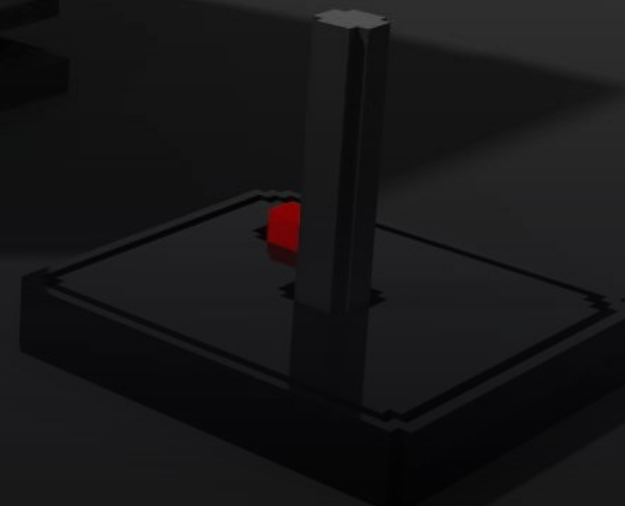
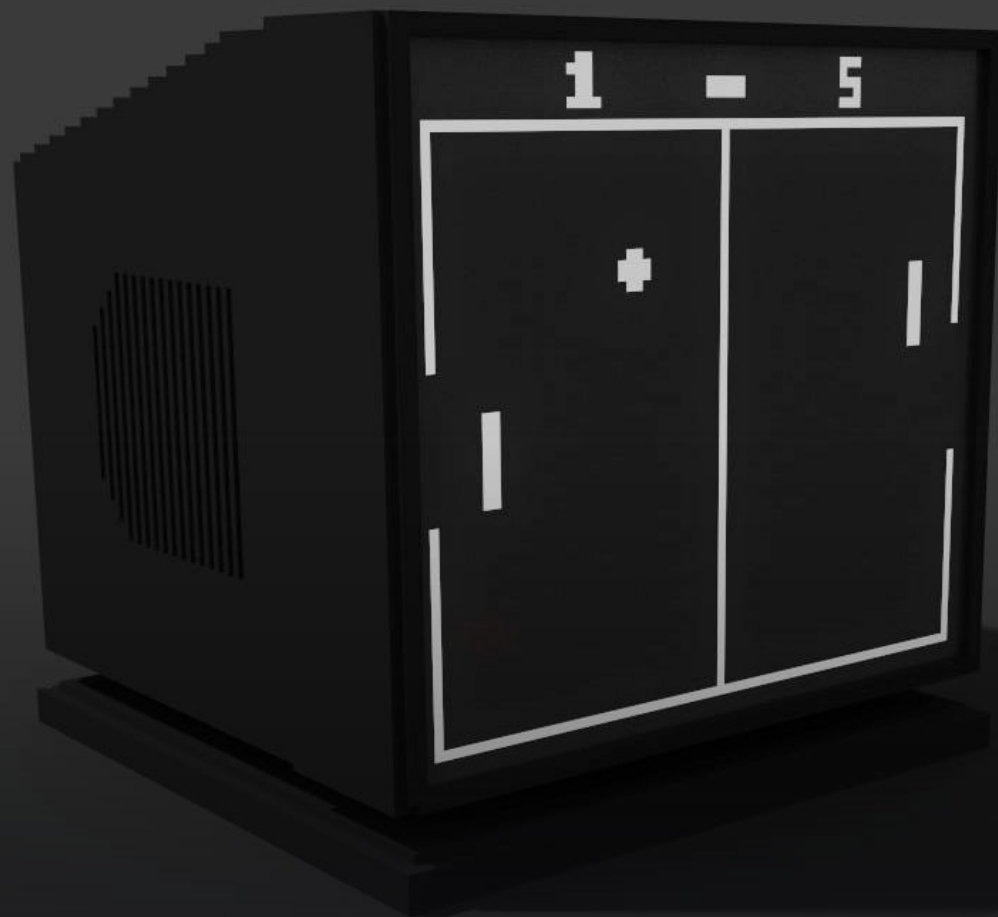


