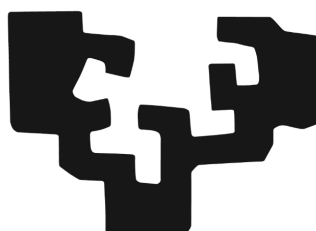


# Rainbow: All together

Informe de trabajo

eman ta zabal zazu



Universidad  
del País Vasco

Euskal Herriko  
Unibertsitatea

Técnicas Avanzadas de la Inteligencia Artificial

26/12/2022

David Bernal Gómez  
Sara Martín Aramburu

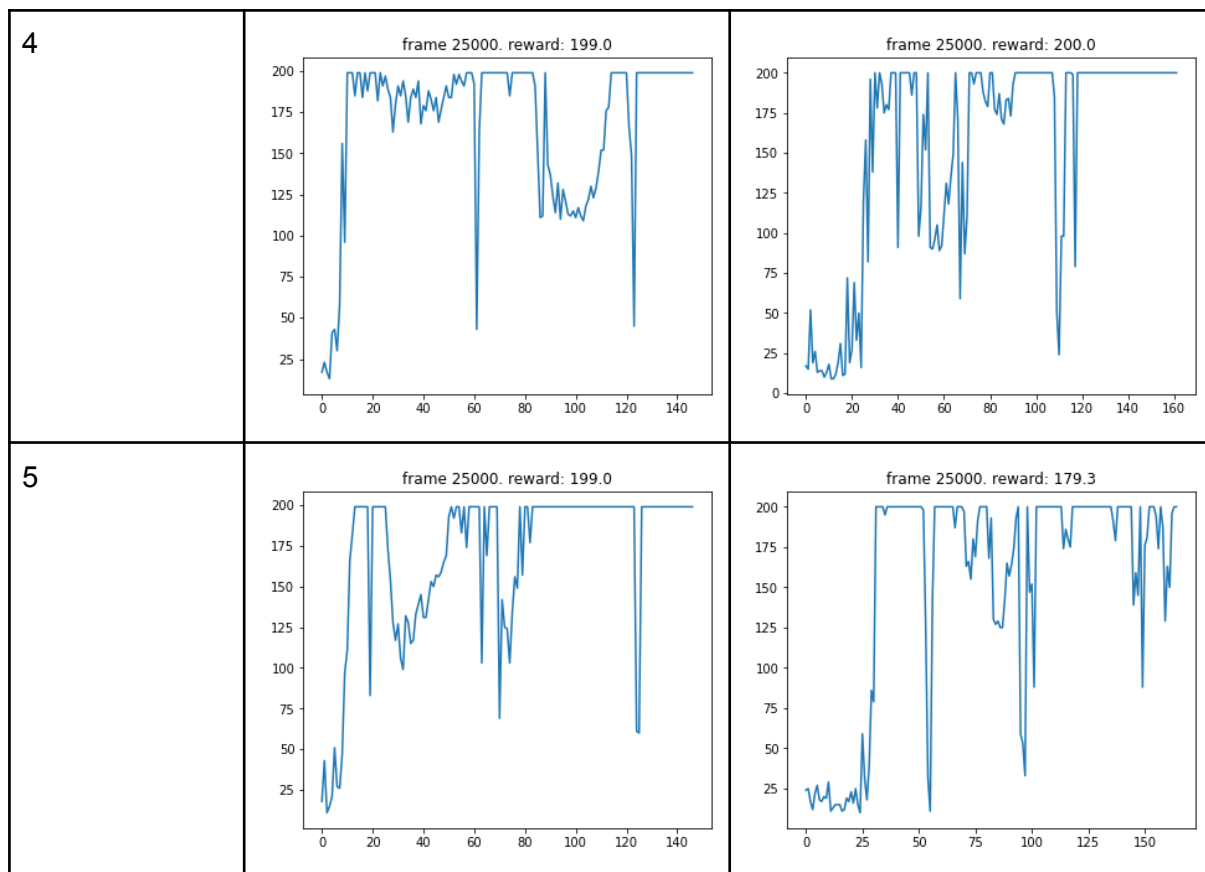
# Índice

<b>Índice</b>	<b>2</b>
<b>Comparación con DQN</b>	<b>3</b>
<b>Comparación con el resto de métodos</b>	<b>5</b>

# Comparación con DQN

[Link al collab](#) en caso de que el código no se pueda ver.

Entrenamiento	All together	DQN
1	<p>frame 25000. reward: 198.2</p>	<p>frame 25000. reward: 192.0</p>
2	<p>frame 25000. reward: 199.0</p>	<p>frame 25000. reward: 117.6</p>
3	<p>frame 25000. reward: 170.2</p>	<p>frame 25000. reward: 200.0</p>



Con estos resultados podemos comparar:

- Las veces que finaliza con la recompensa máxima:
  - All together: 0/5 pruebas
  - DQN: 2/5 pruebas
- El momento en el que la recompensa se dispara:
  - All together:: entre las iteraciones 0 y 2000(10\*200)
  - DQN: entre las iteraciones 0 y 6000 (30\*200)
- Cuánto tiempo se mantienen los modelos con una recompensa máxima:
  - Ambos modelos sufren altibajos en las iteraciones intermedias, pero rainbow (all together) se mantiene más estable.
- Recompensa media a lo largo de los entrenamientos:
  - All together:: media de 169,4 puntos.
  - DQN: media de 140 puntos, ligeramente inferior.

Se puede observar, que aunque no logremos alcanzar la máxima recompensa al final, la recompensa media es superior.

## Comparación con el resto de métodos

A partir de aquí, nos limitaremos a comprar solo las recompensas medias obtenidas, para no llenar todo el documento de gráficas repetidas.

