

Выпускная квалификационная работа
(Дипломная работа) специалиста

Применение генетического алгоритма для управления подвижным экипажем

Выполнил студент 622 группы: Кафанов Олег Игоревич

Научные руководители:

к.ф.-м.н. Шокуров Антон Вячеславович

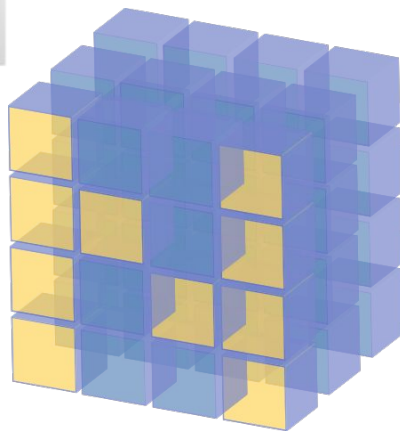
д.ф.-м.н. Косенко Иван Иванович

Актуальность

Обучение с подкреплением

- Задачи, где сложно построить математическую модель.
- Задачи, где необходимо, чтобы система работала на реальных данных, подстраиваясь под изменчивую окружающую среду.

Одной из таких задач является пилотирование автомобиля.



NumPy

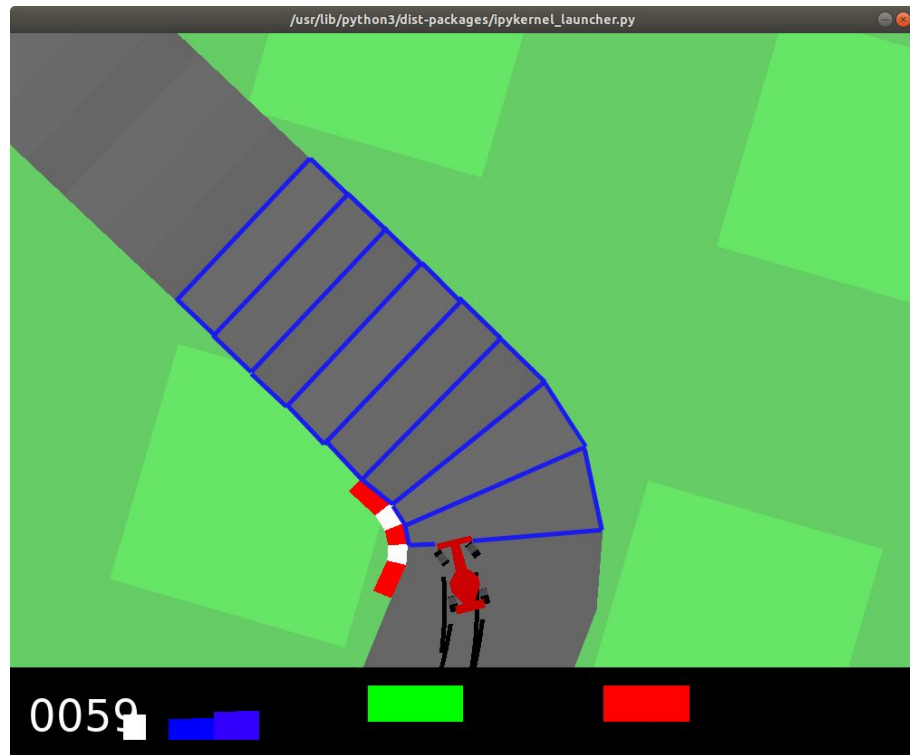


Геймплей и вознаграждение

Задача - пилотировать автомобиль через случайно сгенерированный двумерный мир трека и травы, достигнув финиша за максимально короткое время.