PongAI

FIA 2024/2025

EMANUELE LUIGI AMORE,
05121/16309





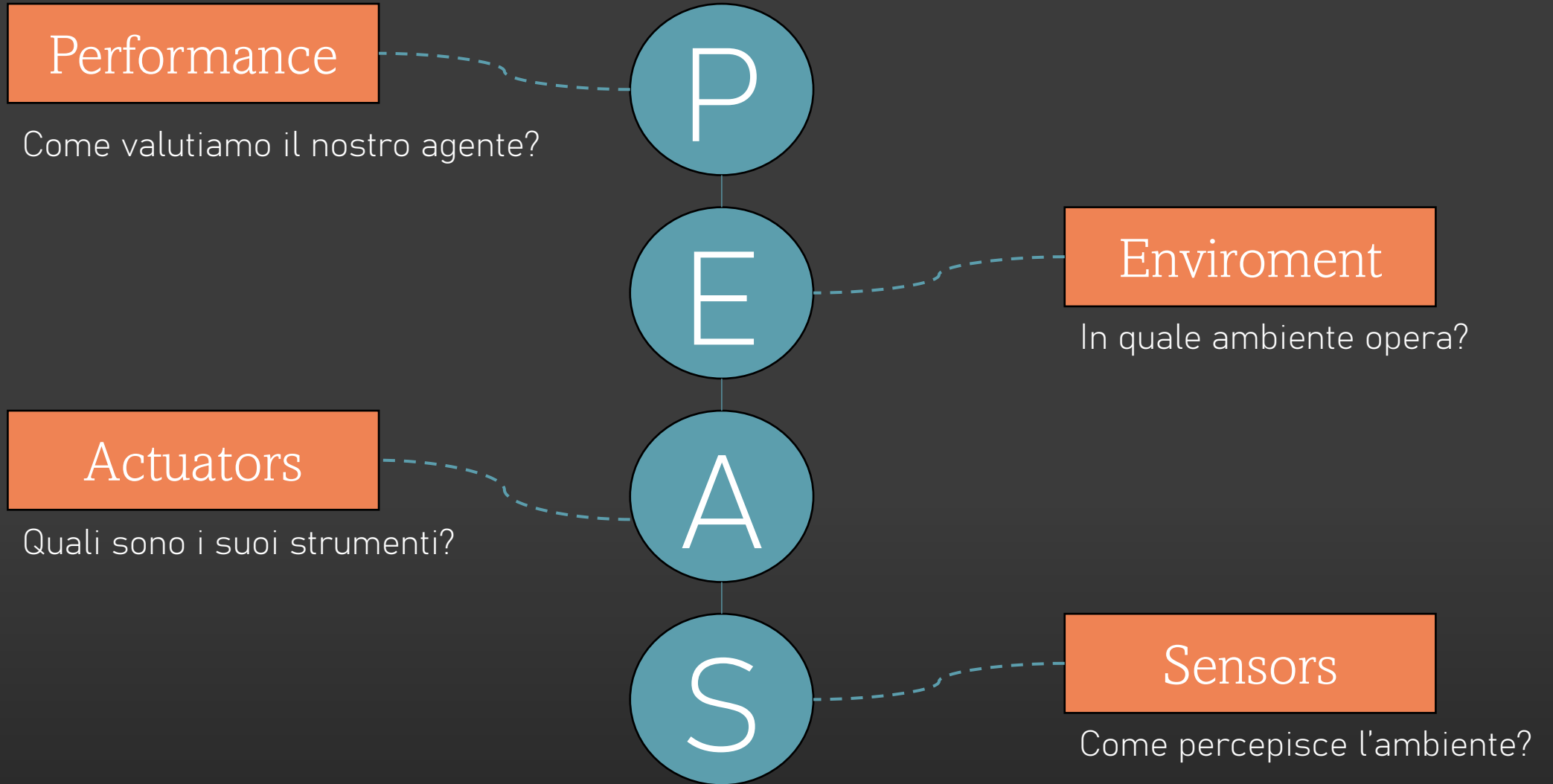
Che cos'è Pong?

Obiettivo del progetto



Creare una IA che permetta a due giocatori di giocare da soli.

Specifica PEAS



Specifica PEAS

Performance

Come valutiamo il nostro agente?



Numero di tocchi



Reward ottenuto



Percentuale vittorie

Specifica PEAS



Multi-Agente e Competitivo



Completamente osservabile



Dinamico



Discreto



Sequenziale



Deterministico



Enviroment

In quale ambiente opera?

Specifica PEAS



Movimento in su



Restare fermo



Movimento in giù

Actuators

Quali sono i suoi strumenti?



Specifica PEAS



Posizione della palla



Posizione dei paddle



Velocità della palla



Sensors

Come percepisce l'ambiente?

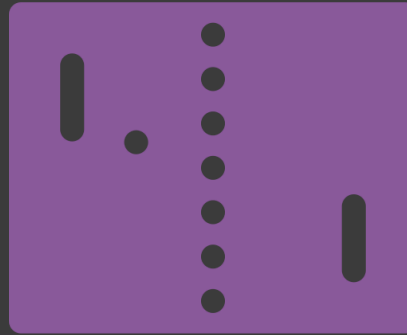
Definizione del Problema

1, Agente



Esplora l'ambiente e
compie delle decisioni

2, Ambiente



Simulazione del gioco

3, Obiettivo



Massimizzare il reward

Q-Learning, che cos'è?

Q-learning è uno dei più conosciuti algoritmi di apprendimento per rinforzo. Fa parte della famiglia di algoritmi adottati nelle tecniche delle differenze temporali, relative ai casi di modelli a inform...

