January 11th, 2019
11:00 am PST
Online via Zoom

AI Research Discussion:

Dopamine - A Framework for Deep RL

Accel.AI
AIDLE
LXAI

AI & Deep Learning Enthusiasts

AIDLE

Meetup

https://www.meetup.com/Artificial-Intelligence-Deep-Learning-Enthusiasts-Meetup/

Github

https://github.com/AccelAI/AI-DL-Enthusiasts-Meetup

Accel.AI

https://www.accel.ai/

LatinX in AI Coalition

http://www.latinxinai.org/

Accel.AI

AI & Deep Learning Enthusiasts

AIDLE

Attendee Guidelines

- Try to read the paper and find a few good questions and/or insights.
- Be on time, latecomers can be disruptive for online discussions
- Use a headset in a quiet location if possible, and have a good internet connection: good audio quality is essential
- Have the discussion paper on your desktop ready to show via screen sharing in the event you need to point to something in a figure e.g.
- There are no silly questions!
- AIDLE has strict guidelines regarding discrimination based on gender, ethnicity, sexual orientation, class, and prior experience or educational background. We are all here to learn and grow! Violators will be removed from the group immediately.

Accel.AI OpenCollective

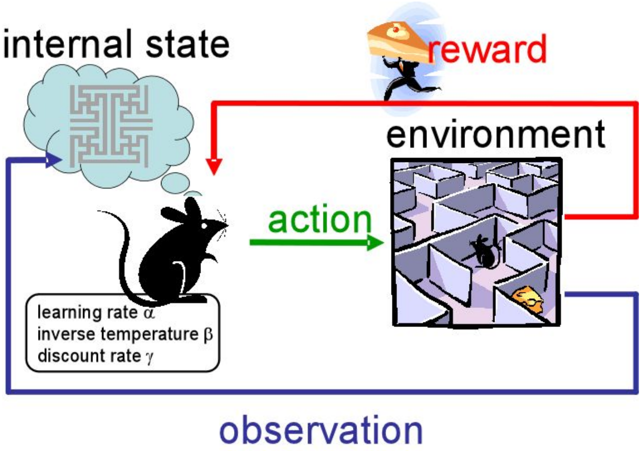
https://opencollective.com/accel-ai

Accel.AI

Dopamine: Introduction

Deep Reinforcement Learning

Complex agent-environment interactions



Tensorflow Based

- Composed of 12 Python Files

Code on Github

https://github.com/google/dopamine

Colab

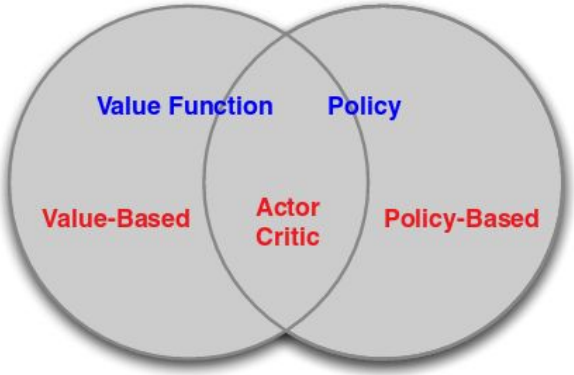
https://colab.research.google.com/github/google/dopamine/blob/master/dopamine/colab/agents.ipynb

Value Based Deep RL

Estimate the value of optimal value function Q*(s,a)

- Maximum value achievable under any policy

→ Architecture Research
→ Comprehensive Studies
→ Visualizations
→ **Algorithmic Research***
→ **Instruction***



Accel.AI

DQN (Deep Q-Network)

