



Reinforcement Learning

Введение

• REC

Проверить, идет ли запись

Меня хорошо видно
&& слышно?



Тема вебинара

Reinforcement learning.

Введение в обучение с подкреплением.

Игорь Стурейко



Руководитель курсов: Reinforcement Learning, ML Professional, ML Basic

Teamlead, главный инженер проекта,
Физический факультет МГУ, PhD теоретическая физика

Опыт:

Более 15 лет занимался прикладной математикой и мат моделированием
(Data Scientist) (Python, C++) в НИИ ПАО Газпром

Анализ временных рядов, эволюционные модели, финансовые модели

@stureiko (TG)

LinkedIn: [igor-stureiko](https://www.linkedin.com/in/igor-stureiko/)



Правила вебинара



Активно
участвуем



Off-topic обсуждаем
в учебной группе #OTUS
RL-2024-03



Задаем вопрос
в чат или голосом



Вопросы вижу в чате,
могу ответить не сразу

Условные обозначения



Индивидуально



Время, необходимое
на активность



Пишем в чат



Говорим голосом



Документ



Ответьте себе или
задайте вопрос

Маршрут вебинара



Цели вебинара

К концу занятия вы сможете

1. Выделить основные отличительные черты Reinforcement Learning
2. Понимать области применения Reinforcement Learning
3. Разобраться, что отличает Reinforcement Learning от других областей машинного обучения

Смысл

Зачем вам это знать

1. Для общего понимания современного состояния Reinforcement Learning
 2. Чтобы заложить основы для дальнейшего обучения на курсе
-

Знакомство

Команда курса



Игорь Стурейко

15+ лет опыта в прикладной математике и математическом моделировании
НИИгэзэкономика ПАО Газпром

Руководил командой, решающей прикладные задачи, вывод готовых моделей в продакшн и поддержку решений у клиента



Артем Голубин

ОЦРВ Руководитель группы, ML-инженер
Промышленный опыт использования RL,
разработка имитационной среды на
железной дороге и обучении RL-агентов
эффективному формированию и
маршрутизации грузовых поездов



Сергей Доронин

12+ лет опыта создания и разработки корпоративных сервисов
(ЭкспрессВолгаБанк, ПАО "Ростелеком" и др.)

Подбирал технологии виртуальных
датацентров, методов защиты информации
(DLP/WAF), проектирования систем
видеонаблюдения, Wi-Fi, VoIP, IPTV и т.д.



Анатолий Чудаков

Teamlead в команде кинотеатра Wink.
Промышленный опыт SOTA разработки RL в
рекомендательных системах и uplift
моделировании

Образование: Прикладная макроэкономика
(СПбГУ) Otus Certified Educator

Расскажите о себе

Напишите, пожалуйста, в чат или скажите голосом



Как вас зовут?



Какой опыт в IT?



Какие ожидания
от курса?



Заполните информацию
в разделе «О себе»
в личном кабинете

Как подготовиться к обучению

Полезный курс о том, как эффективно учиться

Занятия

Тренажеры

1 месяц

Оплачено

Как учиться в OTUS 

Содержание курса

 Предисловие

- 1 [Часть 1. Как подготовиться к обучению в OTUS](#)
- 2 [Часть 2. Как учатся взрослые](#)
- 3 [Часть 3. Что делать во время прохождения обучения](#)
- 4 [Часть 4. Есть ли жизнь после обучения](#)
- 5 [Часть 5. OTUS рекомендует](#)

