

Федеральное государственное автономное  
образовательное учреждение высшего образования  
«Высшая школа экономики»

---

А. И. Панов

# Методы и алгоритмы машинного обучения с подкреплением

*Учебно-методическое пособие*

Москва  
Высшая школа экономики  
2018

В пособии рассмотрены основные

# Оглавление

Оглавление	ii
Введение	1
<b>1 Табличные методы</b>	<b>3</b>
1.1 Марковский процесс принятия решений . . . . .	3
1.2 Динамическое программирование . . . . .	3
1.3 Методы Монте-Карло . . . . .	3
1.4 Q-обучение . . . . .	3
<b>2 Приближенные методы</b>	<b>5</b>
2.1 Предсказание с изменением стратегии . . . . .	5
2.2 Предсказание без изменения стратегии . . . . .	5
2.3 Нейронные сети как аппроксиматоры . . . . .	5
<b>3 Иерархическое обучение с подкреплением</b>	<b>7</b>
3.1 Иерархия действий: Options . . . . .	7
3.2 Иерархия автоматов: НАМ . . . . .	7
3.3 Оптимизация функции оценки: MaxQ . . . . .	7
3.4 Автоматическое формирование иерархий . . . . .	7
<b>4 Обучение с подкреплением и другие науки</b>	<b>9</b>
4.1 Психология . . . . .	9
4.2 Нейрофизиология . . . . .	9
Заключение	11

# Введение

Агент, среда, подкреплением, марковский процесс.



# Глава 1

## Табличные методы

- 1.1 Марковский процесс принятия решений
- 1.2 Динамическое программирование
- 1.3 Методы Монте-Карло
- 1.4 Q-обучение



## Глава 2

# Приближенные методы

- 2.1 Предсказание с изменением стратегии
- 2.2 Предсказание без изменения стратегии
- 2.3 Нейронные сети как аппроксиматоры





## Глава 3

# Иерархическое обучение с подкреплением

3.1 Иерархия действий: Options

3.2 Иерархия автоматов: HAM

3.3 Оптимизация функции оценки: MaxQ

3.4 Автоматическое формирование иерархий



## Глава 4

# Обучение с подкреплением и другие науки

### 4.1 Психология

### 4.2 Нейрофизиология



# Заключение

Немного о целях