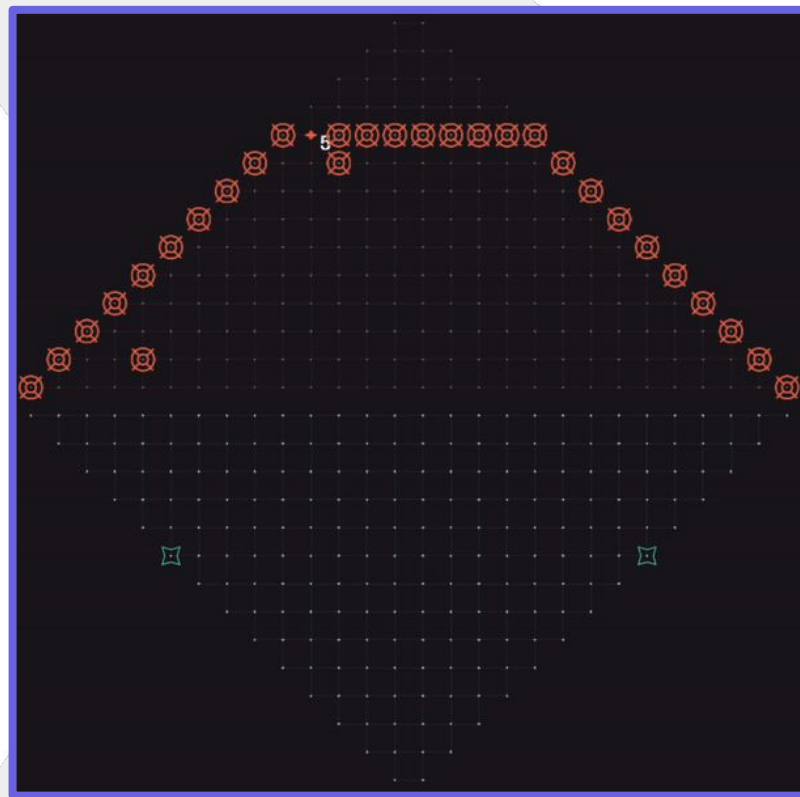
COS 435: Reinforcement Learning

Professor Mengdi Wang, Professor Benjamin Eysenbach

Authors: Yagiz Devre, Ishaan Javali, Shivam Kak

Terminal is a type of Tower Defense game

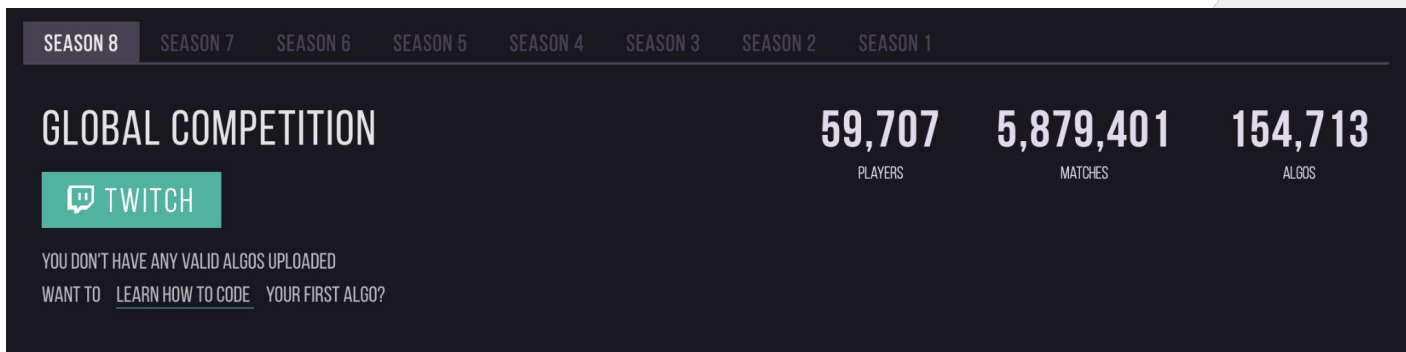
- **Environment:** Diamond-shaped board
- **Actions:** 3 Attack and 3 Defense units:
- **Budget:** Each player has a fixed budget
- **Objective:** Breach the opponent's edges and reduce the opponent's Health Points to zero.



