

Comparison of DQN and it's variants On Atari Game

Ishaan Bassi (2016238)

Pooja Gupta (PhD18018)

Arnav Kumar (2016228)



INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI



Github Repository Link : <https://github.com/ishaanbassi/RL-project.git>

Goal of the Project



- To replicate the state of the art paper “Rainbow”^[1] on Atari games, that combines the following extensions of Deep Q Networks -
 - DQN^[3]
 - DDQN^[6]
 - Prioritized Experience Replay^[5]
 - Dueling^[2]
 - Noisy-nets^[4]
 - N-step learning^[8]
 - Distribution Learning^[7]
- To explore further improvements on Rainbow.

Referenced Papers



- [1] Hessel M., Modayil J., Hasselt, H.V., Schaul T., Ostrovski G., Dabney W., Horgan D., Piot B., Azar M., and Silver D. “Rainbow: Combining Improvements in Deep Reinforcement Learning”. 2018. *In Association for the Advancement of Artificial Intelligence(AAAI)*.
- [2] Wang Z., Schaul T., Hessel M., Hasselt H.V., Lanctot M. and Freitas N.D. “Dueling Network Architectures for Deep Reinforcement Learning”. 2016. *arXiv-pre-print-arXiv : 1511.06581*.
- [3] Mnih V., Kavukcuoglu K., Silver D., Rusu A. A., Veness J., Bellemare M. G., Graves A., Riedmiller M., Fidjeland A. K., Ostrovski G., Petersen S., Beattie C., Sadik A., Antonoglou I, King H., Kumaran D., Wierstra D., Legg S., Hassabis D. “Human-level control through deep reinforcement learning”. 2015. *Nature*, vol 518.
- [4] Fortunato M., Azar M.G., Piot B., Menick J., Hessel M., Osband I., Graves A., Mnih V., Munos R., Hassabis D., Pietquin O., Blundell C., and Legg S. “Noisy Networks for exploration”. 2018. *In International Conference on Learning Representations(ICLR 2018)*.

Referenced Papers Continue



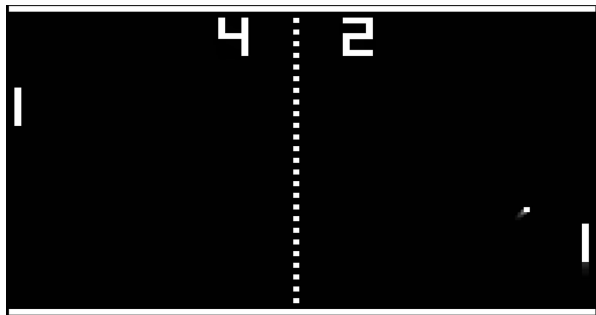
- [5] Schau T., Quan J., Antonoglou I., Silver D. “Prioritized Experience Replay”. 2016, *International Conference on Learning Representations(ICLR 2016)*.
- [6] Hasselt H. V., Guez A., Silver D. “Deep Reinforcement Learning with Double Q-learning”. 2016. *In Association for the Advancement of Artificial Intelligence(AAAI)*.
- [7] Bellemare M. G., Dabney W., Munos R. “A Distributional Perspective on Reinforcement Learning”. *In International Conference on Machine Learning(ICML)*.
- [8] Multi-Step Reinforcement Learning from ‘Reinforcement Learning: An Introduction”. By Sutton R. S. and Barto A.G. Second Edition. Available at :
'<http://incompleteideas.net/book/RLbook2018trimmed.pdf>
- [9] Pedamonti D.“Comparison of non-linear activation functions for deep neural networks on MNIST classification task”. 2018. *arxiv: 1804.02763*.

Novel Contribution by the Team



- Activation functions play a key role in neural networks.
- The improvement proposed by us in this work is that we have replaced the **ReLU** Activation Function with **SELU** (Scaled exponential Linear Unit).
- **Why SELU?**
SELU has the self-normalizing properties which makes the network training *highly robust*, especially in case of many layers architectures^[9].

Pong Atari Game Description



- Pong simulates 2D table tennis. The agent controls an in-game paddle which is used to hit the ball back to the other side.
- **Env** - OpenAI gym, simulated using ALE
- **Agent** - Self player
- **Rewards** - Agent Points at Terminal state
- **States** - 4 frames of 84x84 window (4X84X84)
- **Actions** - 6 actions (0-stay, 1-fire)