 Какой ты OTUS HEROES? Скажи мне, как ты учишься, и я скажу кто ты

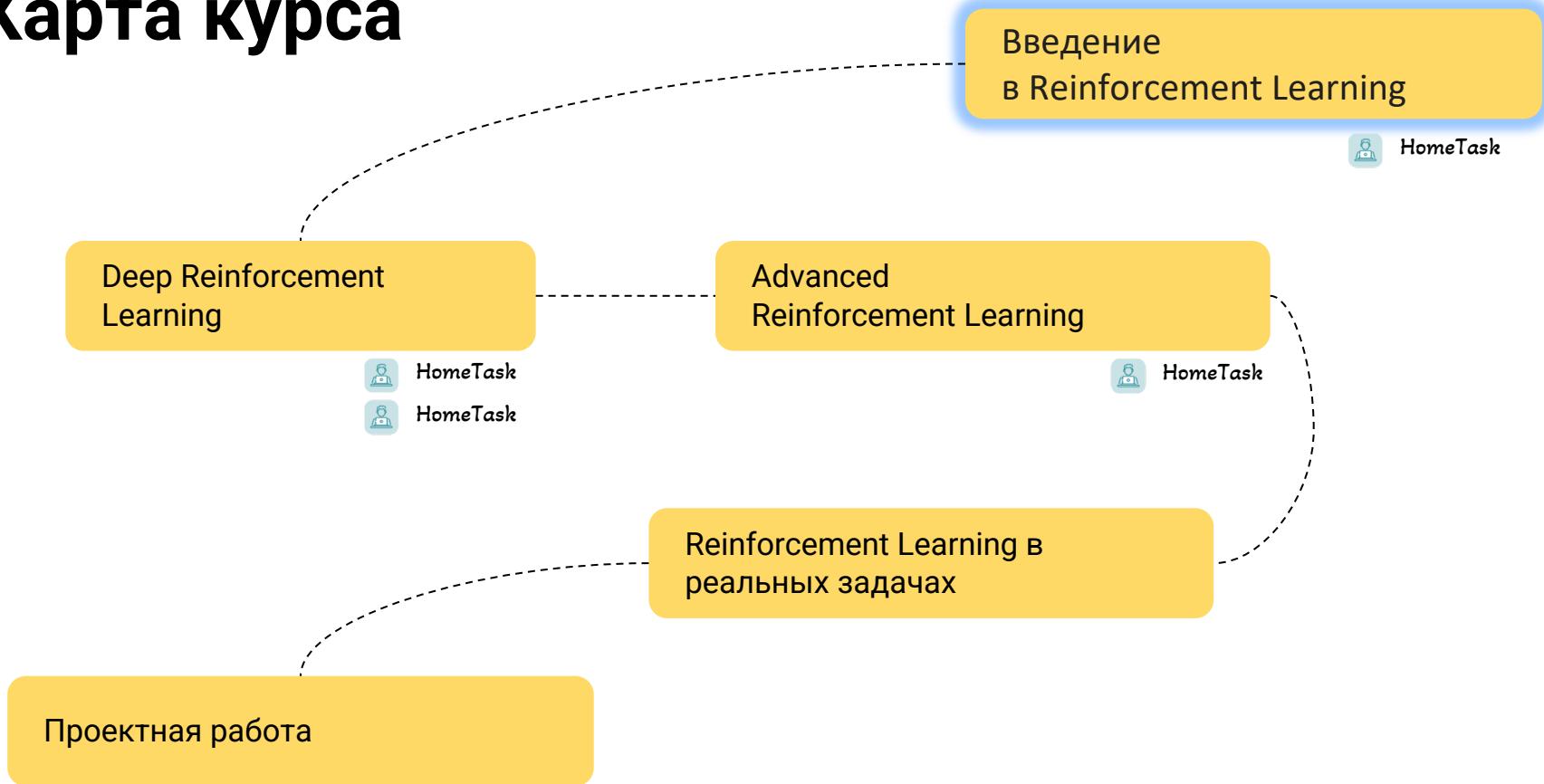
 Заключительное напутствие

 Workbook

 Дополнительные материалы

О курсе

Карта курса



Программа курса



Введение в Reinforcement Learning



27.03 -

Знакомство с Reinforcement Learning



03.04 -

Ключевые понятия RL: агент, среда, награда, политика
Построение модели среды и агента



08.04 -

Основные алгоритмы RL: Value based



10.04 -

Основные алгоритмы RL: Policy based



HomeTask



Программа курса



Deep Reinforcement Learning



15.04 - Введение в Deep Reinforcement Learning



17.04 - Deep Q-Network (DQN) алгоритм



HomeTask



22.04 - Policy Gradient (PG) алгоритм



24.04 - Actor-Critic алгоритм



HomeTask



06.05 - TRPO -> PPO



08.05 - DDPG -> TD3 -> LSTM-TD3



Программа курса



Advanced Reinforcement Learning



13.05 -

Обучение с использованием модели среды, часть 1



HomeTask



15.05 -

Обучение с использованием модели среды, часть 2



20.05 -

Иерархическое обучение с подкреплением



22.05 -

Многоагентное обучение и кооперація агентов



27.05 -

Выбор темы и организация проектной работы



Программа курса



Reinforcement Learning в реальных задачах



29.05 - → Применение RL в игровой индустрии



03.06 - → Применение RL в робототехнике



05.06 - → Применение RL в рекомендательных системах



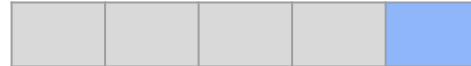
10.06 - → Применение RL в задаче кредитного scoringa



17.06 - → Применение RL в управлении финансовым портфелем



Программа курса



Проектная работа



19.06 -

Консультация по проектам и домашним заданиям



24.06 -

Предзащита



26.06 -

Защита проектных работ



Результат от курса

1) Что получите от курса

- a) познакомитесь с новой областью машинного обучения
 - научитесь разрабатывать решения на базе Reinforcement Learning
 - изучите ключевые особенности и основные алгоритмы Reinforcement Learning
- b) Познакомитесь с коллегами и возможно **создадите группу** для собственного пет-проекта

2) За время курса

- a) Напишете простую библиотеку с основными алгоритмами
 - Value-based алгоритм для решения задачи в дискретной среде
 - Policy-based алгоритм для решения задач с бесконечной областью действий
 - Мультиагентную систему
- b) Создадите финальный проект лично или коллективно с одногруппниками



Как будем учиться?



Вебинары

Вторник, четверг, 20:00.

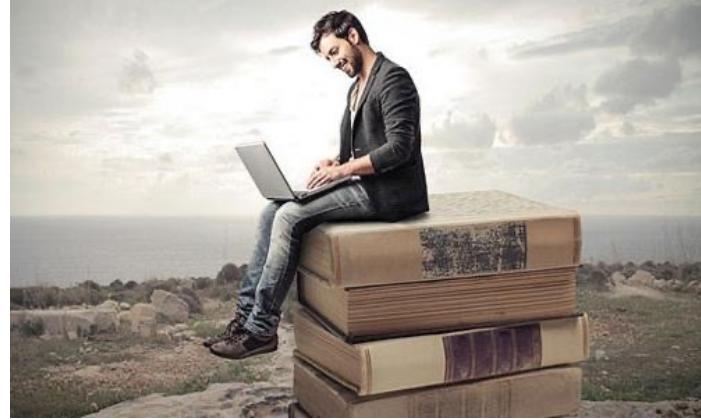
(запись и материалы
выкладывают, как правило,
на следующий день после
вебинара)



Домашние задания

1 дз в неделю.