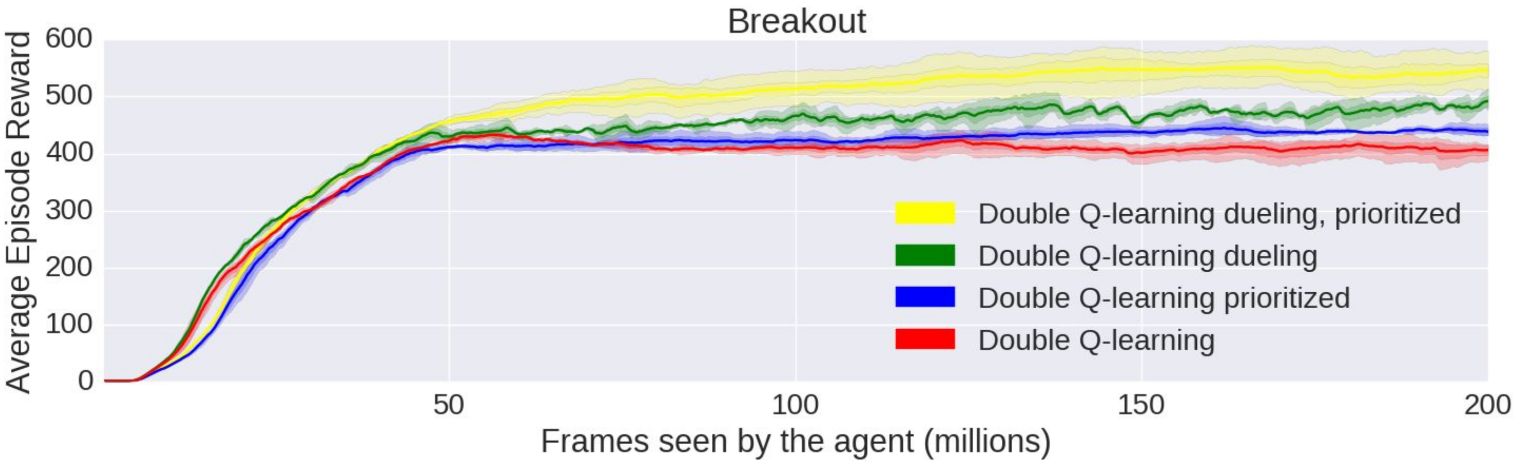
DeepMind (Published in Nature)

OpenAI Implementation

- Baseline for DRL
- Complex - 6 modules

DQN Innovative in use of agent architecture

- Target Network
- Replay Memory
- Atari Specific Preprocessing



Accel.AI

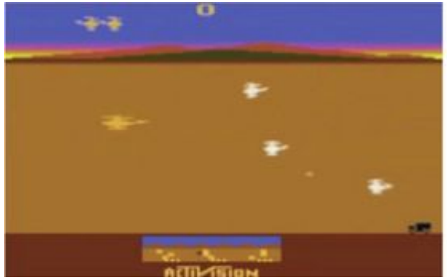
Software for DRL Research

Critical Operations for DRL Research

- ➔ Component Modularity
- ➔ Automatic Differentiation
- ➔ Visualization

Dopamine

- ➔ Fundamental Research
- ➔ Simulated Environments
  - ◆ **Arcade Learning Environments (ALE)***
  - ◆ Continuous Control
  - ◆ First Person Environments
  - ◆ Real Time Strategy Games



Accel.AI

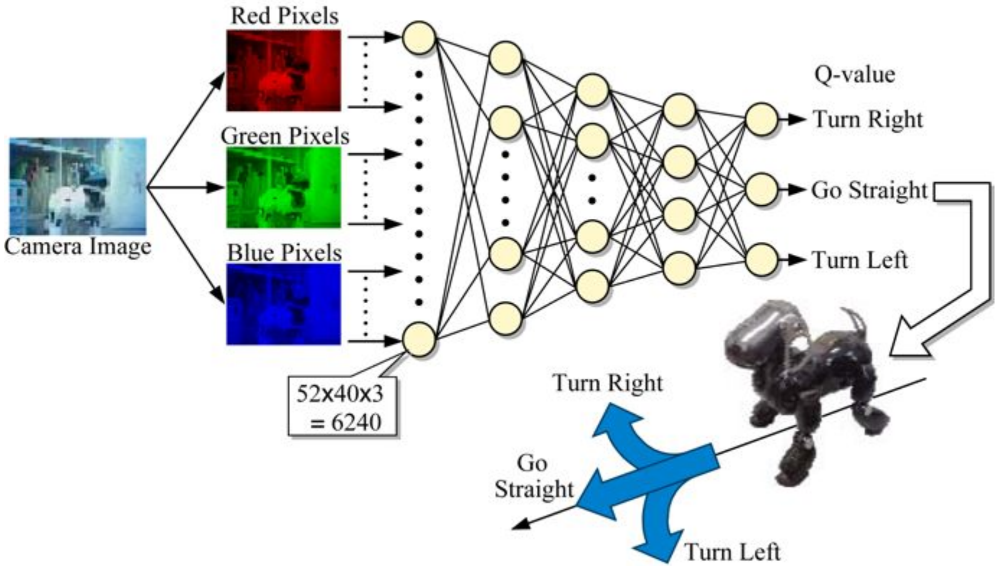
Case Study: Deep Q-Network Architecture

Architecture Research

➔ Interaction between components
  ◆ Network Topologies
➔ Agents composed on multiple interacting components

