

Présentation et pratique

Renforcement : Minecraft

...

Ateliers pratiques



Clément Romac



Vincent Béraud

1h min

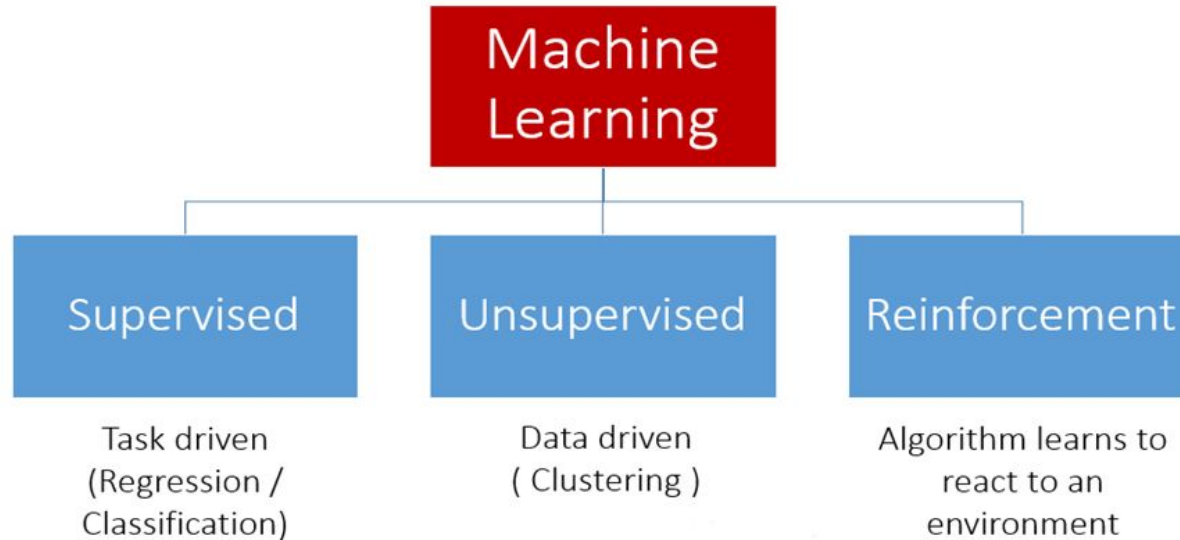


2h ~ K.O. technique

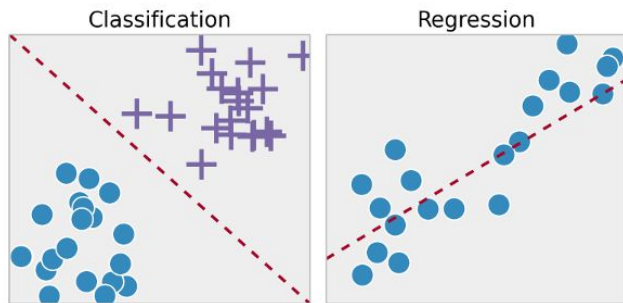
AI AGENT PLAYING MINECRAFT



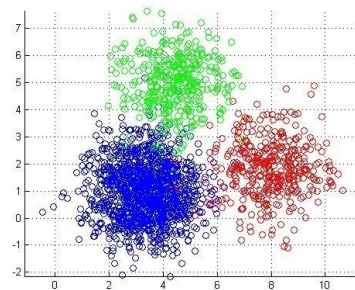
Types of Machine Learning



SUPERVISED

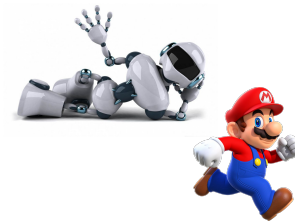
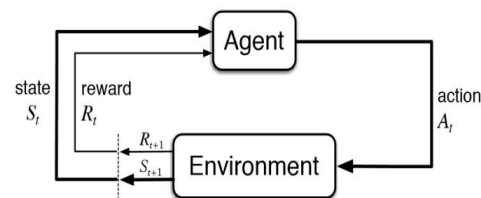