Q-Learning, che cos'è?

Q-learning è uno dei più conosciuti algoritmi di apprendimento per rinforzo. Fa parte della famiglia di algoritmi adottati nelle tecniche delle differenze temporali, relative ai casi di modelli a inform...

Proviamo a riassumerlo in tre punti!



Apprendimento per tentativi ed errori esplorando l'ambiente



Memorizzazione delle esperienze in una Q-Table



Bilanciamento tra esplorazione e sfruttamento

Q-Table, nel dettaglio

Stati della Tabella			
Stato 1	Stato 2	Stato 3	Azioni
0	0	0	Azione 1
0	0	0	Azione 2
0	0	0	Azione 3
0	0	0	Azione 4

Inizialmente abbiamo una tabella inizializzata a 0!

Q-Table, nel dettaglio

Stati della Tabella			
Stato 1	Stato 2	Stato 3	Azioni
+1	0	0	Azione 1
+1	-1	-1	Azione 2
-5	-1	+5	Azione 3
+1	+5	+1	Azione 4

...Durante il training viene poi popolata con valori utili.

La nostra Q-Table

Stati della tabella					
Stato 1	Stato 2	Stato 3	Stato 4	Stato 5	Azioni
0.25	-0.90	0.46	0.25	0.46	Su
0.12	-0.14	0.59	-0.94	0.59	Giù
0.85	-0.90	-0.23	0.60	-0.23	Fermo

I dati, ovviamente, sono di esempio.

Exploration vs Exploitation

Exploration

In questa fase l'agente sceglie un'azione in maniera casuale.

In questo modo, ha la possibilità di esplorare l'ambiente e di apprendere le mosse migliori!



Exploitation

In questa fase, invece, l'agente sfrutta quello che ha appreso.

In questa fase non dovrà più esplorare l'ambiente, ma solo sfruttare le conoscenze senza aggiornare la sua tabella!

Sistema di Rewarding

Tipi di Reward		
-5	<i>DefensivePenalty</i>	<i>Caso in cui si subisce un goal</i>
+1	<i>HitBonus</i>	<i>Caso in cui si tocca la palla</i>
+1	<i>AggressiveBonus</i>	<i>Caso in cui si fa goal</i>
+2	<i>PrecisionBonus</i>	<i>Caso in cui si tocca la palla, ma con gli spigoli</i>