Algorithmic Research

➔ Improve Underlying Algorithms
  ◆ Learning
  ◆ Behavior
➔ Not tied to Agent Architecture



Case Study: Deep Q-Network Architecture

Q-Learning Robustness Increased

- ➔ Double DQN & Gap Increasing Methods
  - ◆ Prioritized Experience Replay
  - ◆ Frequent Sampling of State
  - ◆ High Prediction Error
- ➔ Retrace
  - ◆ Computes sum of n-step returns
  - ◆ Weighted by truncated correction ratio derived from policy probabilities
- ➔ Dueling Algorithm
  - ◆ Separates estimation of Q-values
  - ◆ Advantage & Baseline Components
- ➔ Distributional Methods
  - ◆ Introduces new losses to algorithm
  - ◆ Replaces scalar made by DQN with value distribution

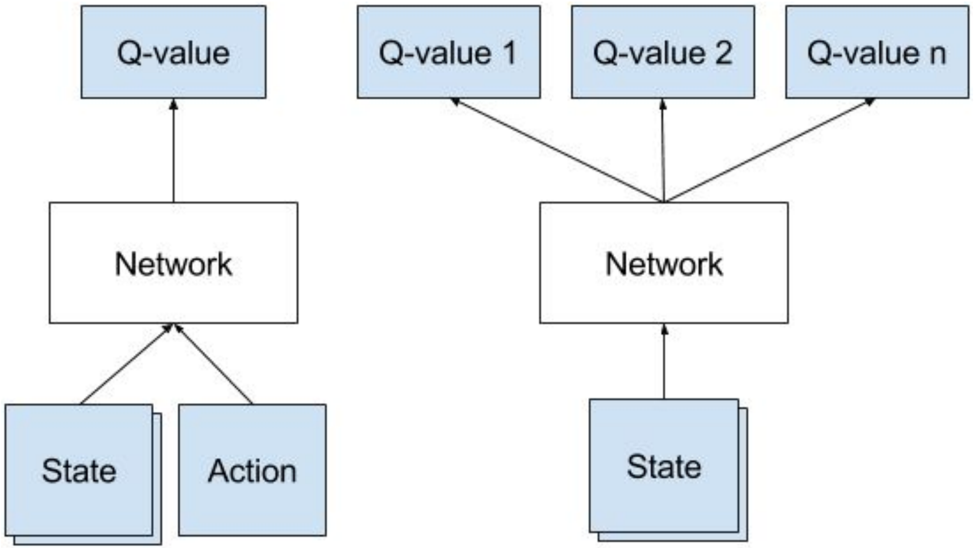
Accel.AI

Case Study: Deep Q-Network Architecture

Visualization

- ➔ Focused Interventions
  - ◆ Deeper understanding of Deep RL Methods
- ➔ Performance Not Emphasized
- ➔ Not Restricted to Visual Analysis

*Although majority of research focuses on visualizations



Accel.AI

How are research objectives enabled by software?

- ➔ Existing code availability vs new code written
- ➔ Shelf life of code written during research
- ➔ Value of high performance code

- ➔ Code Complexity
  - ◆ Complex Framework (large, modular, abstract)
  - ◆ Simple Framework (compact, monolithic, readily understood)

*Dopamine positions itself on the side of simplicity, aiming to facilitate algorithmic research and instructional purposes

Accel.AI

Dopamine: Design Principles

- *Easy experimentation*: Make it easy for new users to run benchmark experiments.
- *Compact and reliable*: Provide implementations for a few, battle-tested algorithms.
- *Flexible development*: Make it easy for new users to try out research ideas.
- *Reproducible*: Facilitate reproducibility in results.