Дедлайна нет, кроме
окончания курса :)
Типовой срок проверки:
2-3 дня



Чат в телеграм

Задавайте вопросы,
обменивайтесь инсайтами.
Наш чат – это
#OTUS RL-2023-10



Отзывы



Всегда рады вашим
конструктивным отзывам :)



Вебинары во многом адаптивны.
Мы следим за результатами опросов,
обсуждениями в телеграм и ДЗ.



Лайфхаки



Определитесь зачем оно вам?
Зачем вы пришли на курс?



Сделайте упор на тему,
которая вам важна



Старайтесь полученные знания
применять на практике



Задавайте вопросы,
так материал лучше
усваивается



Регулярное выполняйте ДЗ
(наверстать пропуски тяжело)



Обсуждайте материал и
задачи в учебной группе



Объединяйтесь в группе и
выполняйте проект
совместно

Мотивация

Приготовьтесь к волнообразной реакции
от «ух ты как интересно», до «куда я попал?»

* в такие моменты вспоминайте
про цели обучения, которые вас мотивируют)



Первые уроки позволяют адаптироваться
и «влиться» в процесс и могут показаться
вам слишком простыми, особенно если вы
имеете опыт разработки

Это нормально :)



Организационные вопросы



Про технические моменты:

- Python 3, virtual environment (venv / conda / poetry)
- Пакетные менеджеры pip / conda
- Git / Github



Какие инструменты будем использовать:

- Gymnasium (<https://gymnasium.farama.org/index.html#>)
- PyTorch
- Jupyter notebook / VSCode / PyCharm



Практика на курсе:

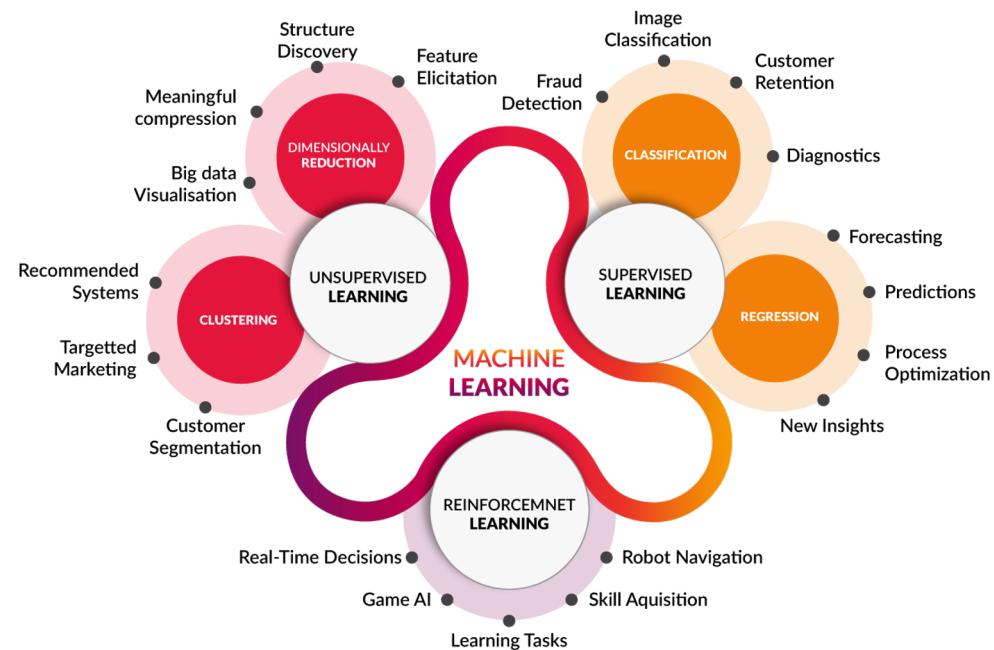
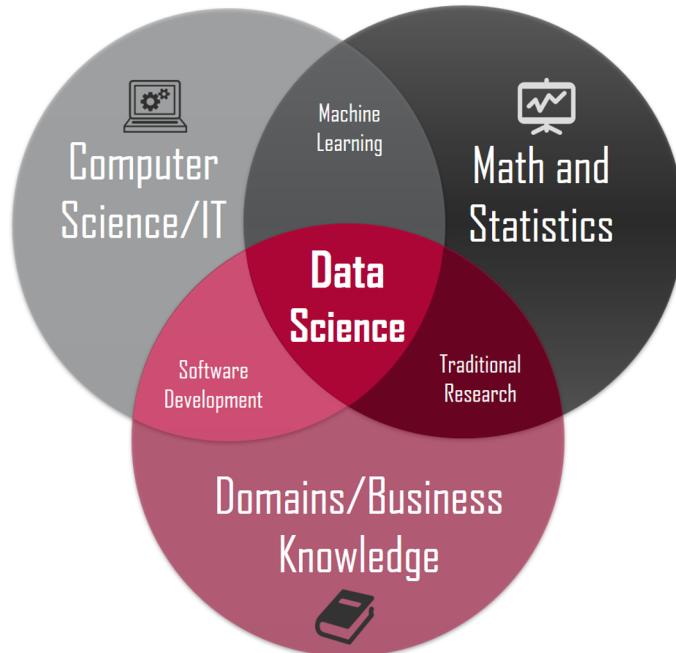
- разработка алгоритмов



Ваши вопросы?

Основы и зарождение Reinforcement Learning

Третья парадигма машинного обучения



Третья парадигма машинного обучения



Обучение с учителем

Модель обучается на основе прецедентов (пары стимул-реакция), с целью восстановления / приближения потенциальной функциональной связи между входными и выходными сигналами.