Risultati ottenuti, le metriche



Numero di tocchi



Reward ottenuto

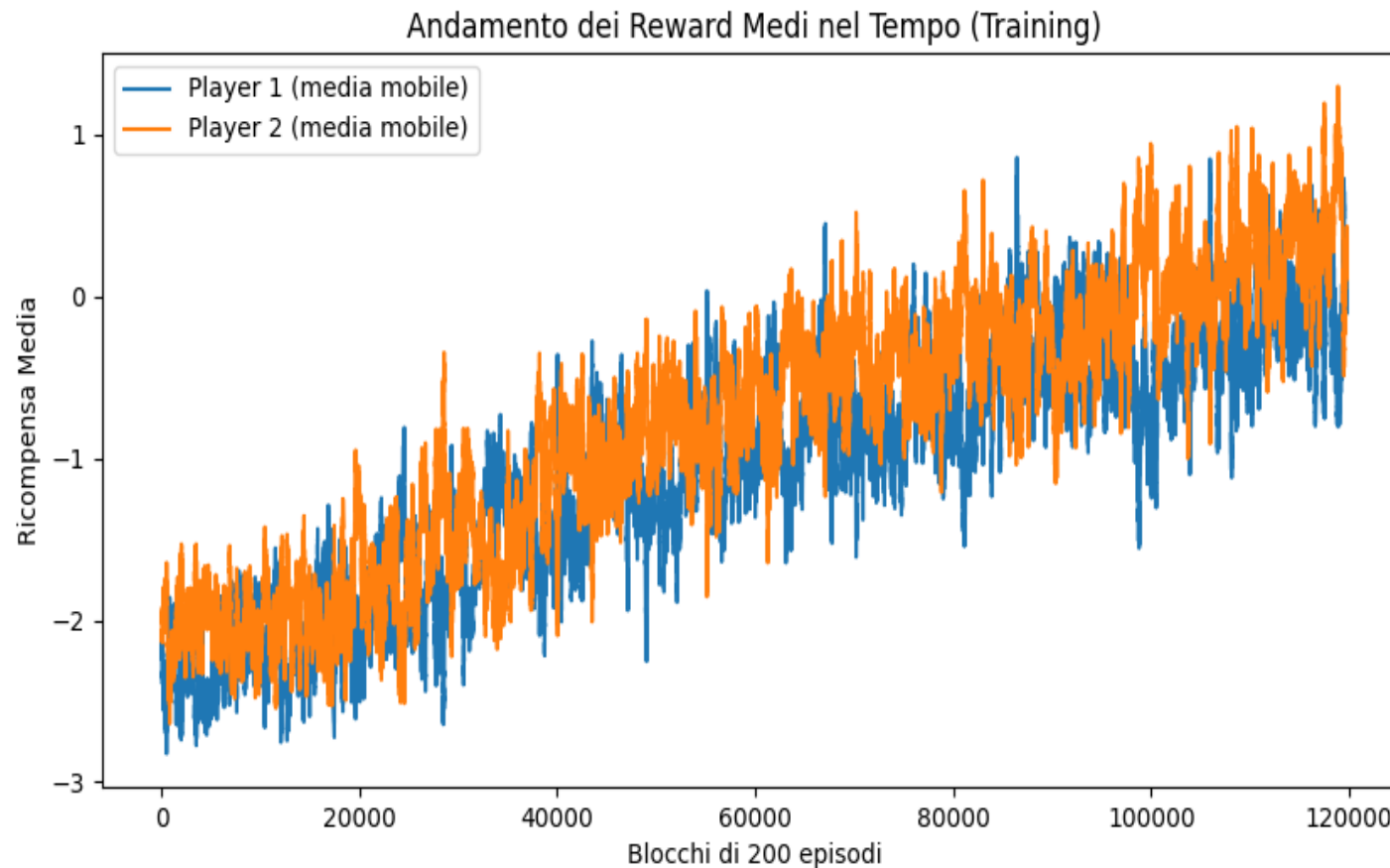


Percentuale vittorie

