

Skillbox

Введение в машинное обучение

Андрей Мещеряков

Data Scientist

План модуля

- Что такое машинное обучение.
- Чем занимаются инженеры машинного обучения.
- Домашнее задание.

Skillbox

Данные вокруг нас



Data Science и Machine Learning

- **Data Scientist**
Извлекает важную информацию из данных для решения реальных задач и принятия решений.
- **Инженер машинного обучения**
Учит машину самостоятельно принимать решения и обосновывать выводы.

Skillbox

Где используется машинное обучение

Онлайн-маркетинг и
персонализированная
реклама.



Skillbox

Где используется машинное обучение

Искусственная креативность (заготовка для скринкаста)

Skillbox

Где используется машинное обучение

Медицина.



Ограничения машинного обучения

- **Этические ограничения**

Мы доверяем принятию решений алгоритмам больше, чем себе, но что если принятое решение плохо повлияет на человека?

- **Юридические ограничения**

Персональные данные человека должны охраняться и не использоваться для его дискриминации.

Выводы

- Инженеры машинного обучения учат машину самостоятельно принимать и обосновывать решение.
- Данные – это основа машинного обучения.
- Возможности машинного обучения практически не ограничены.

Skillbox

Работа инженера машинного обучения

Андрей Мещеряков

Data Scientist

План урока

- Как инженеры машинного обучения строят работу.
- Создание модели машинного обучения.
- Домашнее задание.

Работа инженера машинного обучения

- Инженеры учат машину самостоятельно принимать решение.
- Для этого инженеры создают модель машинного обучения.
- Модель должна быть качественной.

Как создать качественную модель?

Найти ответы на следующие вопросы:

- Какие задачи будет решать модель?
- Какие данные мы можем использовать?
- Какую именно модель мы будем обучать?
- Как мы проверим модель?

Создание модели машинного обучения

Этап 1: Business Understanding. На этом этапе важно понять:

На этом этапе важно понять:

1 Какую задачу хочет решить заказчик?

3 Как понять, что мы хорошо решили задачу?

2 Как записать задачу на языке машинного обучения?

Skillbox

Создание модели машинного обучения

Этап 1: Business Understanding.

На этом этапе важно понять:

- 1 Какую задачу хочет решить заказчик?
- 2 Как записать задачу на языке машинного обучения?
- 3 Как понять, что мы хорошо решили задачу?

Skillbox

Создание модели машинного обучения



Этап 1: Business Understanding.

На этом этапе важно понять:

1 Какую задачу хочет решить заказчик?

Удалять фотографии, на которых изображены кошки.

2 Как записать задачу на языке машинного обучения?

Выполнять классификацию фотографии на 2 класса: кошек и собак.

3 Как понять, что мы хорошо решили задачу?

Мы удаляем минимальное количество фотографий с собаками.

Этап 2: Data Understanding.

На этом этапе важно понять:

1 Какие данные нужны системе?

Фотографии, поскольку приложение заказчика работает с фотографиями.

3 Сколько нужно фотографий?

1000 фотографий кошек и 1000 фотографий собак.

2 Что должно быть на фотографиях?

Кошки, чтобы знать, кого удалять, и собаки, чтобы знать, кого не удалять.

Skillbox

Этап 3: Data Preparation.

На этом этапе выполняется подготовка данных.

Скринкаст

Skillbox

Этап 4: Modelling.

На этом этапе строится модель.

Скринкаст

Skillbox

Создание модели машинного обучения

99,5%

Точность полученной модели.

Выводы

- Инженеры машинного обучения разбивают свою работу на этапы.
- Этапы можно использовать как чек-листы при подготовке моделей.
- В конце работы модель нужно протестировать.

Skillbox

Домашнее задание



Этапы создания модели машинного обучения

- Определить показатель качества будущей модели, который будет важен заказчику.
- Скачать и распаковать данные заказчика, проверить, что тестовые данные отделены от основных, посмотреть несколько фотографий в основном наборе данных, чтобы убедиться, что заказчик правильно разложил фотографии по папкам.
- Создать на Teachable Machines новый проект по классификации изображений с двумя классами, hot-dog и not-hot-dog. В первый класс добавить изображения хот-догов, во второй — остальной еды.
- Запустить обучение модели, после обучения проверить качество модели на нескольких изображениях из папки test.

Skillbox

**Спасибо за
внимание!**