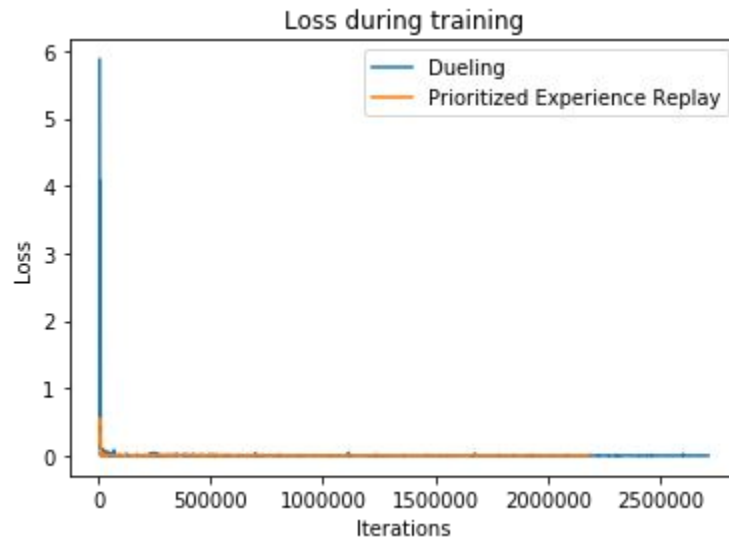
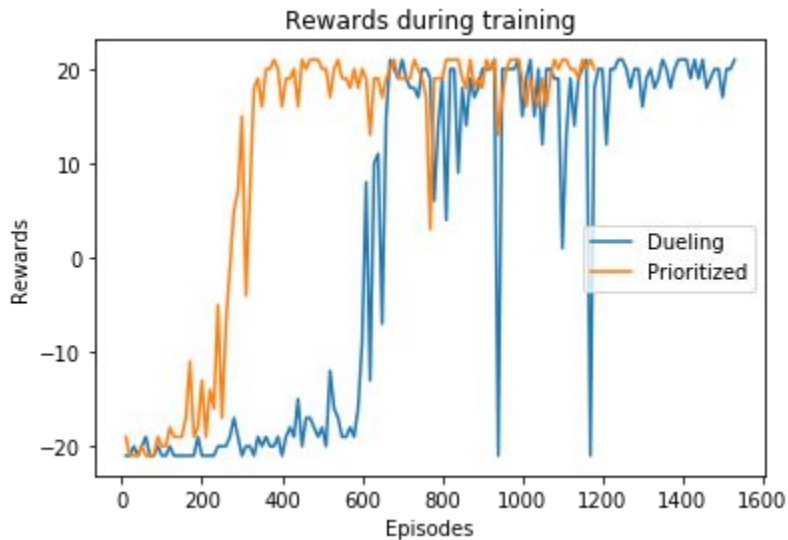
Rainbow combines the following algorithms:

1. DQN - deep Q-learning network
2. DDQN - avoids overestimated values of value functions
3. Prioritized DDQN - uses experience replay buffer with priorities
4. Dueling DDQN - two estimators:
 - a) value-function
 - b) advantage function
5. Multi-Step - Q learning based on returns from multiple steps
6. Distributional DQN - learns the distribution of the value function.
7. Noisy DQN - adds noise to the weights aiding exploration.

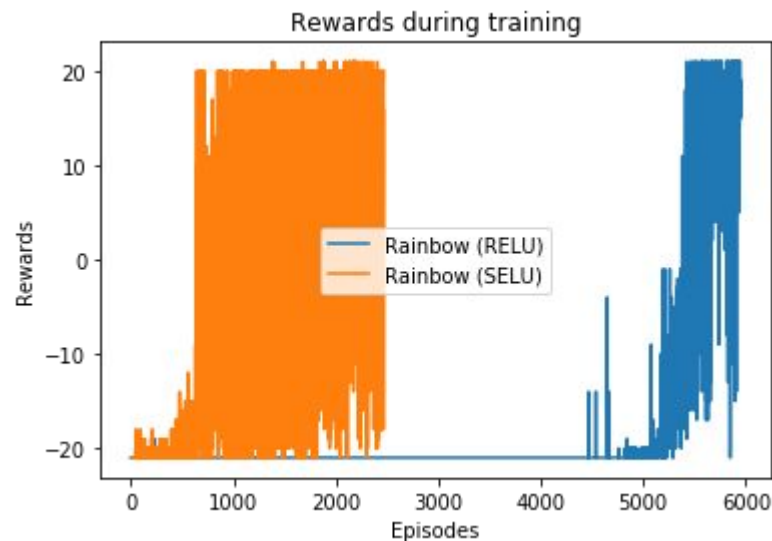
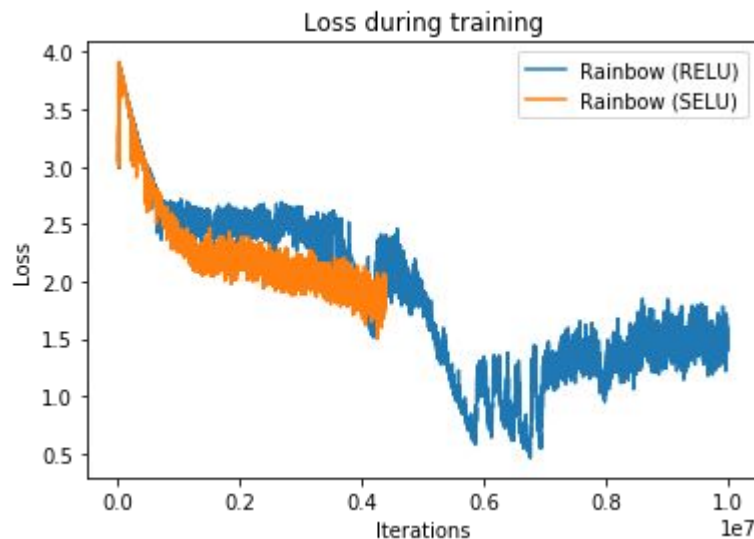
Training Curves



Dueling and Prioritized Experience Replay



RELU vs SELU



Average Episodic Rewards



| S.no | <u>Algorithm</u> | <u>Average Episodic Reward</u> | <u>Standard Deviation</u> |
|------|-------------------------------|--------------------------------|---------------------------|
| 1 | Dueling DQN | 20.4 | .756 |
| 2 | Prioritized Experience Replay | 21 | 0 |
| 3 | Rainbow (RELU) | 15.36 | 6.355 |
| 4 | Rainbow (SELU) | 18.26 | 0.44 |

Average Episodic Rewards



| S.no | <u>Algorithm</u> | <u>Average Episodic Reward</u> | <u>Standard Deviation</u> |
|------|-------------------------------|--------------------------------|---------------------------|
| 1 | Dueling DQN | 20.4 | |
| 2 | Prioritized Experience Replay | 21 | |
| 3 | Rainbow (RELU) | | |
| 4 | Rainbow (SELU) | | |