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**

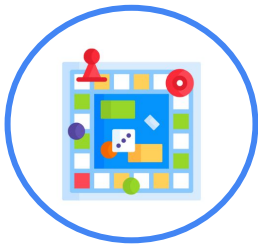


**Export
to CSV**

Data Pipeline

Data Collection Phase

Data Preparation



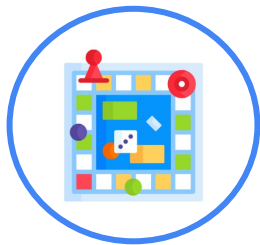
Prepare Gameboard State



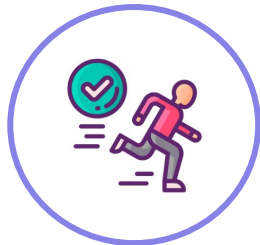
Data Pipeline

Data Collection Phase

Data Preparation



**Prepare
Gameboard
State**

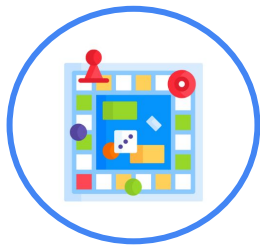


**Correlate
Action
Response**

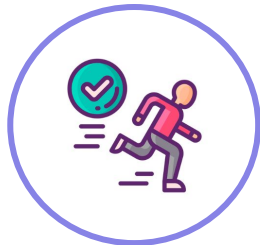
Data Pipeline

Data Collection Phase

Data Preparation



**Prepare
Gameboard
State**



**Correlate
Action
Response**

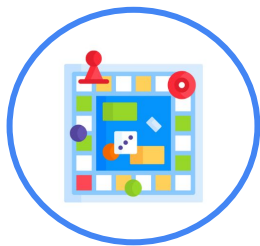


**Save as tensors
By turn and
trajectories**

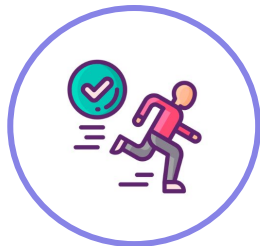
Data Pipeline

Data Collection Phase

Data Preparation



**Prepare
Gameboard
State**



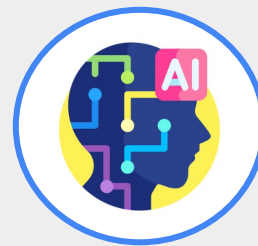
**Correlate
Action
Response**



**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)*

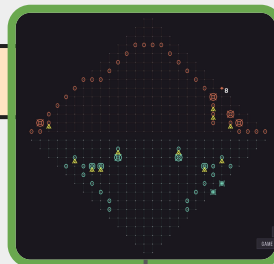
Player Stats:

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

Round #

Model Output

The probability it should play the 714 possible actions with



Convolutional Layer / Residual Block

.....

Convolutional Layer / Residual Block

Flatten

Player Stats + Round Data

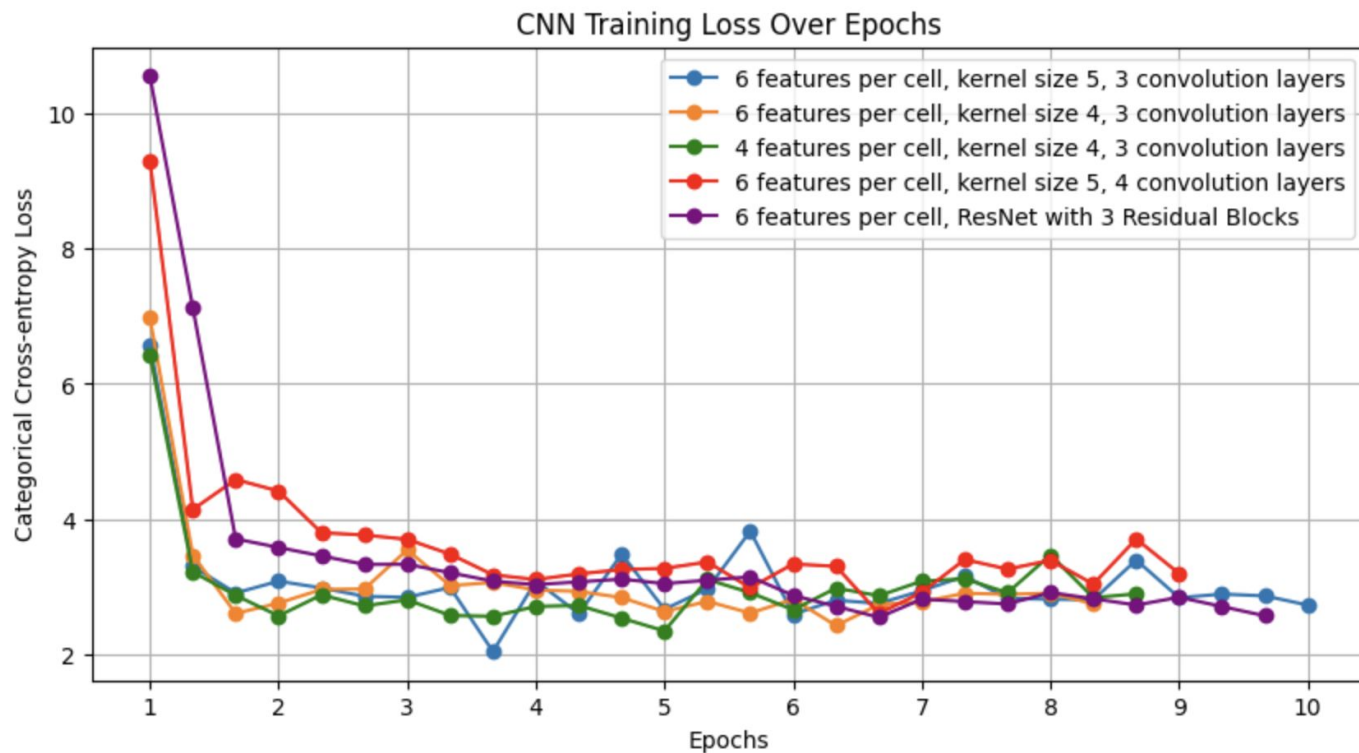
Fully Connected Layer

.....