Обучение без учителя

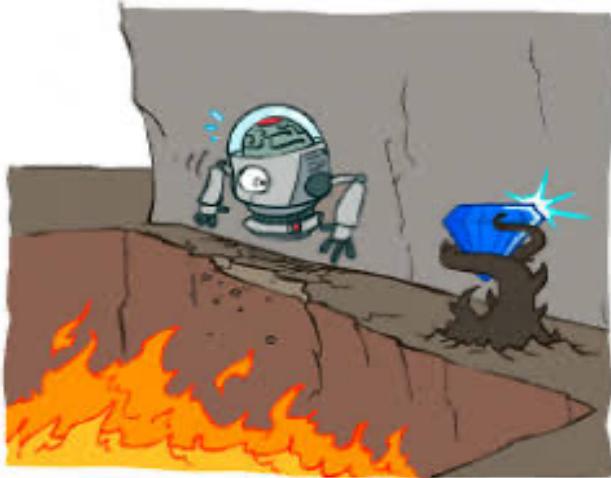
Модель обучается с целью обнаружения структуры в некотором пространстве объектов (внутренние взаимосвязи, зависимости, закономерности).

Обучение с подкреплением

Агент взаимодействует со средой методом проб и ошибок, чтобы выработать наиболее оптимальное поведение для максимизации получаемого вознаграждения.

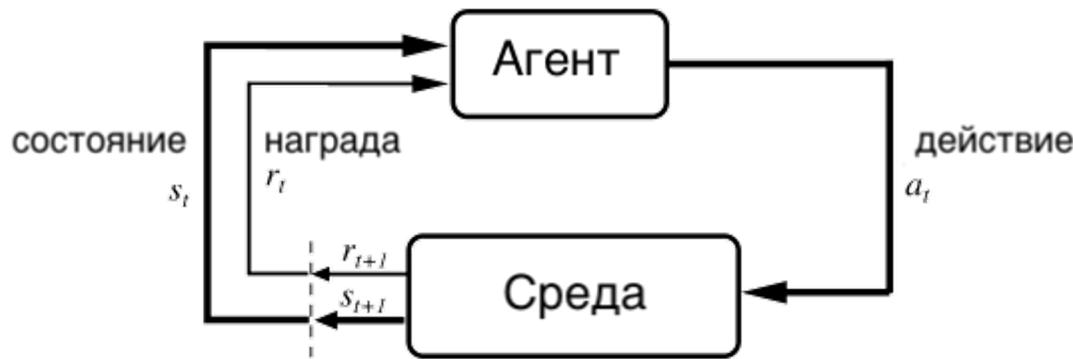
Агент своими действиями может влиять на последующие состояния среды и получаемое вознаграждение (в том числе отложенное).

Что такое Обучение с Подкреплением?



Этот метод обучения схож с тем, как люди и животные учатся на основе положительного и отрицательного подкрепления.

Обучение с Подкреплением - это направление машинного обучения, где агент (искусственный интеллект) обучается, принимая решения и действуя в некоторой среде получая награду и штрафы с целью максимизировать некую награду.



Вопросы?



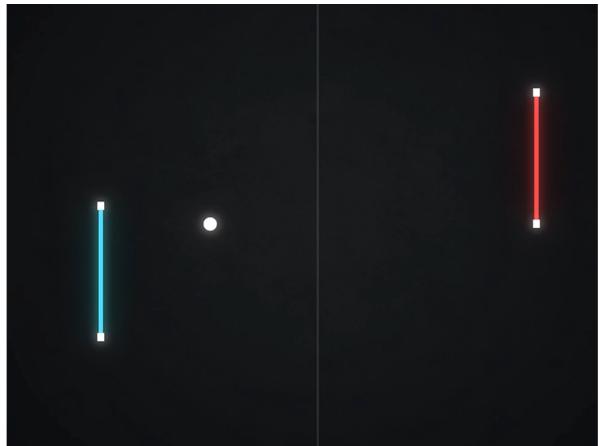
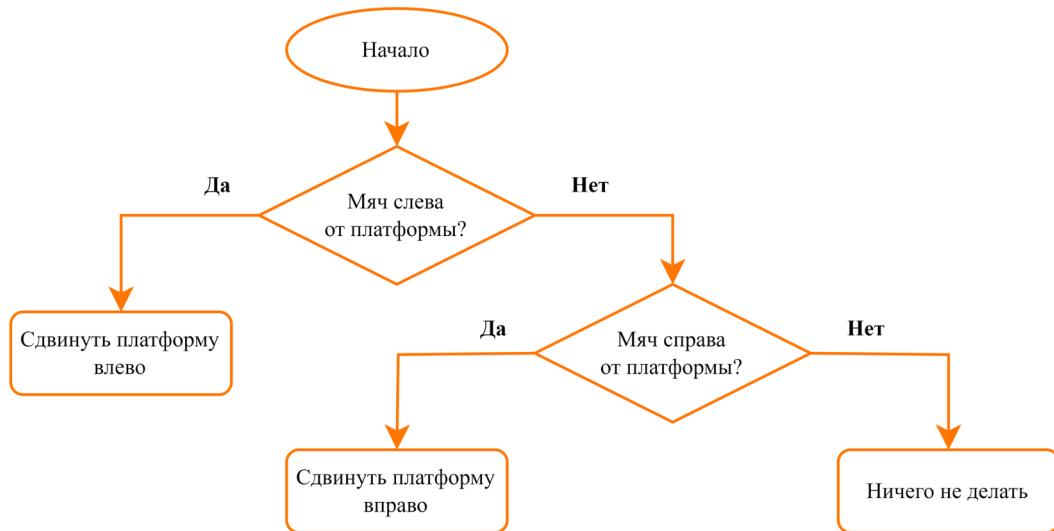
Задаем
вопросы