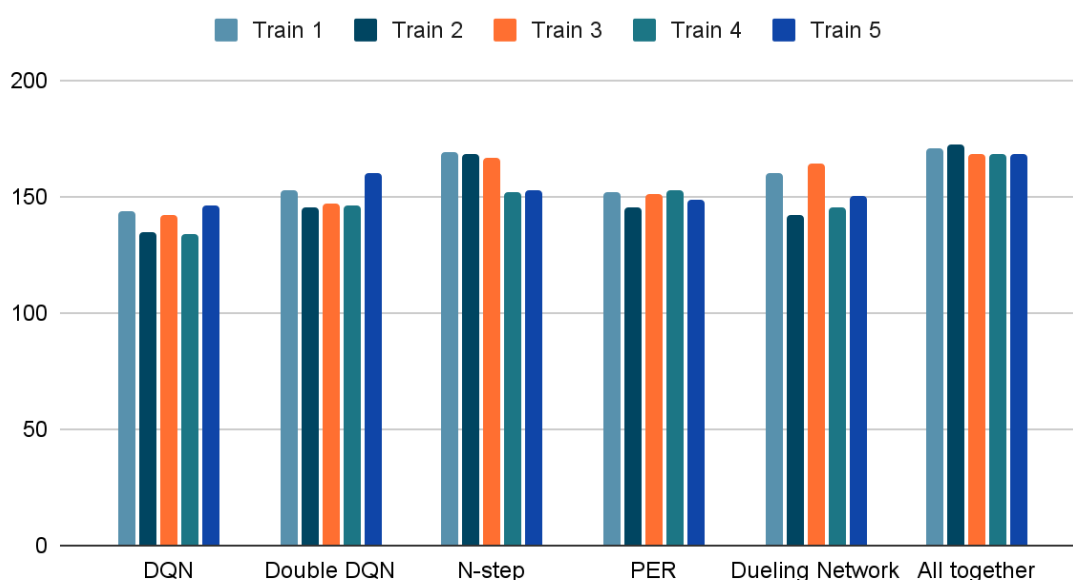
La siguiente gráfica muestra las recompensas medias obtenidas en todos los entrenamientos, para cada uno de los distintos algoritmos. Además incluimos la media general de cada algoritmo

Entrenamiento	All together	DDQN	N-step	PER	Dueling Network
1	171	153	169	152	160
2	172	145	168	145	142
3	168	147	167	151	164
4	168	146	152	153	145
5	168	160	153	149	150
Media total	169,4	150,2	162	150	152

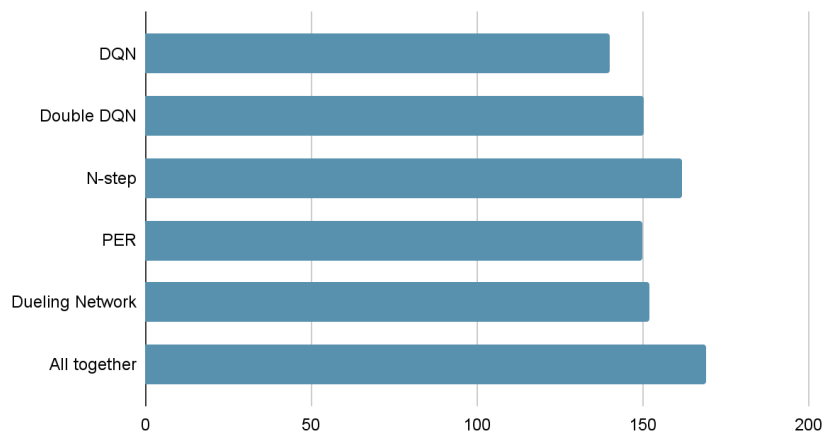
Como podemos observar, usando la técnica all together obtenemos resultados mucho mejores.

Las siguientes gráficas muestran los resultados de forma más visual:

### Recompensas medias en los 5 entrenamientos



Media general de las recompensas obtenidas



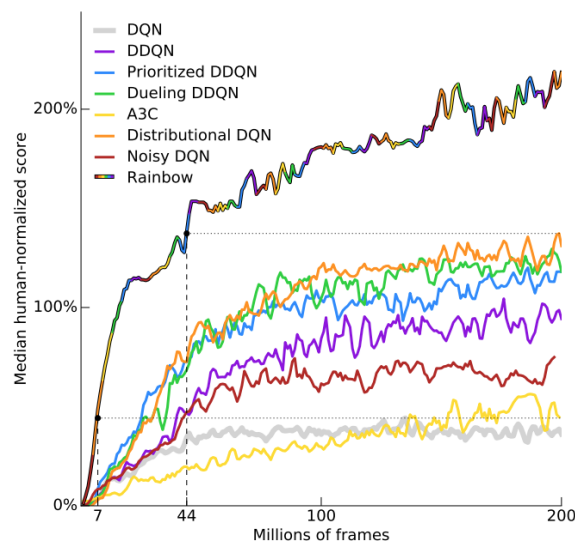
Teniendo en cuenta las recompensas obtenidas en las distintas pruebas, los algoritmos que mayor resultado proporcionan son, en este orden:

1. All together
2. N-step
3. Dueling Network
4. Double DQN y PER
5. DQN

Comparando con los resultados que se obtienen en el paper, los mejores agentes son:

1. All together (rainbow)
2. Dueling Network
3. Prioritized Experience Replay (PER)
4. Double DQN
5. DQN

(N-step no figura en el paper)



Podemos ver como nuestras implementaciones y resultados en las pruebas nos dan un rendimiento similar, siendo Rainbow la mejor opción.