**Министерство**

**науки**

**и**

**высшего**

**образования**

**Российской**

**Федерации**

**Федеральное**

**государственное**

**бюджетное**

**образовательное**

**учреждение**

**высшего**

**образования**

**«Московский**

**государственный**

**технический**

**университет**

**имени**

**Н.Э.**

**Баумана**

**(**

**национальный**

**исследовательский**

**университет)»**

**(**

**МГТУ**

**им.**

**Н.Э.**

**Баумана)**



ФАКУЛЬТЕТ Информатика, искусственный интеллект и системы управления

КАФЕДРА Системы обработки информации и управления

**Лабораторная работа №5**

**«Обучение на основе временных различий»**

ИСПОЛНИТЕЛЬ: Ролдугин Е.В.

ФИО

группа ИУ5-25М \_\_\_\_\_\_\_\_\_\_\_

подпись

"\_\_"\_\_\_\_\_\_\_\_\_2023 г.

ПРЕПОДАВАТЕЛЬ: Гапанюк Ю.Е.

ФИО

\_\_\_\_\_\_\_\_\_\_\_\_

подпись

"\_\_"\_\_\_\_\_\_\_\_\_2023 г.

Москва – 2023

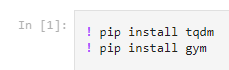
**Задание**

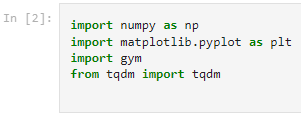
На основе рассмотренного на лекции примера реализуйте следующие алгоритмы:

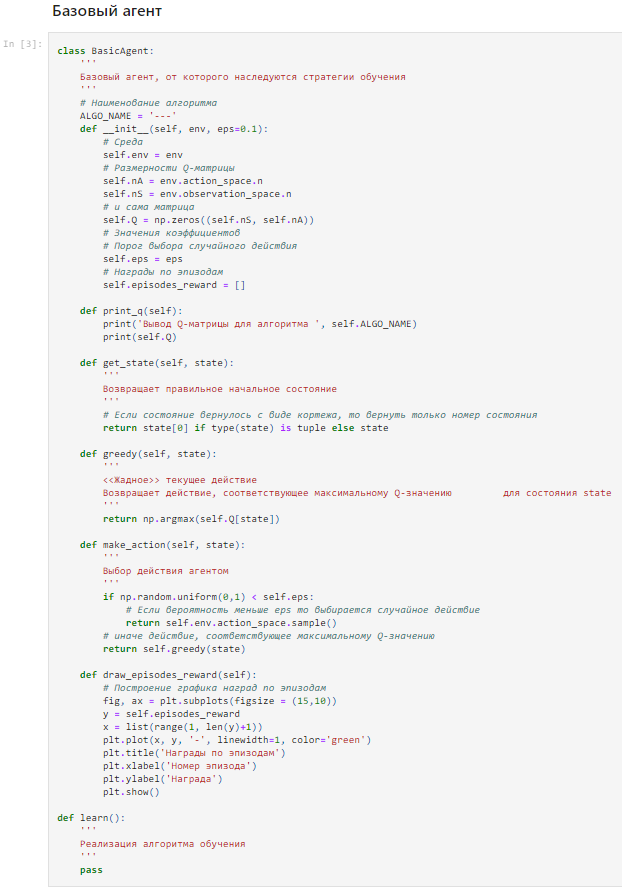
* SARSA
* Q-обучение
* Двойное Q-обучение

для любой среды обучения с подкреплением (кроме рассмотренной на лекции среды Toy Text / Frozen Lake) из библиотеки Gym (или аналогичной библиотеки).

**Решение**

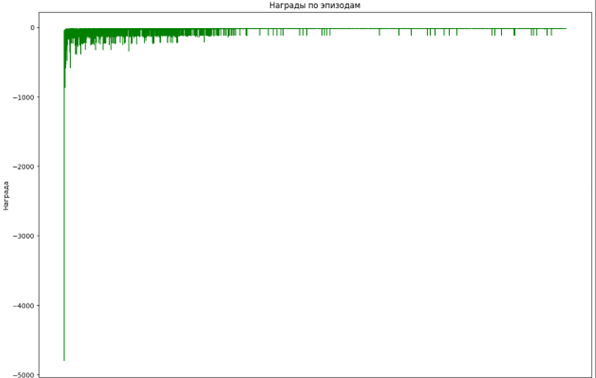






1. SARSA

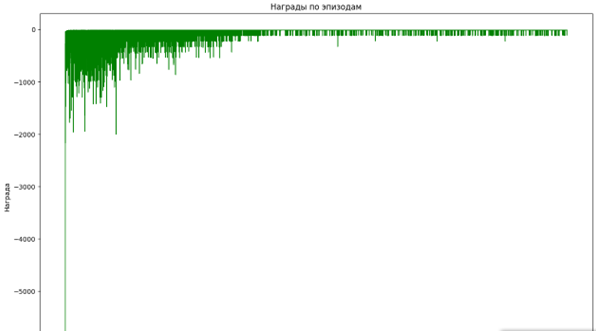






1. Q-обучение





1. Двойное Q-обучение