Трек разделен на фрагменты-плитки ;

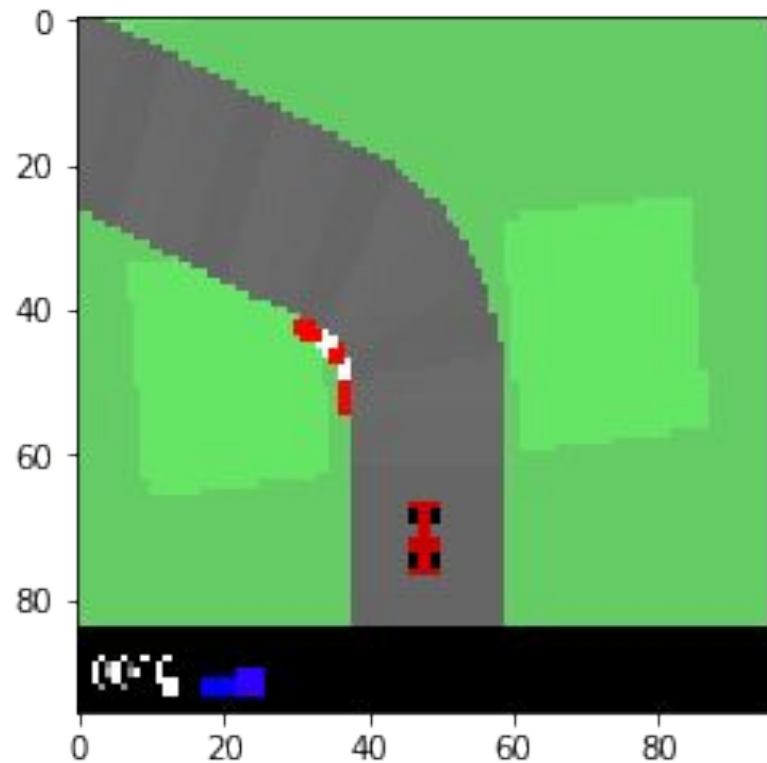
Вознаграждение составляет $-0, 1$ за каждый кадр и $+1000/N$ за каждый посещенный фрагмент трека (плитку), где N - общее количество фрагментов в треке.



Пространство состояний и пространство действий

Пространство состояний - последовательность кадров для игрового экрана. Каждый представлен в виде сетки $96 \times 96 \times 3$ значений RGB

Пространство действий представляет собой множество троек $(s, a, d) \in [-1, 1] \times [0, 1] \times [0, 1]$, где рулевое управление - s , газ - a , тормоз - d