Fase di Training



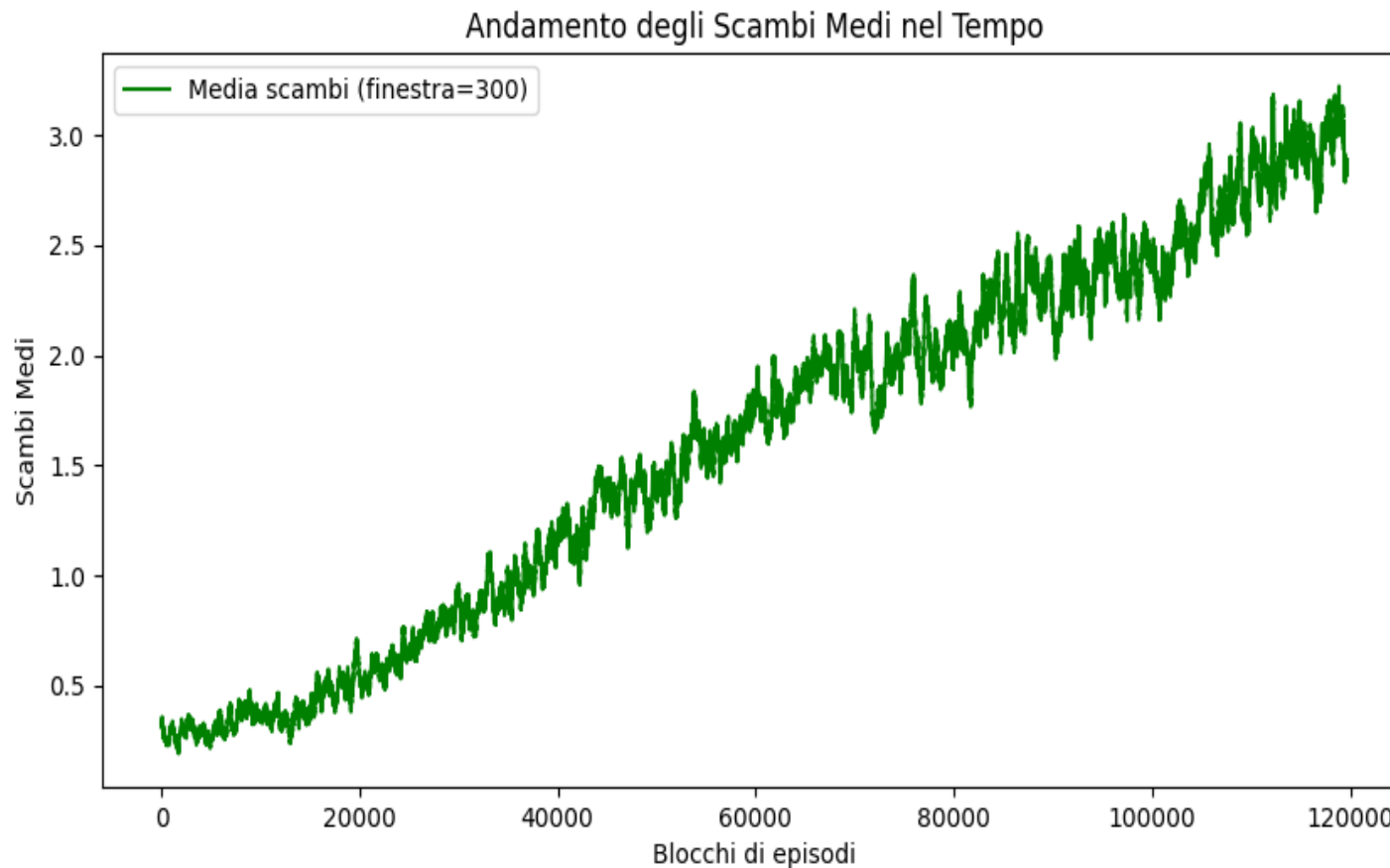
Reward ottenuto



Fase di Training



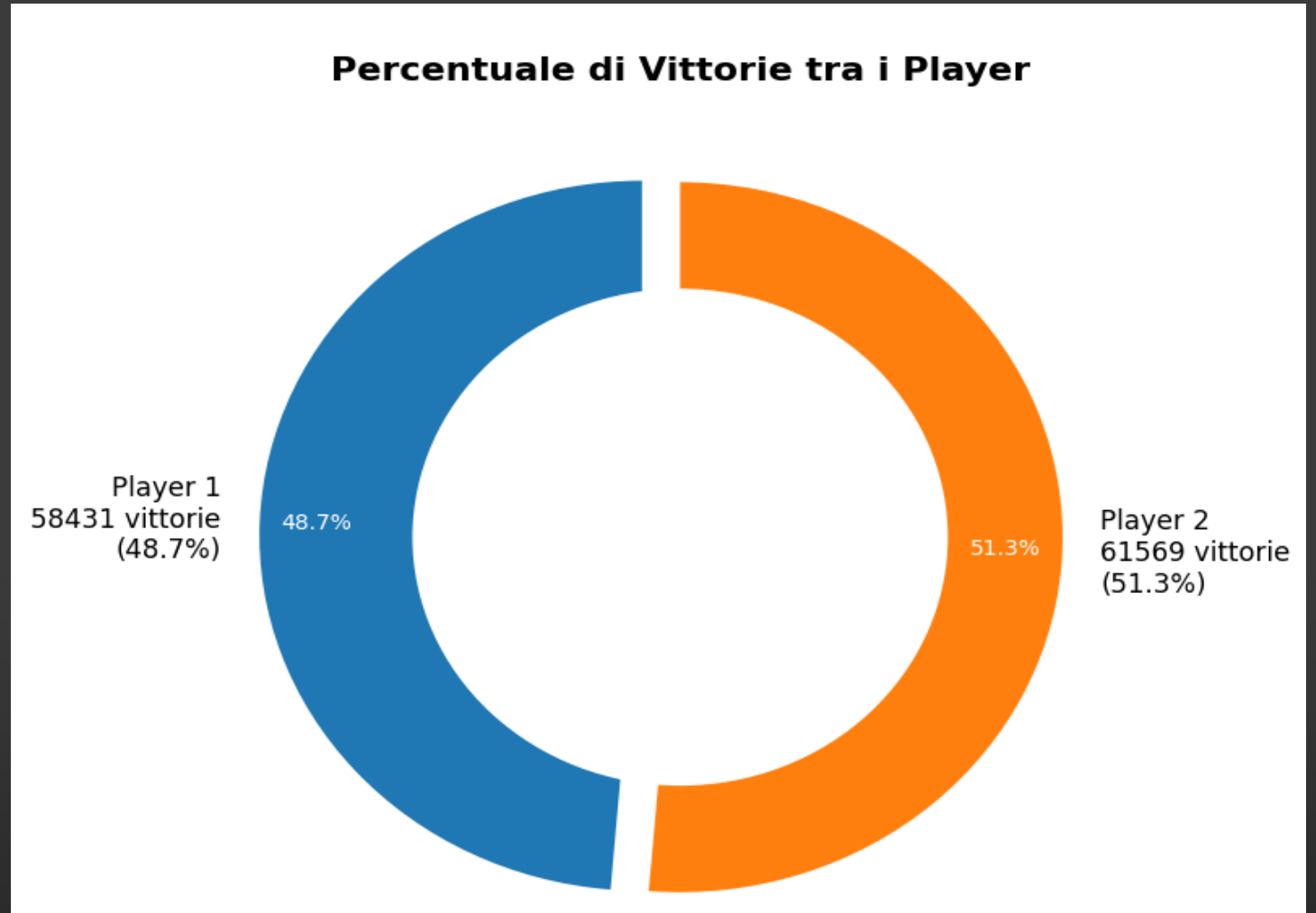