Dopamine: Self Contained & Compact

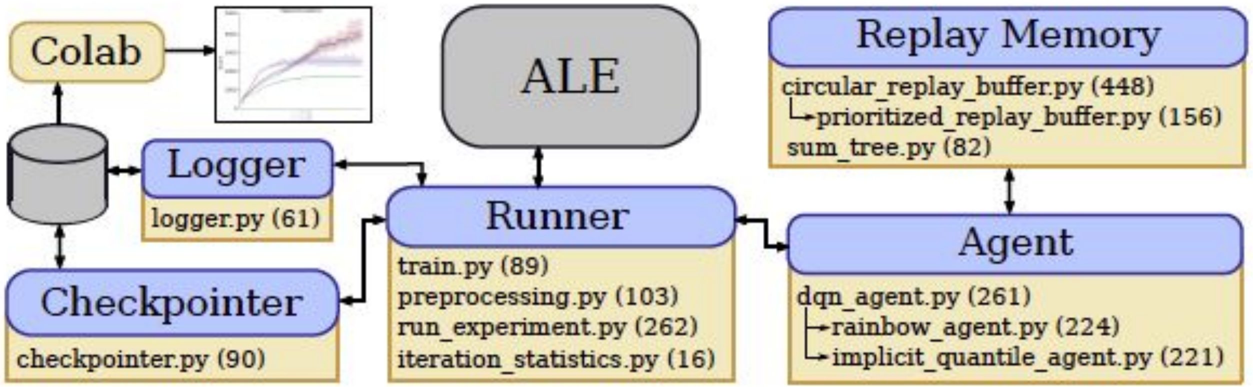


Figure 1: The design of Dopamine. Blue boxes are software components. Yellow boxes indicate the file names composing the software components; directed arrows indicate class inheritance, while the numbers in parentheses indicate the number of non-comment Python lines.

Accel.AI

Dopamine: Self Contained & Compact

```
DQNAgent.gamma = 0.99
DQNAgent.epsilon_train = 0.01
DQNAgent.epsilon_decay_period = 250000  # agent steps
DQNAgent.optimizer = @tf.train.RMSPropOptimizer()
tf.train.RMSPropOptimizer.learning_rate = 0.00025
Runner.sticky_actions = True
WrappedReplayBuffer.replay_capacity = 1000000
WrappedReplayBuffer.batch_size = 32
```

Figure 2: A few lines from the DQN gin-config file implementing the default Dopamine settings.

Accel.AI

Dopamine: Reliable & Reproducible

	Dopamine	DQN	C51	Rainbow	IQN
Sticky actions	Yes	No	No	No	No
Epis. termination	Game Over	Life Loss	Life Loss	Life Loss	Life Loss
Training ϵ	0.01	0.1	0.01	0.01	0.01
Evaluation ϵ	0.001	0.01	0.001	0.001	0.001
ϵ decay schedule (frames)	1M	4M	4M	1M	4M
Min. history to learn (frames)	80K	200K	200K	80K	200K
Target net. update freq. (frames)	32K	40K	40K	32K	40K

Table 1: Comparison of the hyperparameters used by published agents and our default settings.

5

Accel.AI

Dopamine: Reliable & Reproducible

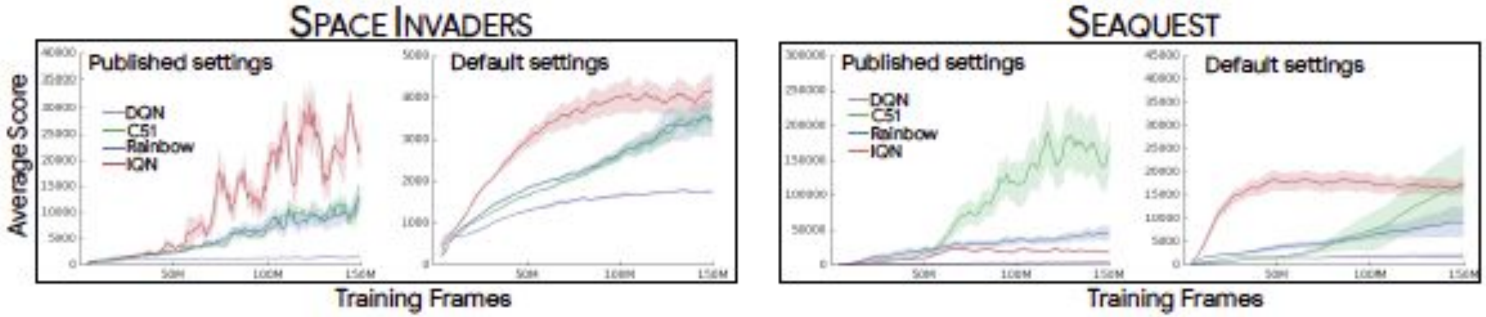


Figure 3: Comparing results of the published settings against the default settings, averaged over 5 runs. In each game, note that the y-scales between published and default settings are different; this is due mostly to the use of sticky actions in our default setting. The relative dynamics between the different algorithms appear unchanged for SpaceInvaders, but are quite different in Seaquest.