Сокращение пространства состояний

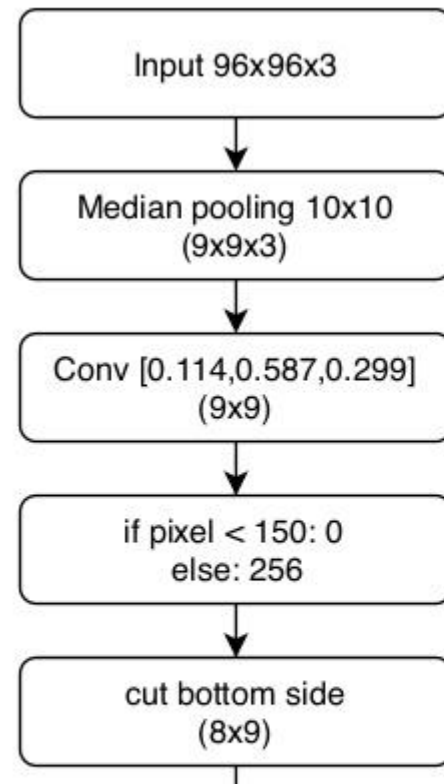
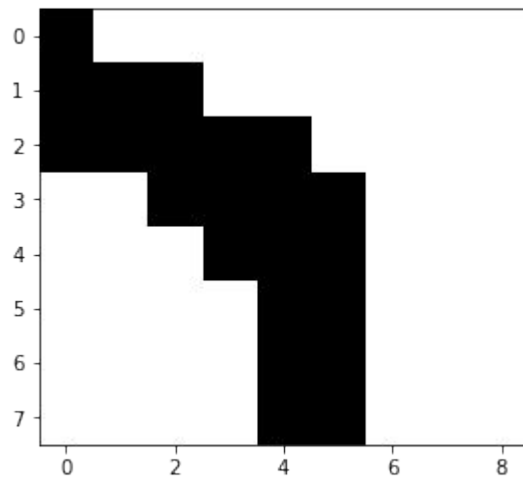
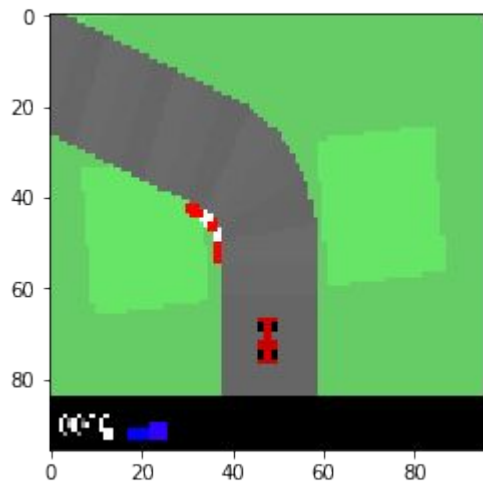
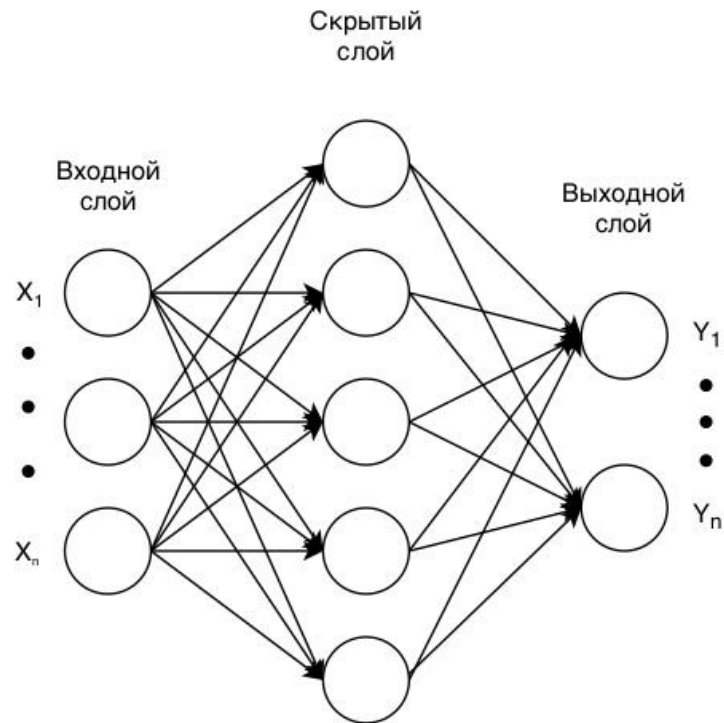
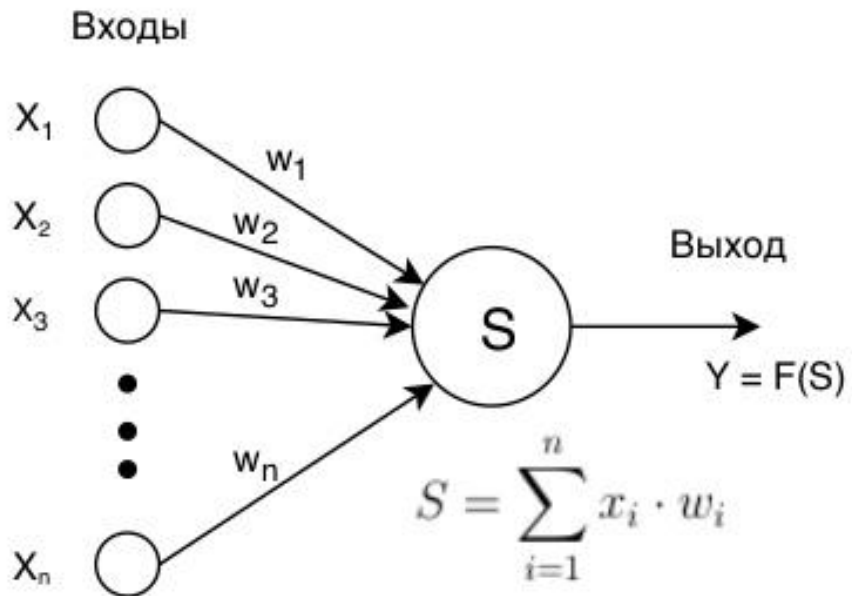


