**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)**

Курс «Методы машинного обучения»

Отчет по рубежному контролю №2

Выполнила:

студентка группы ИУ5-24М

Мащенко Е. И.

Проверил:

преподаватель каф. ИУ5

Гапанюк Ю.Е.

*2023*

**Задание**

Для одного из алгоритмов временных различий, реализованных Вами в соответствующей лабораторная работе:

* SARSA
* Q-обучение
* Двойное Q-обучение

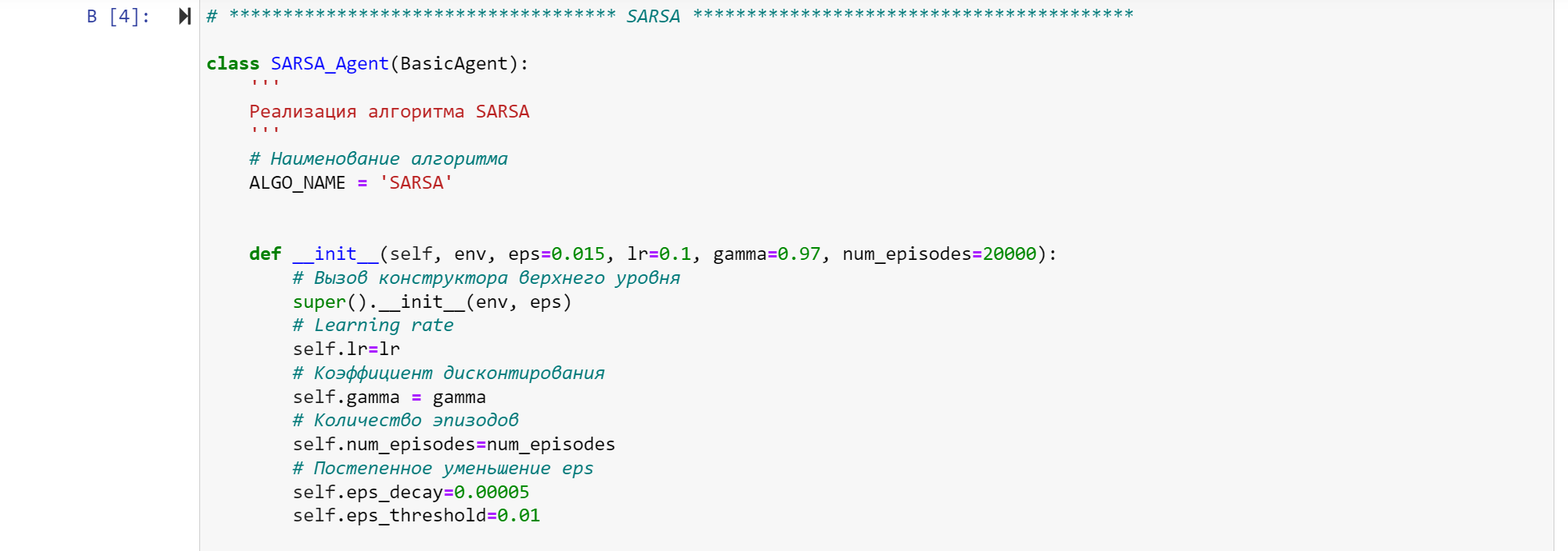
Осуществите подбор гиперпараметров. Критерием оптимизации должна являться суммарная награда.

**Выполнение работы**

Листинг:

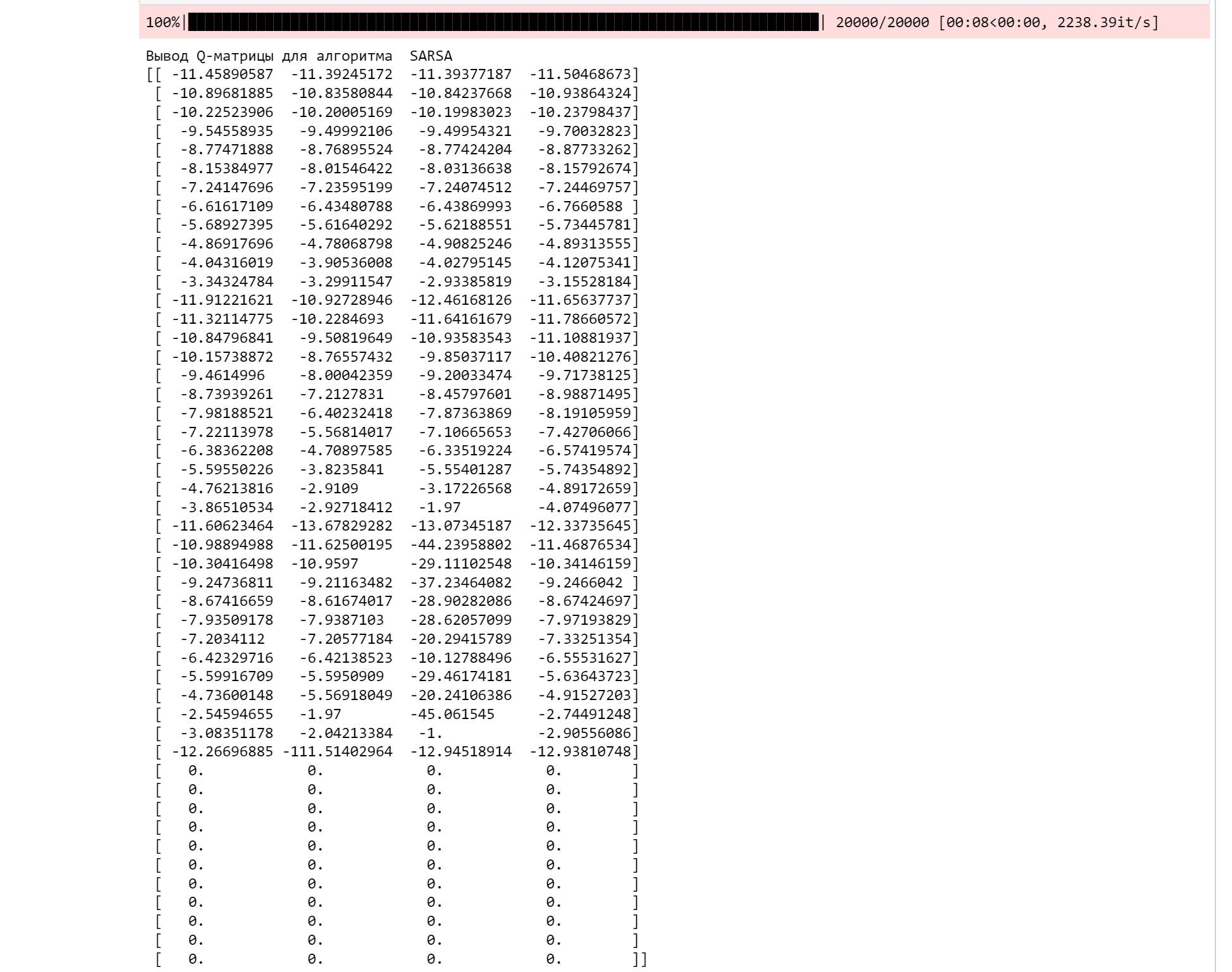


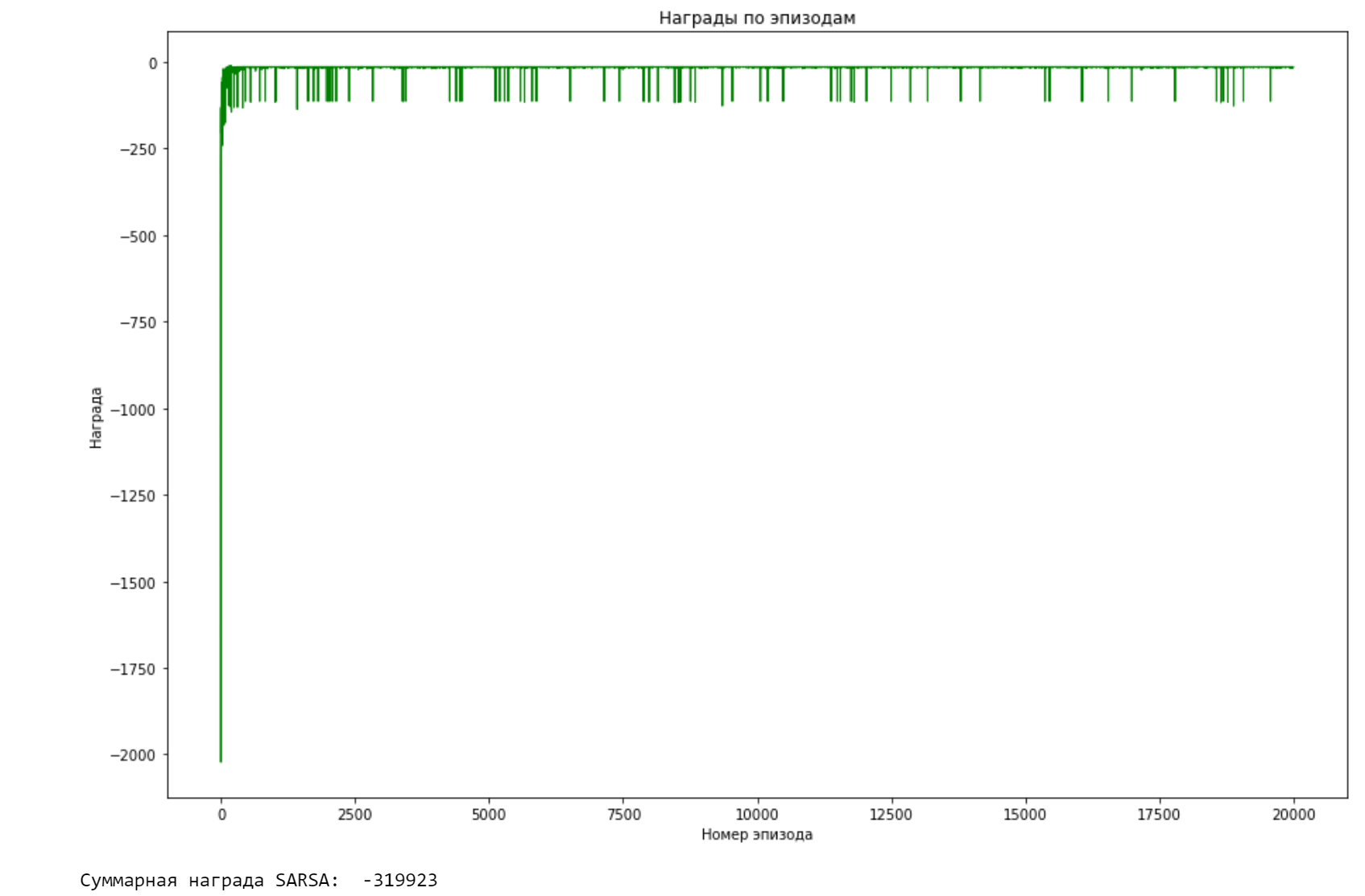










******

Результаты изменения суммарной награды от скорости обучения lr представлены в Таблице 1. График зависимости суммарной награды от lr представлен на рис.1.

Таблица 1. Результаты изменения суммарной награды от скорости обучения lr

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Суммарная награда | Награда за последний эпизод | lr | eps | gamma | episodes |
| -529192 | -17 | 0,025 | 0,4 | 0,98 | 20000 |
| -499310 | -17 | 0,05 | 0,4 | 0,98 | 20000 |
| -484401 | -17 | 0,1 | 0,4 | 0,98 | 20000 |
| -488352 | -17 | 0,15 | 0,4 | 0,98 | 20000 |
| -488035 | -19 | 0,2 | 0,4 | 0,98 | 20000 |
| -497641 | -17 | 0,25 | 0,4 | 0,98 | 20000 |