UNSUPERVISED

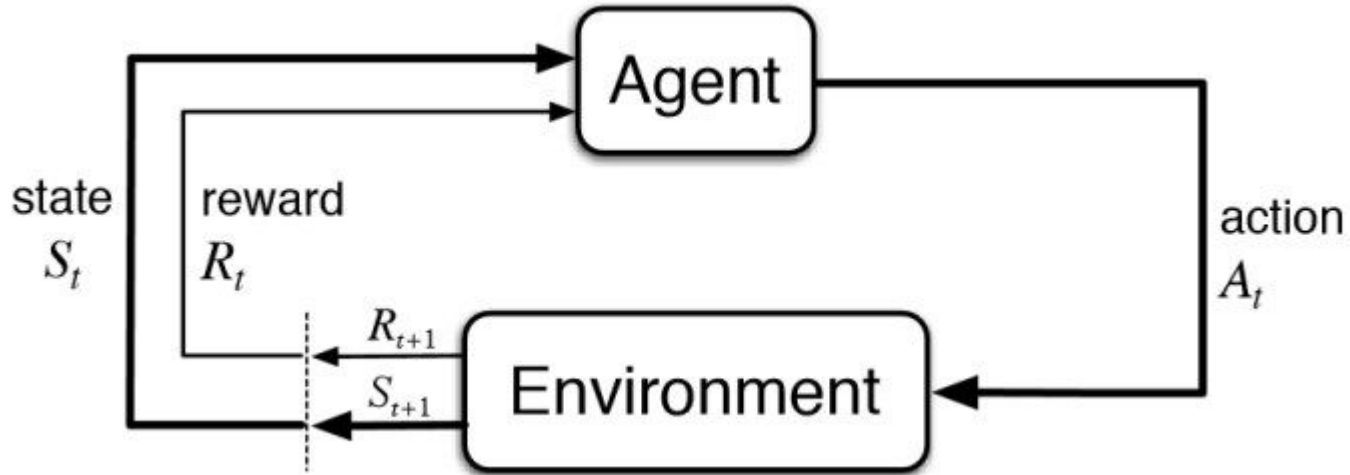


NETFLIX
me^otic

REINFORCEMENT

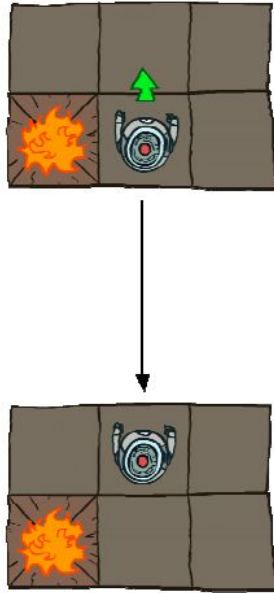


WHAT IS REINFORCEMENT LEARNING ?