Ставим “-”,
если вопросов нет

Применение RL в играх

Система правил



Игра Pong



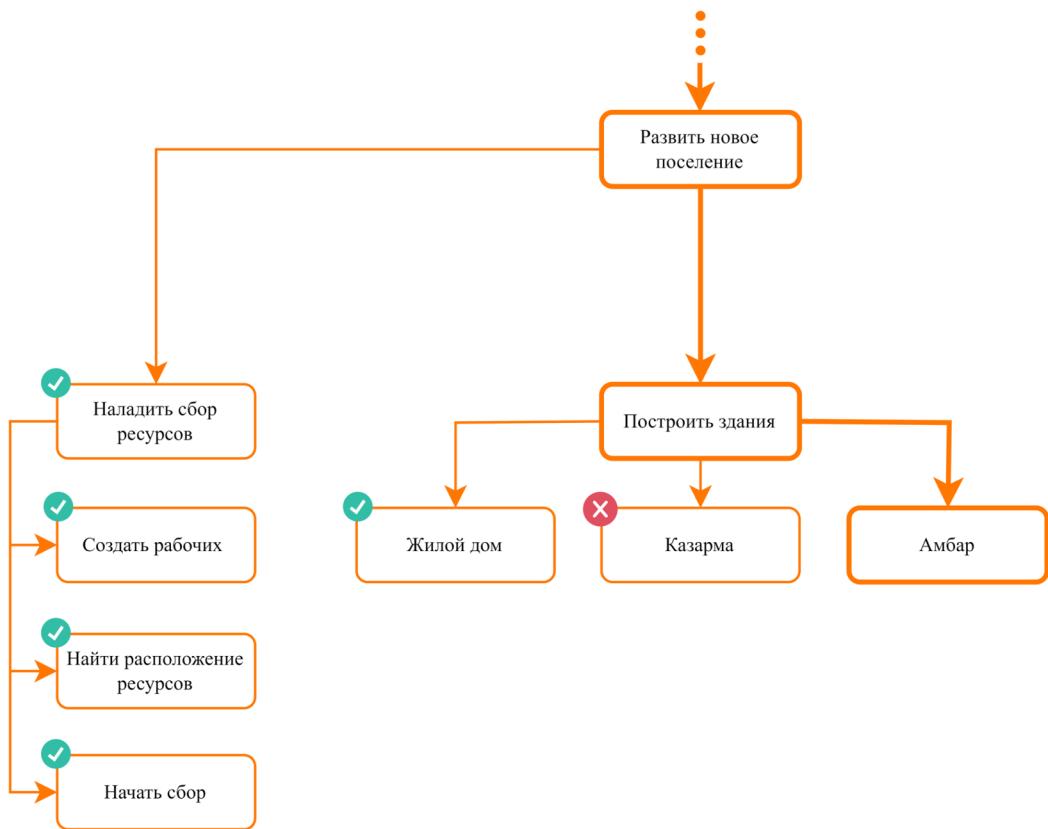
Конечные автоматы



Игра Doom



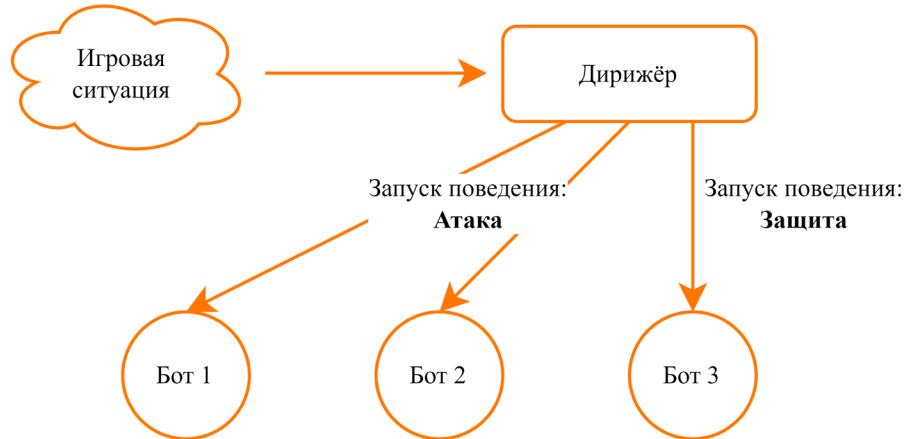
Деревья поведения



Игра Age of Empires



Коллективное поведение



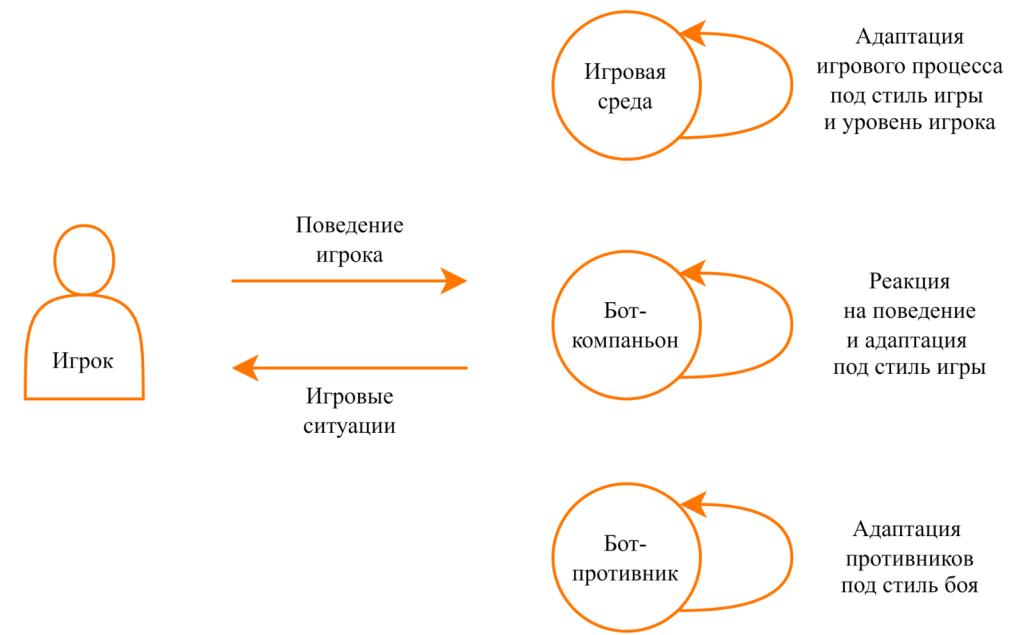
Игра F.E.A.R.