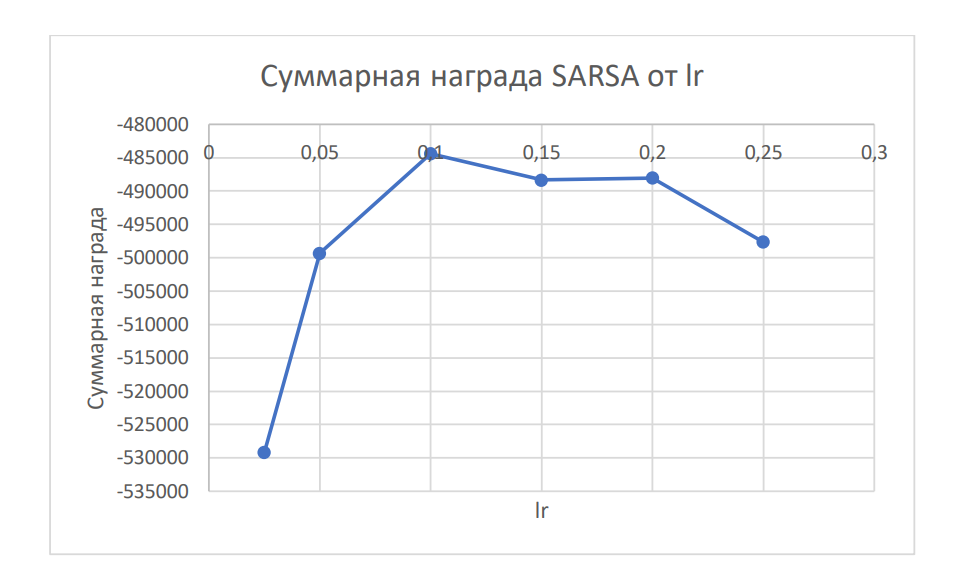


Рис.1. График зависимости суммарной награды от lr

Результаты изменения суммарной награды от параметра eps представлены в Таблице 2. График зависимости суммарной награды от параметра eps представлен на рис.2.

Таблица 2. Результаты изменения суммарной награды от параметра eps

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Суммарная награда | Награда за последний эпизод | lr | eps | gamma | episodes |
| -321854 | -15 | 0,1 | 0,015 | 0,98 | 20000 |
| -322896 | -15 | 0,1 | 0,025 | 0,98 | 20000 |
| -322810 | -15 | 0,1 | 0,05 | 0,98 | 20000 |
| -329450 | -15 | 0,1 | 0,1 | 0,98 | 20000 |
| -351044 | -15 | 0,1 | 0,2 | 0,98 | 20000 |
| -422912 | -17 | 0,1 | 0,3 | 0,98 | 20000 |
| -484401 | -17 | 0,1 | 0,4 | 0,98 | 20000 |
| -586762 | -17 | 0,1 | 0,5 | 0,98 | 20000 |
| -806757 | -17 | 0,1 | 0,6 | 0,98 | 20000 |