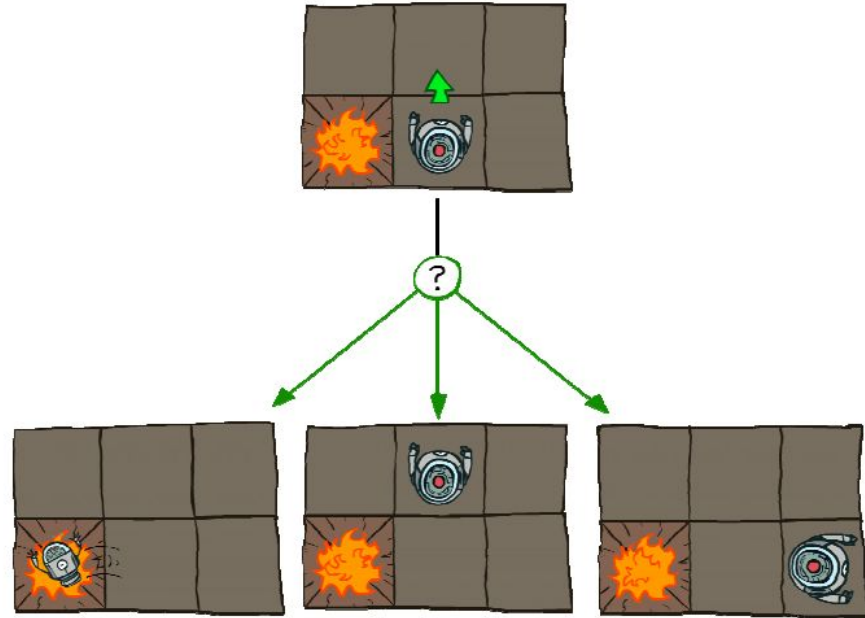


WHY MARKOV DECISION PROCESS ?

Deterministic Grid World



Stochastic Grid World



ACTIONS & REWARDS



STATES VALUES

V=0.81	V=0.9	V=1	
V=0.73		 V=0.9	
V=0.66			

$$V(\pi, s) = R(s) + \gamma \sum_{s' \in S} P(s' | s, \pi(s)) V(\pi, s')$$

BELLMAN EQUATION

$$V^*(s) = R(s) + \max_a \gamma \sum_{s' \in S} P(s' | s, a) V^*(s') \quad \forall s \in S$$

↑
State's reward

↑
Discount factor

↑
Next state's value

ALGORITHMS

VALUE ITERATION

POLICY ITERATION

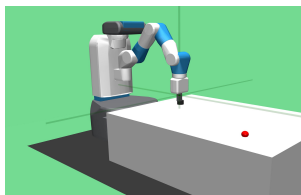
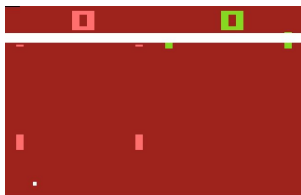
Q-LEARNING

$$Q(s,a) \leftarrow Q(s,a) + \alpha (R(s) + \gamma \max_{a'} Q(s',a') - Q(s,a))$$