Numero di tocchi



Fase di Training

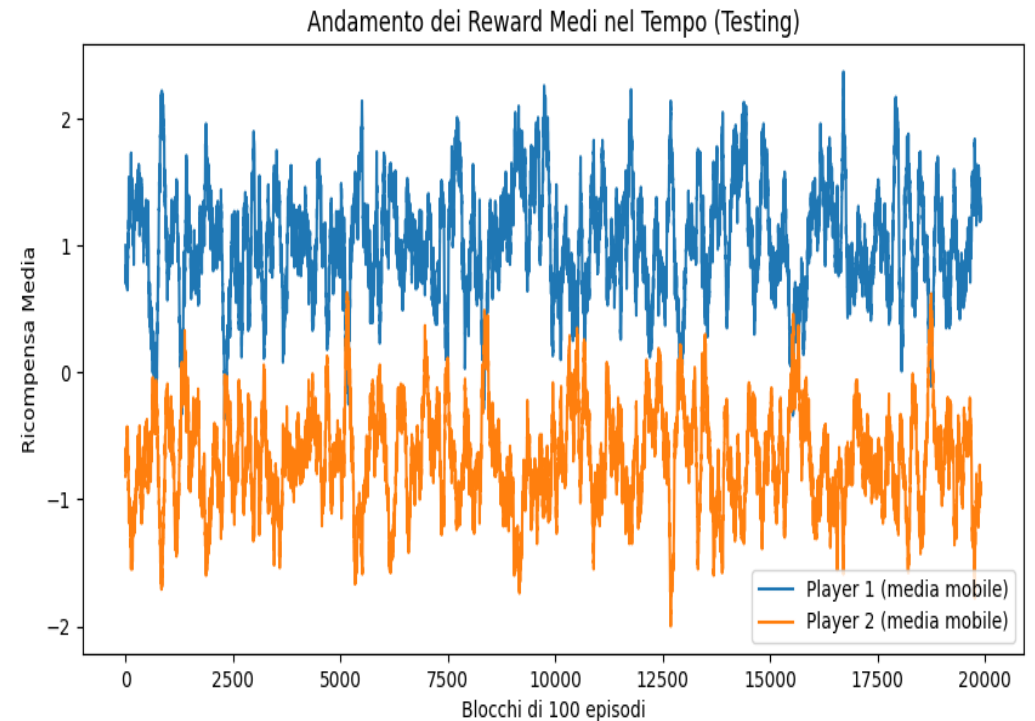
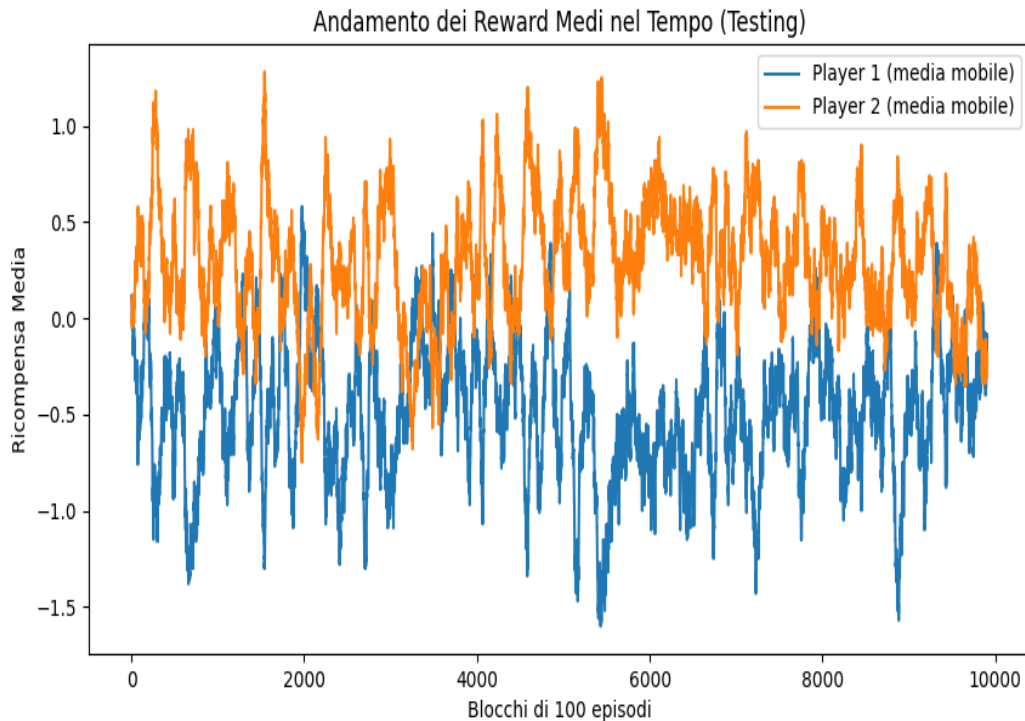