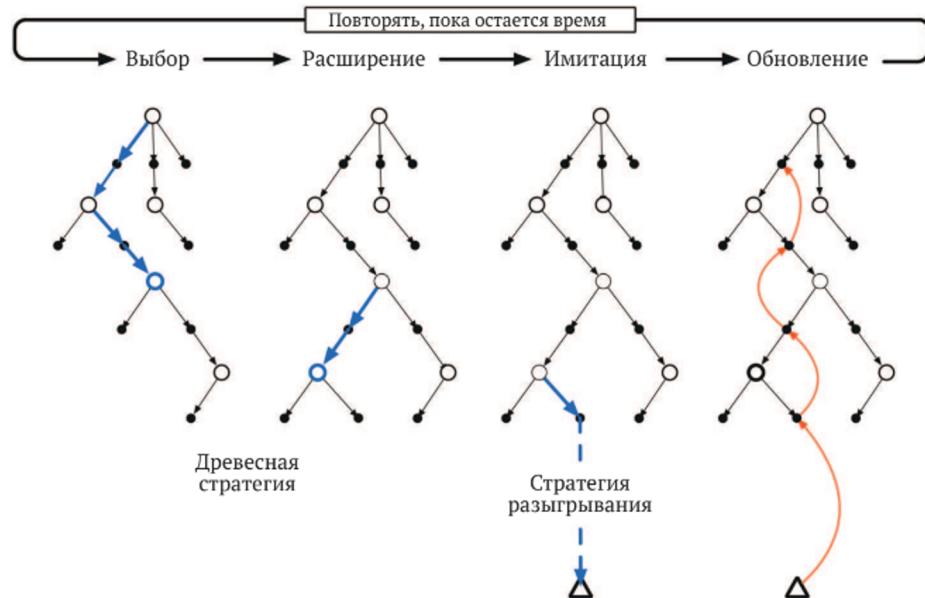
Игра Horizon Zero Dawn



Адаптивное поведение

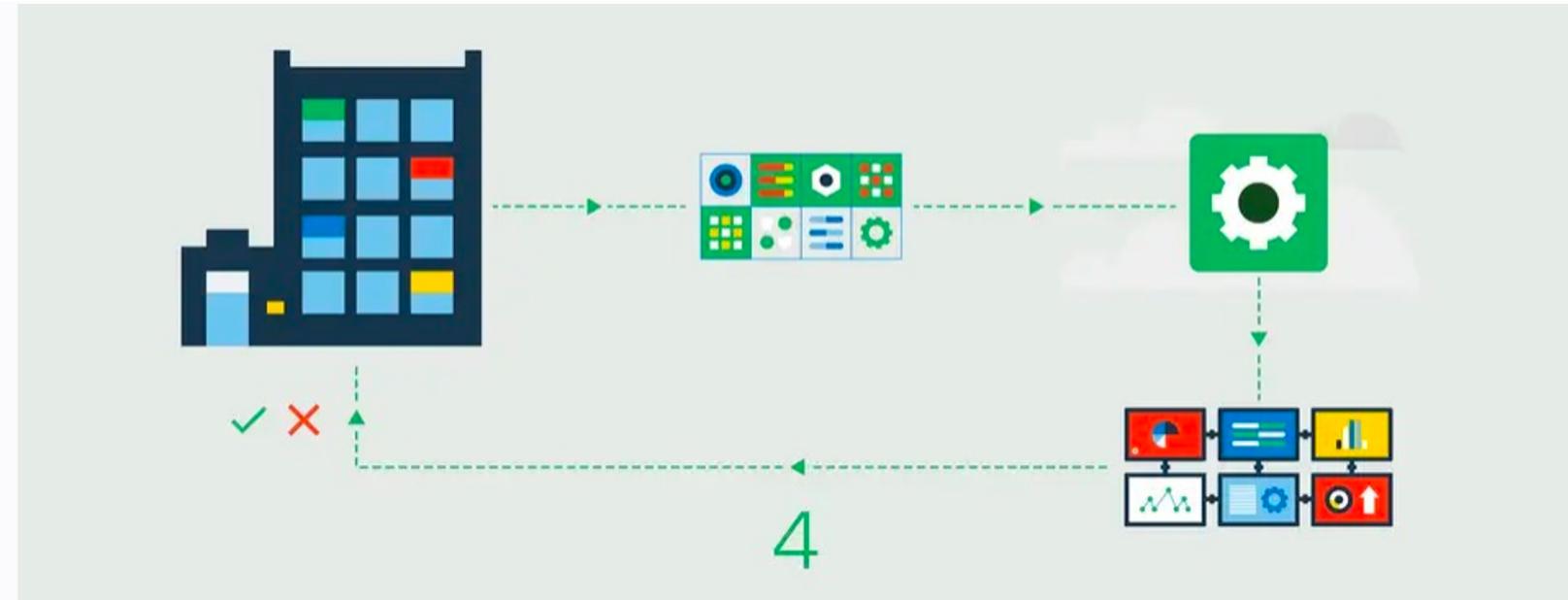


Игра на основе правил



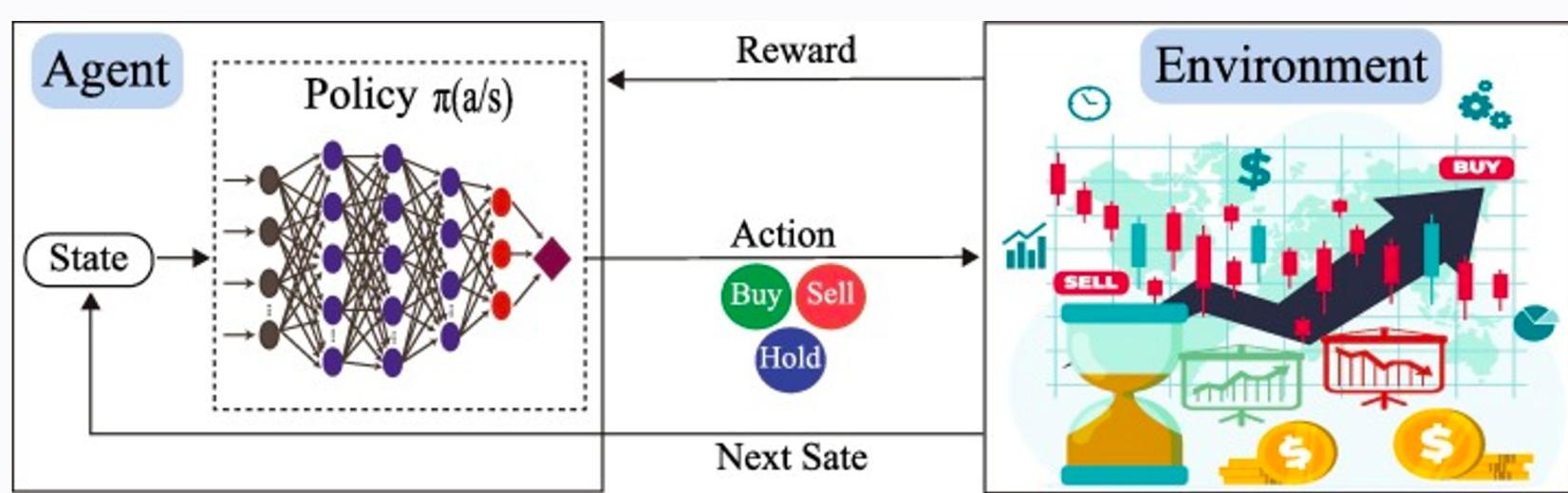
Современные применения

Охлаждение дата-центров



Источник: <https://www.deepmind.com/blog/safety-first-ai-for-autonomous-data-centre-cooling-and-industrial-control>

Торговля на бирже



АВТОНОМНОЕ ВОЖДЕНИЕ