Fully Connected Layer

Distribution over 714 actions

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

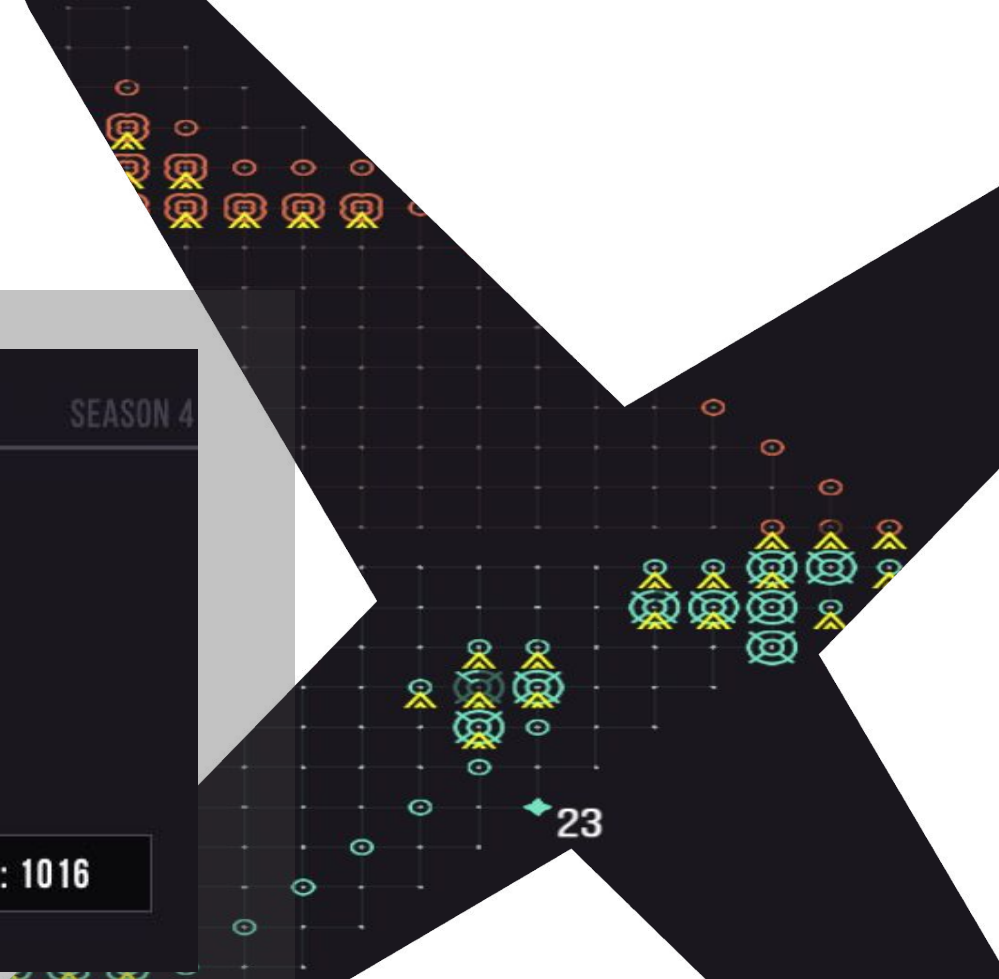
GLOBAL COMPETITION

 TWITCH

MY BEST ALGO

# 339	PYTHON-ALGO-13	RATING: 1016
-------	----------------	--------------

The screenshot shows a dark-themed user interface for a competition. At the top, there are tabs for seasons 4 through 8, with 'SEASON 8' selected. Below the tabs is the title 'GLOBAL COMPETITION'. A teal button with the Twitch logo and text is present. Underneath is the text 'MY BEST ALGO'. At the bottom, a table displays user information: rank '# 339', username 'PYTHON-ALGO-13', and 'RATING: 1016'.



Live Demo!

>_TERMINAL

ROUND/00 PHASE/DEPLOY

YDEVRE
PYTHON-ALGO-14

HEALTH  30

STRUCTURE POINTS 40 MOBILE POINTS 5

DETAILS

SKALERM0
PYTHON-ALGO

HEALTH  30

STRUCTURE POINTS 40 MOBILE POINTS 5

DETAILS

FRAME 0 OF 867 / ROUND 0

 8X

YDEVRE

EDIT CODE <

GAME REFERENCE <



RL is **TERMINAL**

5/1/2024

COS 435: Reinforcement Learning

Professor Mengdi Wang, Professor Benjamin Eysenbach

Authors: Yagiz Devre, Ishaan Javali, Shivam Kak