Percentuale vittorie

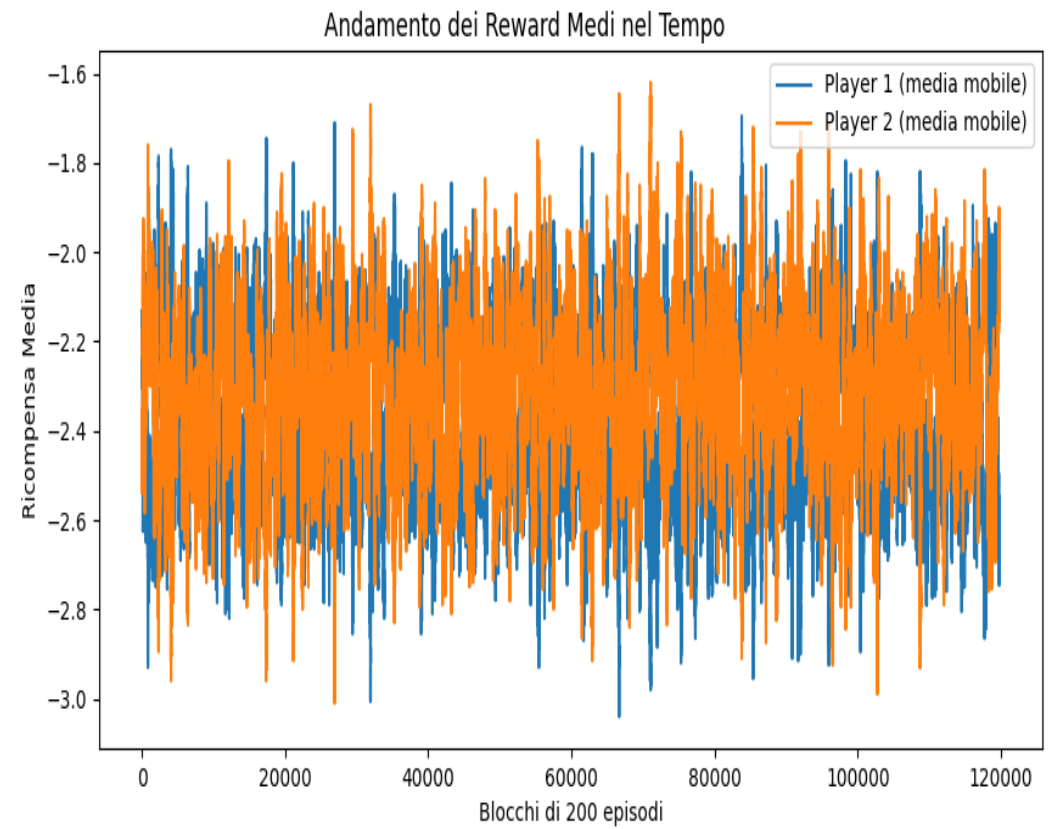
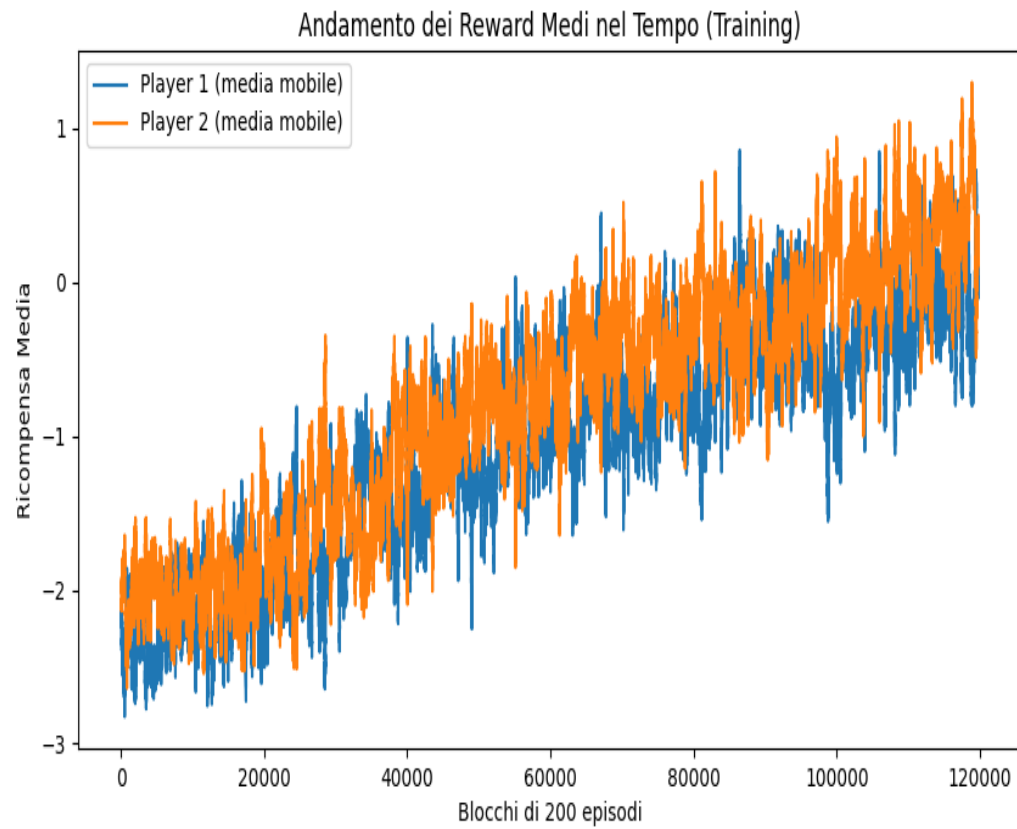


Fase di Testing

In fase di testing possiamo notare come gli agenti ed il modello possono performare in maniera differente!