Схема нейросети

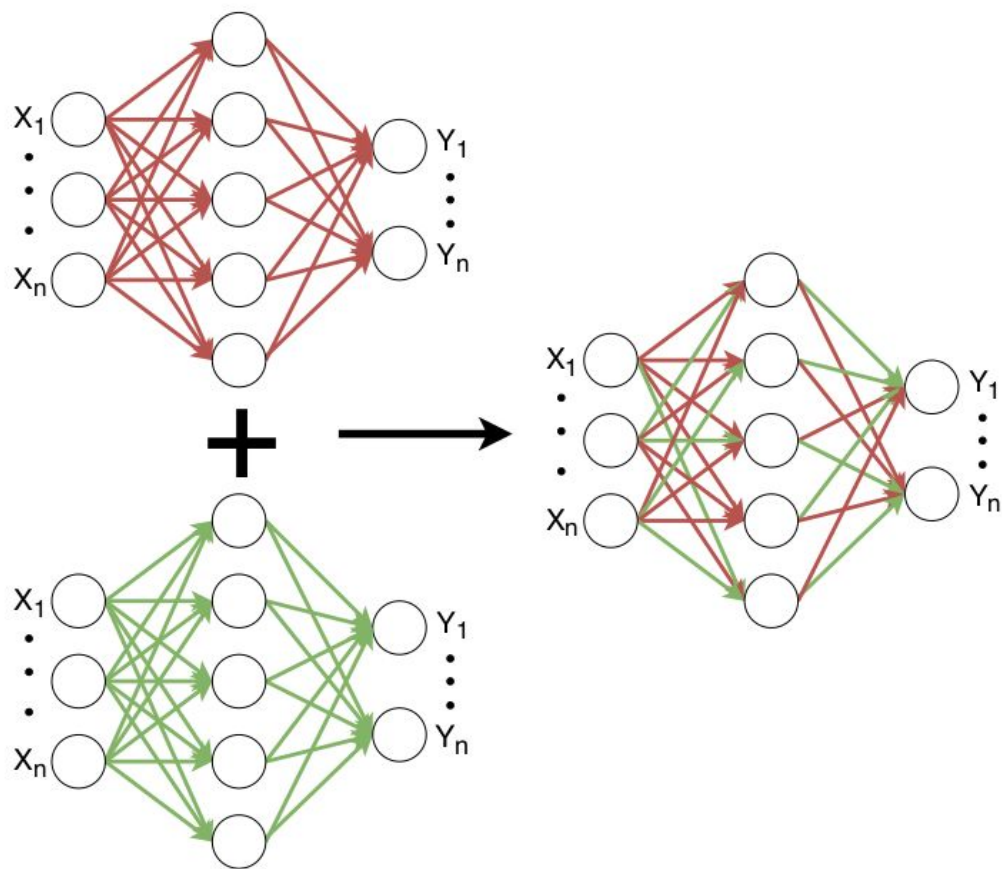


Генетический алгоритм

Нейросеть обучается при помощи генетического алгоритма. В качестве генома выступают веса нейросети