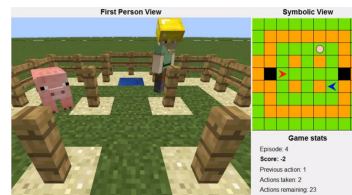
MINECRAFT ENVIRONMENT



Gym OpenAI



Project Malmo



MINECRAFT ENVIRONMENT

GYM MINECRAFT

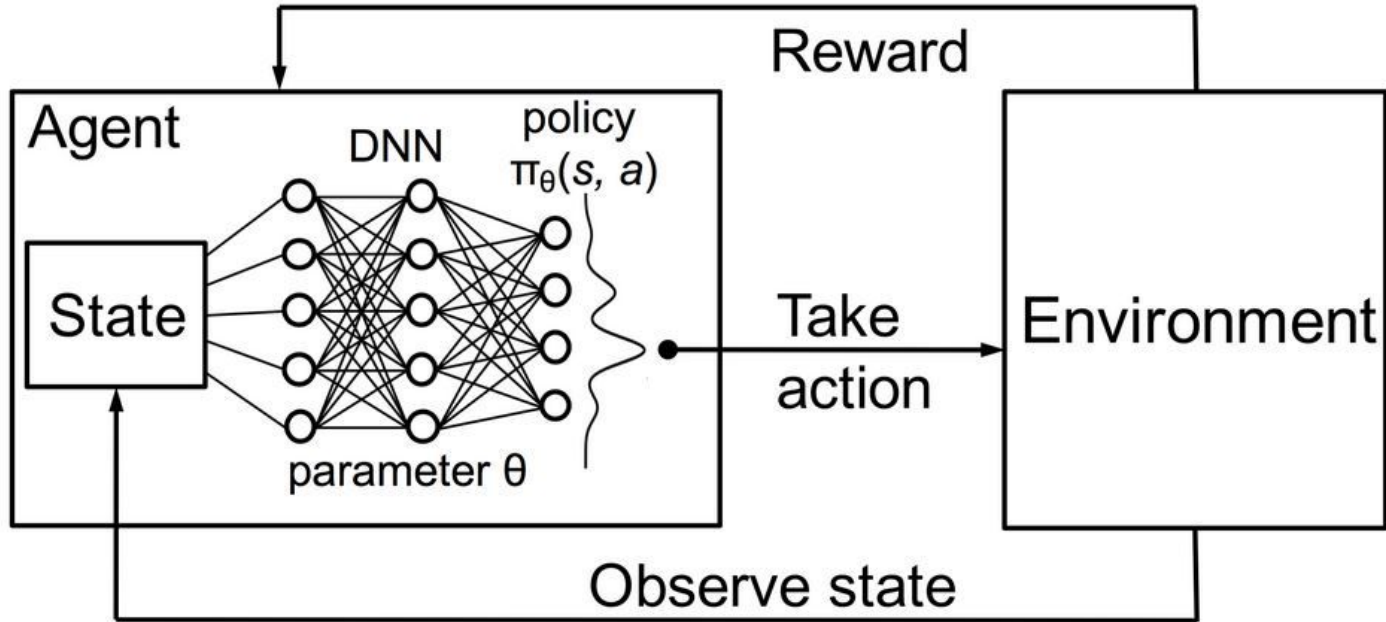
by Tambet Matiisen



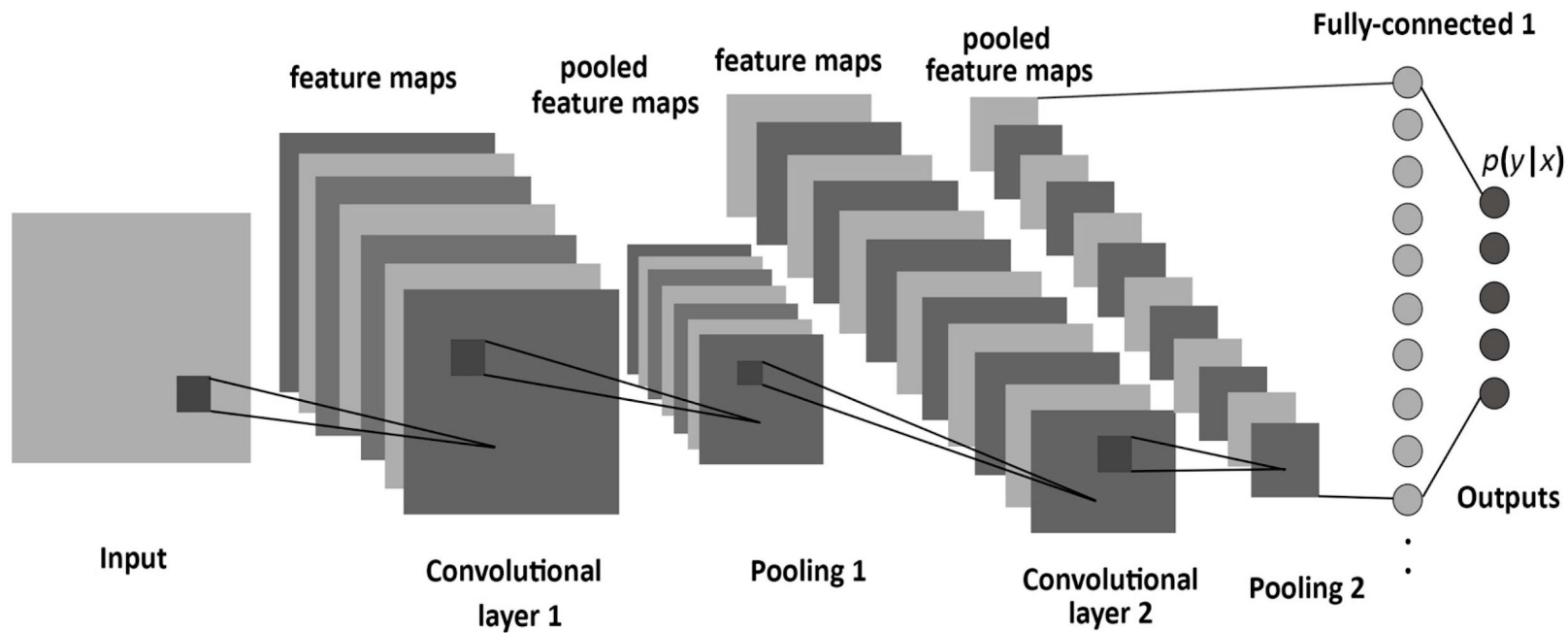
+



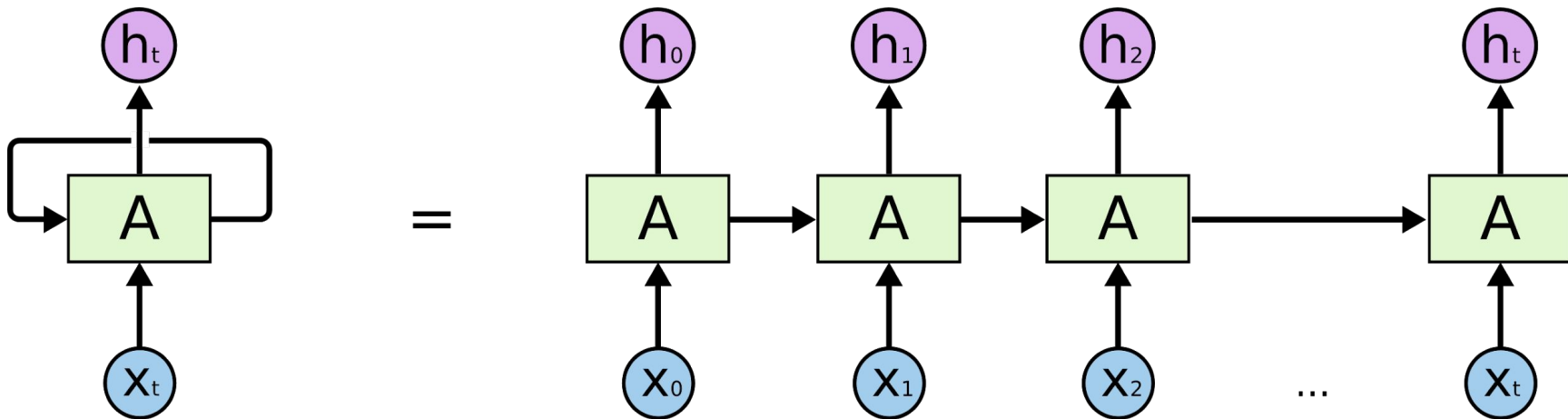
DQN



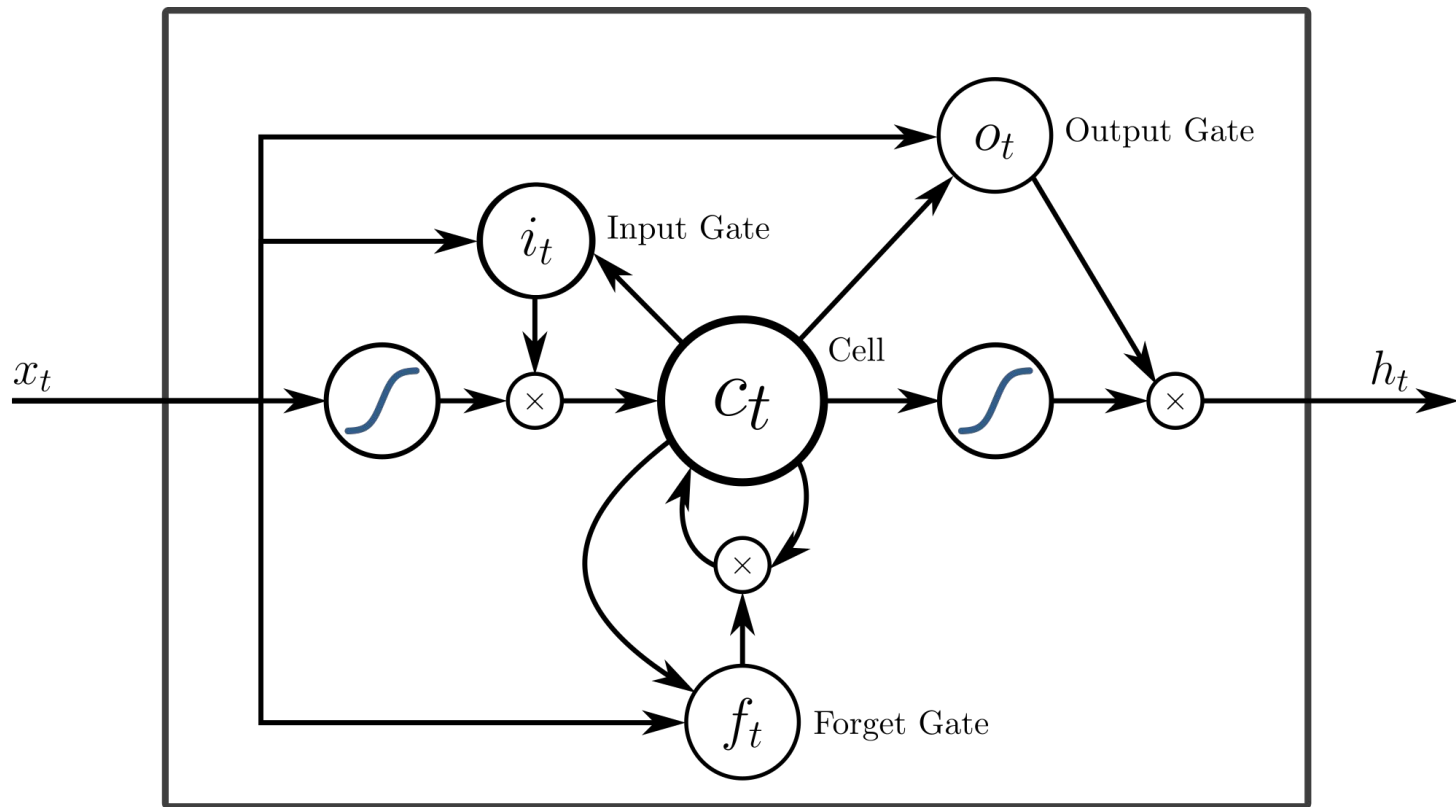
CNN



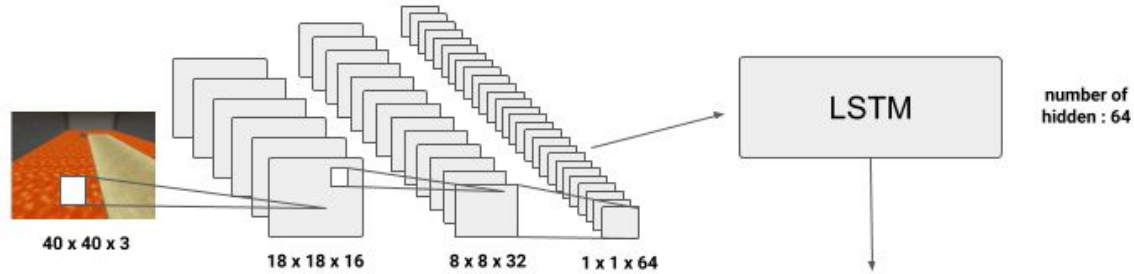
RNN



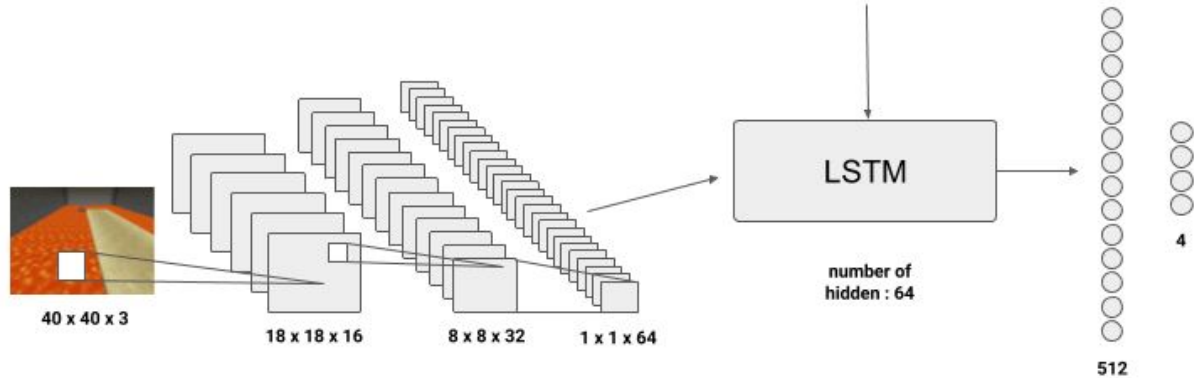
LSTM



ARCHITECTURE



... x8



EXPLORATION / EXPLOITATION

PRE TRAIN STEPS

ϵ GREEDY EXPLORATION

TIPS

EXPERIENCE REPLAY