Accel.AI

Dopamine: Reliable & Reproducible

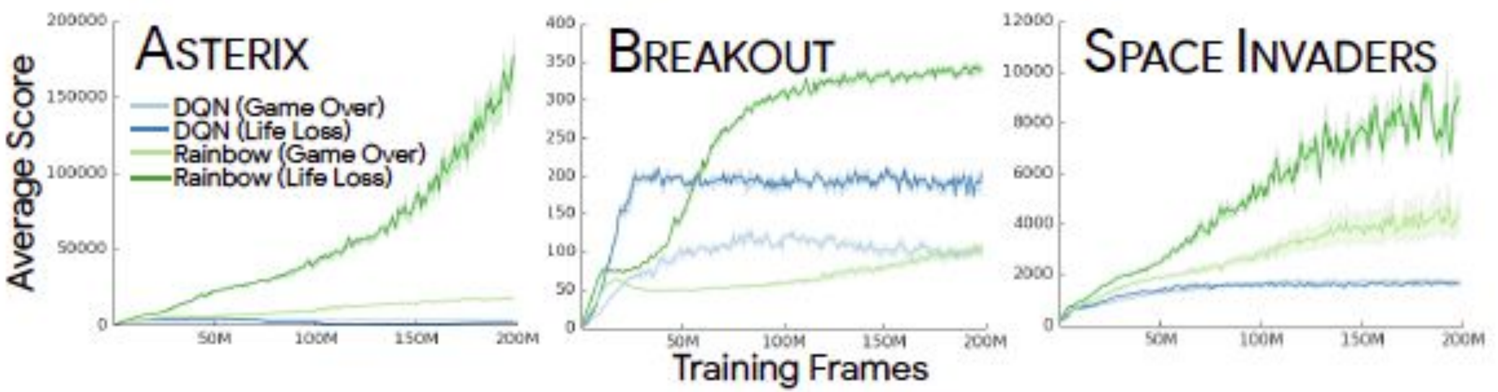


Figure 4: Effect of terminating episodes on Life Loss or Game Over, averaged over 5 runs; this choice can have a dramatic difference in reported performance, although less so for DQN.

Accel.AI

Dopamine: Training & Performance

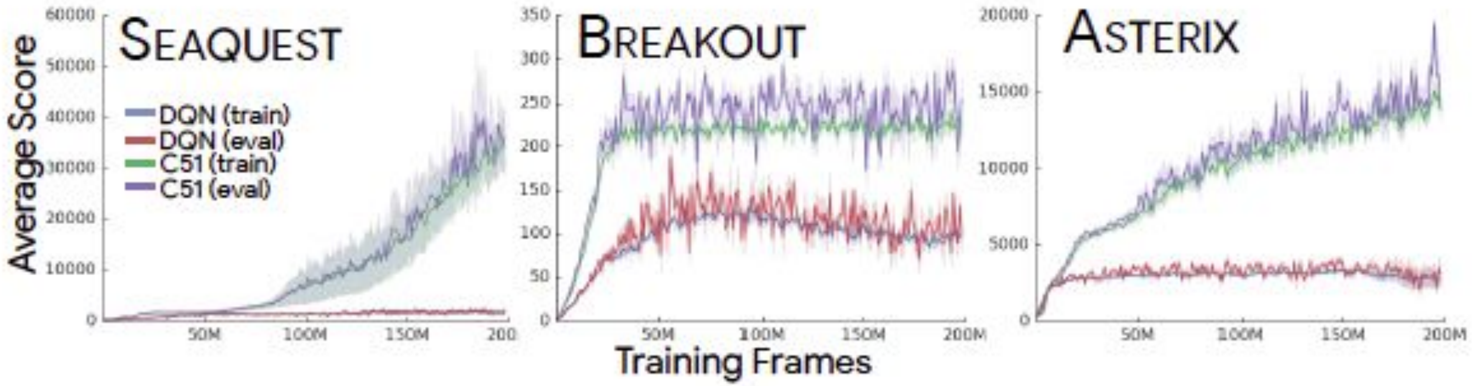


Figure 5: Training and evaluation scores, averaged over 5 runs. For our default values of training and evaluation ϵ, there is a minimal difference between the two settings.