DeepTraffic
Deep Reinforcement Learning Competition

Name:
Lex Fridman

Highest Average Speed:
69.38 mph

Highest Ranking:
5 out of 1,871
On Jan 19, 2017 with 68.97 mph

Current Ranking:
2,276 out of 22,687
On Jan 08, 2018 with 69.38 mph



Sensing:
Side Sensing: 3
Forward Sensing: 30
Backward Sensing: 10
Temporal Window: 0

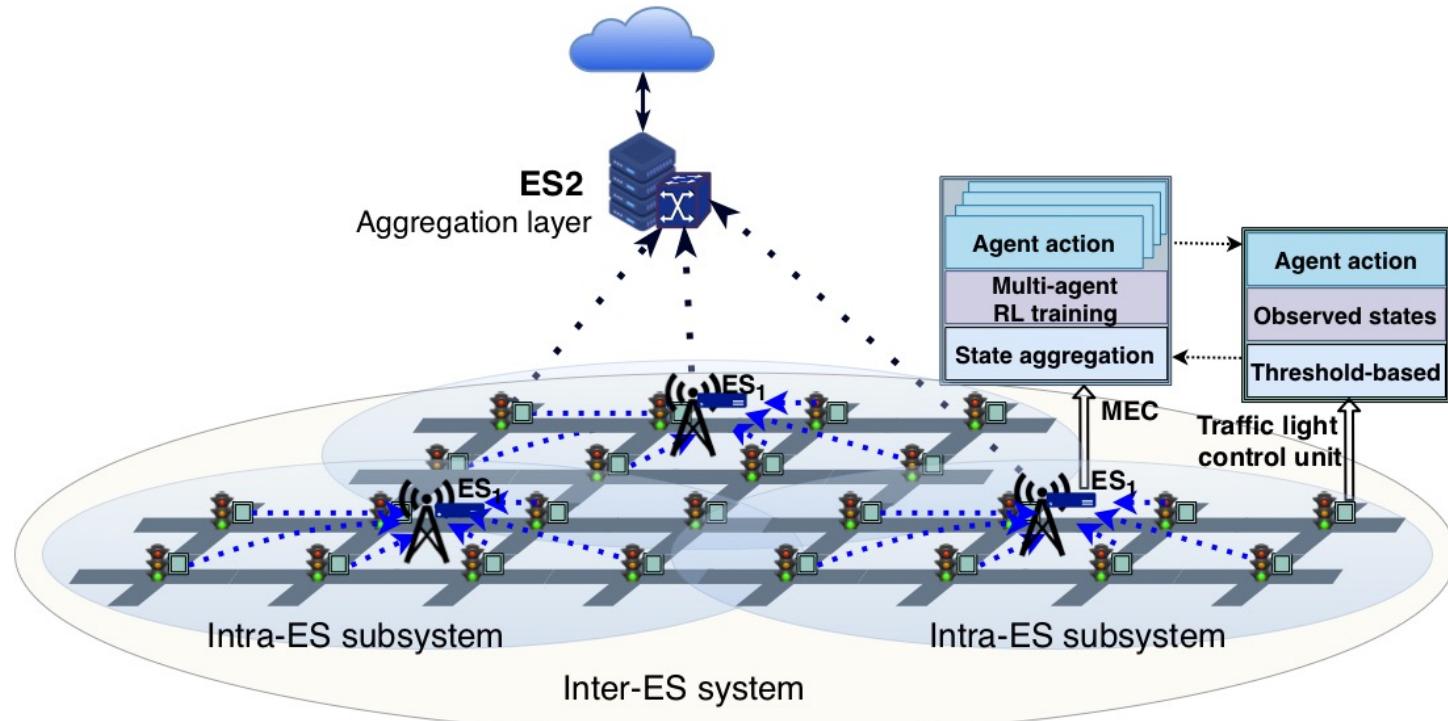
Network Architecture:
Layers: 3
Parameters: 11,445