DOUBLE Q LEARNING

EXPERIMENTS



TRAIN

Victory: 6%



TEST

Victory: 54%

TIME: 2000ms

r_WIN: 100

r_OT: -100

r_step: -3

EXPERIMENTS



GITHUB

Deep RL Agents Examples - Arthur Juliani
<https://github.com/awjuliani/DeepRL-Agents>

Gym Minecraft - Tambet Matiisen
<https://github.com/tambetm/gym-minecraft>

Minecraft DQN - Clément Romac & Vincent Beraud
<https://github.com/vincentberaud/Minecraft-Reinforcement-Learning>

Tic Tac Toe DQN - Clément Romac
<https://github.com/ClementRomac/DeepQLearning-TicTacToe>

SOURCES

Image Sources

- <https://medium.com>
- <https://blog.openai.com>
- <https://www.microsoft.com/>
- <https://stats.stackexchange.com/>
- <https://towardsdatascience.com/>
- <https://www.kickstarter.com>
- <https://www.theverge.com>
- <https://supermariorun.com/>
- <https://www.kdnuggets.com/>
- <https://www.networkauctions.co.uk>
- <https://pvoodoo.blogspot.com/>
- <https://www.prestoeng.co.uk/>
- <http://www.historiquedesjeuxvideo.com/>
- <https://leonardoaraujosantos.gitbooks.io>
- <https://en.wikipedia.org>
- <http://colah.github.io>

Papers

- [Human-level control through deep reinforcement learning](#)
- [Deep Recurrent Q-Learning for Partially Observable MDPs - Matthew Hausknecht and Peter Stone](#)
- [Deep Learning for Video Game Playing - Niels Justesen and Philip Bontrager and Julian Togelius and Sebastian Risi](#)
- [Playing FPS Games with Deep Reinforcement Learning](#)
- [Deep Reinforcement Learning with Double Q-learning - Hado van Hasselt and Arthur Guez and David Silver](#)
- [Teacher-Student Curriculum Learning - Tabet Matiisen and Avital Oliver and Taco Cohen and John Schulman](#)