



> Конспект > 9 урок > Практическое применение многоруких бандитов

> Оглавление

- > [Оглавление](#)
- > [Контекстуальные бандиты](#)
- > [Дополнительные материалы](#)

Ноутбук с комментариями по лекции находится тут: [скачать](#)

В формате `html`: [скачать](#)

> Контекстуальные бандиты

Отличие от многоруких бандитов: ковариаты доступны на каждой итерации и могут использоваться для выбора ручки.

Проблема: итеративный процесс генерации данных. На каждом раунде мир создаёт наблюдения, состоящие из ковариат X (описание шага и действий) фиксированной размерности.

r — вектор вознаграждений.

m — количество рук.

КУПЛЕНО НА
SKLADCHIK.COM

Агент должен выбрать одну из m рук. Среда возвращает награду **только за выбранную ручку**.

Цель — создать стратегию, которая максимизирует вознаграждение, получаемое агентом.

Ручки могут пропадать и появляться, что приводит к уже знакомой проблеме баланса Exploitation/Exploration.

+ 15
Есть сходство с проблемой классификации, однако нам видны только те метки, которые соответствуют выбранной на определённом шаге ручке.

В реальной жизни такое можно наблюдать, например, **в клинических испытаниях**: мы знаем, как человек отреагировал на определённое лечение, но не знаем ничего о других способах. Ещё один пример — **онлайн реклама**. Мы знаем, что пользователь нажал на баннер, который увидел, но не знаем, нажал ли бы он на какие-то другие.

> **Дополнительные материалы**

1. [contextualbandits](#)
2. [space-bandits](#)