Trainato vs Non Trainato



Cosa si potrebbe provare per migliorare?



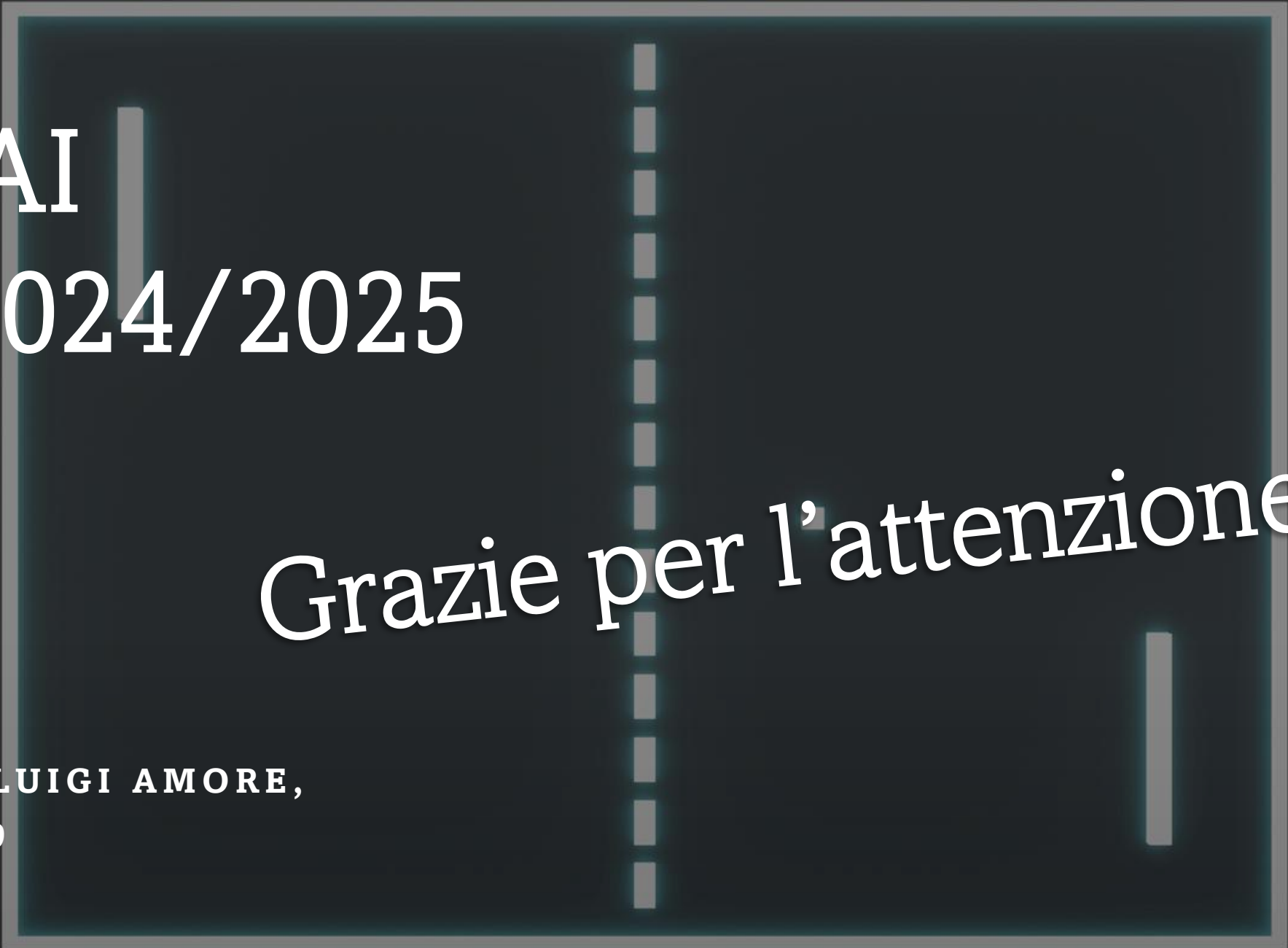
Differenziare il
Training



Double Q-Learning



Rete Neurale

The background of the slide is a dark blue Pong game screen. It features a central vertical dashed line representing the net, and two vertical solid lines on either side representing the paddles. The text is overlaid on this background.

PongAI

FIA 2024/2025

Grazie per l'attenzione!

EMANUELE LUIGI AMORE,
05121/16309