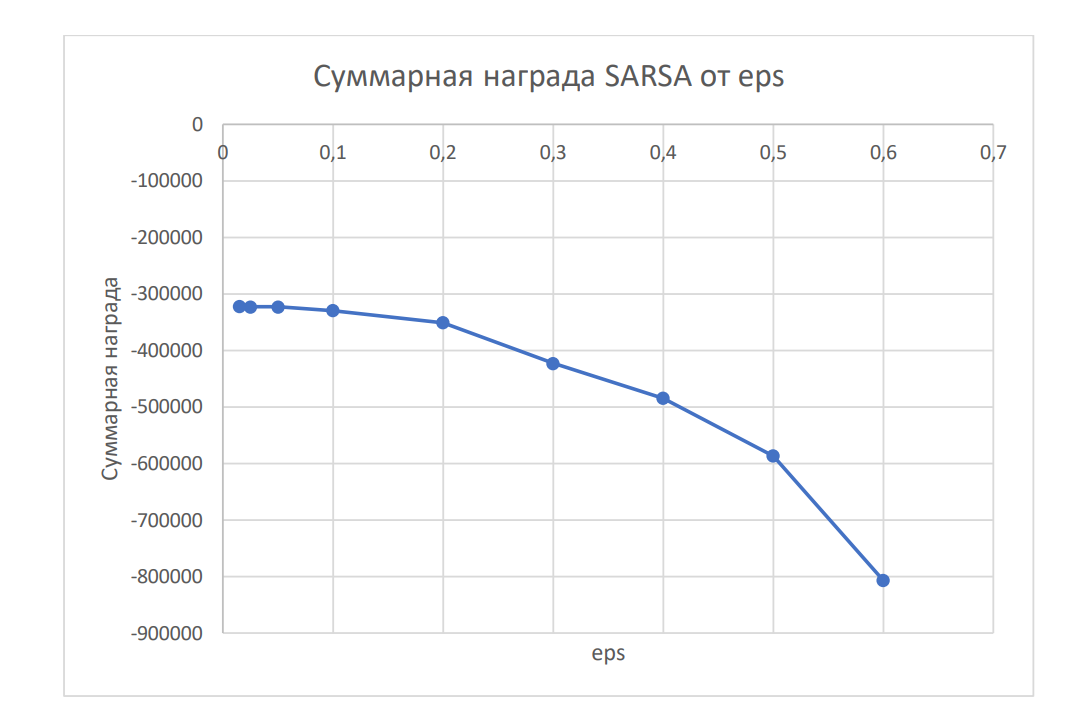


Рис.2. График зависимости суммарной награды от параметра eps

Результаты изменения суммарной награды от параметра gamma представлены в Таблице 3. График зависимости суммарной награды от параметра gamma представлен на рис.3.

Таблица 3. Результаты изменения суммарной награды от параметра gamma

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Суммарная награда | Награда за последний эпизод | lr | eps | gamma | episodes |
| -321112 | -15 | 0,1 | 0,015 | 0,96 | 20000 |
| -319923 | -15 | 0,1 | 0,015 | 0,97 | 20000 |
| -321854 | -15 | 0,1 | 0,015 | 0,98 | 20000 |
| -321521 | -15 | 0,1 | 0,015 | 0,99 | 20000 |
| -321573 | -15 | 0,1 | 0,015 | 0,995 | 20000 |

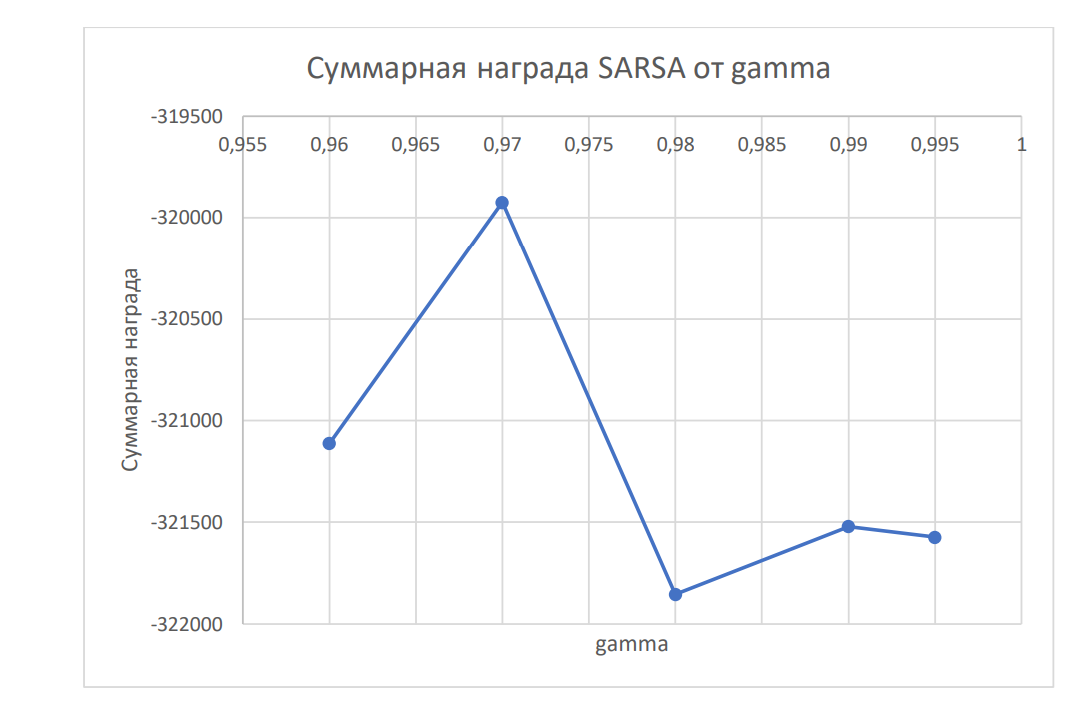


Рис.3. График зависимости суммарной награды от параметра gamma

**Вывод**

В результате подбора гиперпараметров лучшими значениями оказались: eps=0.015, lr=0.1, gamma=0.97, num\_episodes=20000. При этом, при уменьшении eps стратегия агента приближалась к максимальной (к движению по краю обрыва).