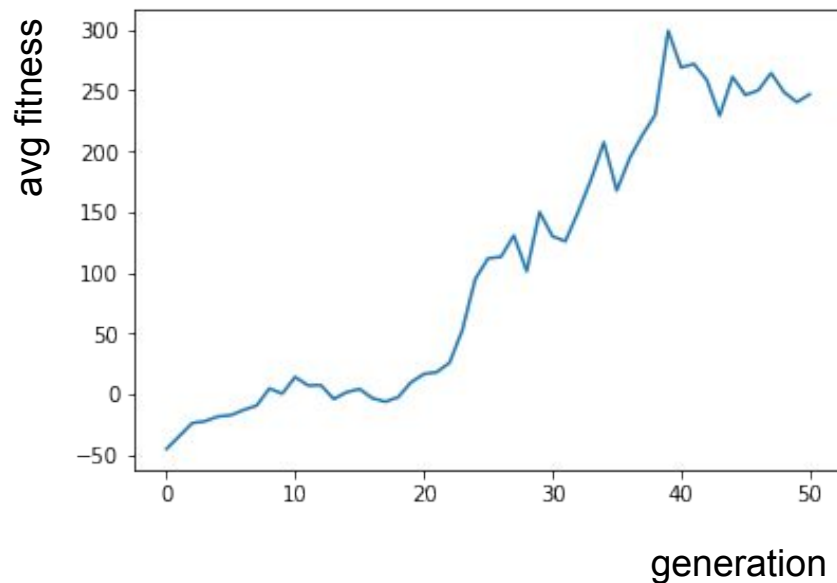
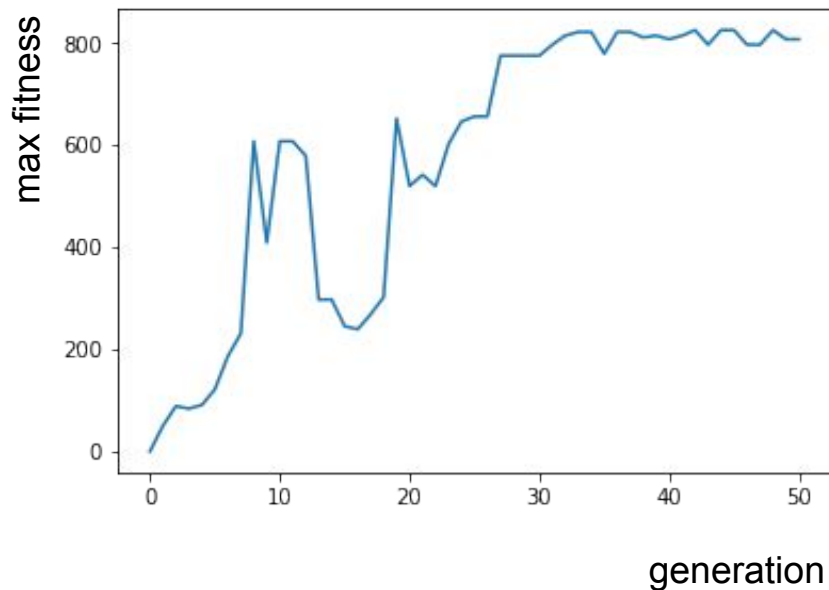
1. Начальная популяция
2. Симуляция и подсчет функции приспособленности (суммарного вознаграждения) для каждого представителя популяции
3. Селекция
4. Формирование нового поколения путем скрещивания и мутации
5. Повторять 2,3,4 пока не сойдемся к максимуму функции приспособленности

Формирование нового поколения



Эксперименты

На графиках представлены максимальное и среднее значение функции приспособленности по всем представителям поколения в зависимости от поколения



Список литературы

1. Pablo Aldape Samuel Sowell “Reinforcement Learning for a Simple Racing Game”
Department of Statistics Stanford University, Department of Electrical
Engineering Stanford University, December 8, 2018
2. Henry Teigar, Miron Storožev, Janar Saks “2D Racing game using reinforcement
learning and supervised learning”
University of Tartu, Institute of Computer Science Neural Networks
3. John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, Oleg Klimov Proximal
“Policy Optimization Algorithms”
OpenAI, 28 Aug 2017, arXiv:1707.06347v2
4. Zhangbo Liu “A Guided Genetic Algorithm for the Planning in Lunar
Lander Games”
Department of Computer Science University of British Columbia
5. “Proximal Policy Optimization”
https://github.com/xtma/pytorch_car_caring

Спасибо за внимание!