Learning Parameters:
Training Iteration: 10,000
Momentum: 0.0
Batch Size: 64
L2 Decay: 0.01
Learning Rate: 0.001

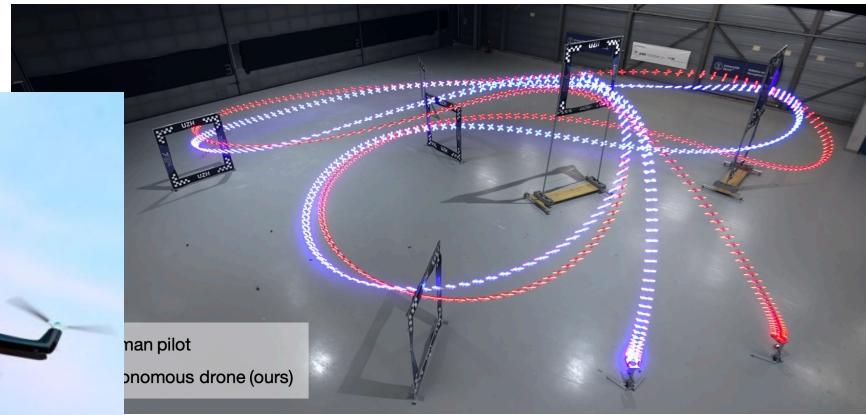
Reinforcement Learning:
Experience Size: 3,000
Gamma: 0.7
Number of Intelligent Cars: 10

MIT
selfdrivingcars.mit.edu

Контроль движения в городе



Многоагентные системы - рой дронов



Вопросы?



Задаем
вопросы



Ставим “-”,
если вопросов нет

Ключевые тезисы

1. Reinforcement Learning широко применяется в самых разных областях
2. Специалисты находят всё больше применений Reinforcement Learning
3. Reinforcement Learning можно выделить в отдельную группу алгоритмов машинного обучения
4. Существуют различные типы сред, задач и алгоритмов

Необходимые инструменты

Что установить?

- **Python**
 - Используя инструмент **pyenv**: <https://realpython.com/intro-to-pyenv/>
 - Используя инструмент **conda**: https://developers.google.com/earth-engine/guides/python_install-conda
- **Gymnasium**
 - pip install gymnasium
 - pip install "gymnasium[toy-text]"
 - <https://gymnasium.farama.org/>
- **Pytorch** (начиная со второго модуля)
 - <https://pytorch.org/>

Дополнительные инструменты

- **Jupyter** <https://jupyter.org/install>
- **OpenCV** (по желанию, для отрисовки игр из Gymnasium)
 - <https://pypi.org/project/opencv-python/>
- **Poetry** (по желанию, крайне желателен опыт разработки ПО)
 - <https://python-poetry.org/docs/#installation>

Список материалов для изучения

1. Обучение с подкреплением для самых маленьких

<https://habr.com/ru/articles/308094/>

2. Обучение с подкреплением: неформальное знакомство

<https://habr.com/ru/companies/wunderfund/articles/667654/>

3. Введение в обучение с подкреплением

<https://habr.com/ru/companies/otus/articles/429090/>

4. Мотивационное видео

<https://www.youtube.com/watch?v=9C06M2Hs0IA>

Вопросы?



Ставим “+”,
если вопросы есть



Ставим “-”,
если вопросов нет

Рефлексия

Цели вебинара

Проверка достижения целей

1. Выделить основные отличительные черты Reinforcement Learning
2. Понимать области применения Reinforcement Learning
3. Разобраться, что отличает Reinforcement Learning от других областей машинного обучения

Тезисы

Подведем итоги

1. Reinforcement Learning зарождался в психологии
2. Reinforcement Learning можно выделить в отдельную группу алгоритмов машинного обучения
3. Reinforcement Learning широко применяется в самых разных областях: от игр до управления охлаждения data-центров и от торговли на бирже до роевого поведения дронов

Рефлексия



С какими впечатлениями уходите с вебинара?
Какие примеры использования RL вам
показались наиболее интересными?



Что показалось сложным?

Следующий вебинар



3 апреля 2024

Ключевые понятия RL: агент, среда, награда, политика. Построение среды.



Ссылка на вебинар
будет в ЛК за 15 минут



Материалы
к занятию в ЛК –
можно изучать



Обязательный
материал обозначен
красной лентой



**Заполните, пожалуйста,
опрос о занятии
по ссылке в чате**