Accel.AI

Dopamine: Sticky Actions & Performance

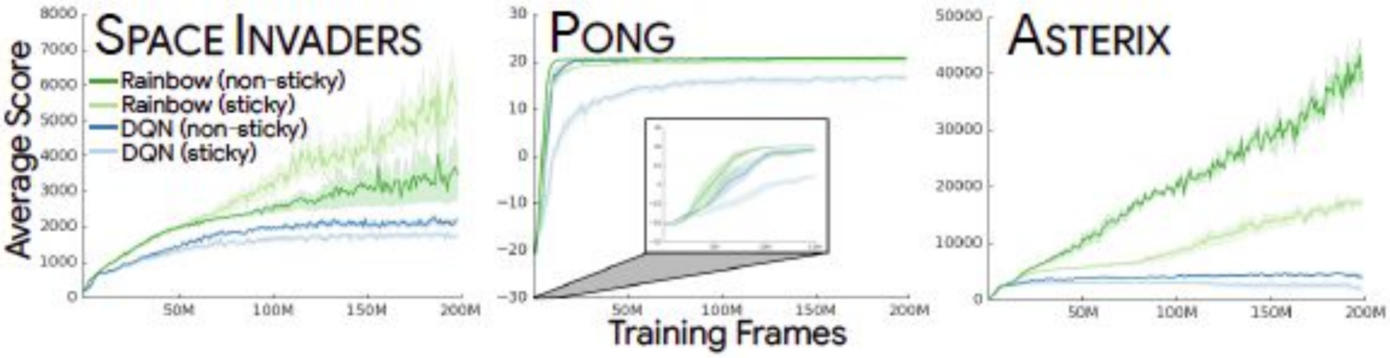


Figure 6: Effect of sticky vs. non-sticky actions. Sticky actions affect the performance of both Rainbow and DQN, but preserve the qualitative differences in performance between the two.

Accel.AI

Dopamine: Related Work

Deep RL Frameworks

- ➔ OpenAI baselines (Dhariwal et al., 2017)
- ➔ Coach from Intel (Caspi et al., 2017)
- ➔ Tensorforce (Schaarschmidt et al., 2017)
- ➔ Keras RL (Plappert, 2016)
- ➔ RLLab (Duan et al., 2016)
- ➔ RLLib (Liang et al., 2018)
- ➔ ELF (Tian et al., 2017)

Individual Agents

- ➔ DQN in Lua
- ➔ PPO implementations (e.g. Hafner et al., 2017)
- ➔ IMPALA open-sourced (Espeholt et al., 2018)

RL Frameworks

- ➔ RLGlue (Tanner & White, 2009)
- ➔ BURLAP (MacGlashan, 2016)
- ➔ PyBrain (Schaul et al., 2010)
- ➔ RLPy (Geramifard et al., 2015)

Accel.AI

Additional Resources

Pablo Samuel Castro, PhD

- ➔ Google AI Blog
- ➔ NeurIPS Slides

Siraj Rival

- ➔ Dopamine Live
  - ◆ Python notebook on Github
- ➔ Mathematics of Dopamine
  - ◆ Code on Github

Lex Fridman - MIT

- ➔ Deep Reinforcement Learning

Arxiv Insights with Zander

- ➔ An introduction to Reinforcement Learning

Jabrils

- ➔ Making my First Machine Learning Game

Dopamine Neurotransmitter

- ➔ 2 Min Neuroscience - Dopamine
- ➔ What is Dopamine?

Accel.AI

AI & Deep Learning Enthusiasts

AIDLE

Meetup

https://www.meetup.com/Artificial-Intelligence-Deep-Learning-Enthusiasts-Meetup/

Github

https://github.com/AccelAI/AI-DL-Enthusiasts-Meetup

Accel.AI

https://www.accel.ai/

LatinX in AI Coalition

http://www.latinxinai.org/

Accel.AI OpenCollective

https://opencollective.com/accel-ai

Accel.AI

Accel.AI

Join Us!

Accel AI Institute 501c3 Non Profit

info@accel.ai

@AccelerateAI

/accelai/

/accel.ai/

/company/accel.ai

www.accel.ai

www.latinxinai.org