

Компьютерные методы обработки изображений

Лекция 0

Прошлый семестр

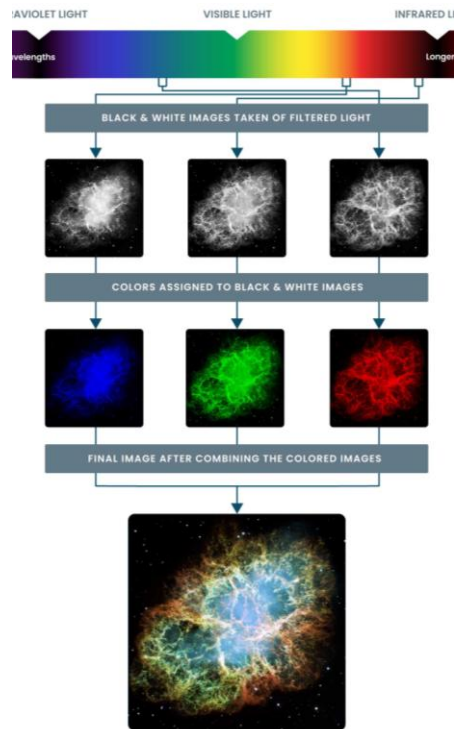
- Элементы теории обработки сигналов
- Получение и регистрация цифровых изображений
- Фильтрация и преобразования изображений
- Цветные изображения
- Сжатие и хранение изображений
- ...



Original image



Transformed image



Noisy Image



Denosed (Default)



Denosed (Calibrated)



В этом семестре

Машинное обучение и глубокое обучение для обработки изображений

Распознавание объектов/образов, более высокоуровневая обработка изображений
(компьютерное зрение)



Classification



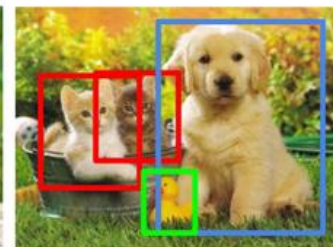
CAT

Classification
+ Localization

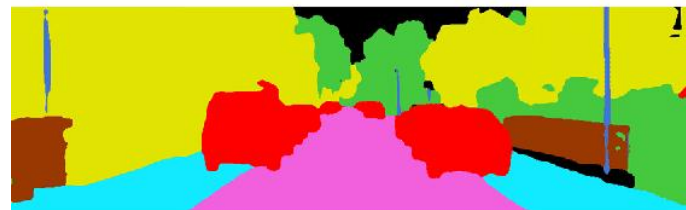


CAT

Object Detection



CAT, DOG, DUCK

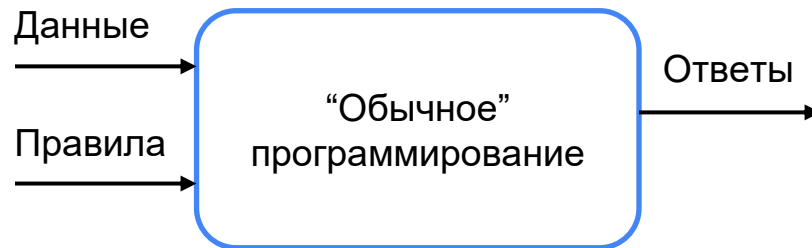


Road	Sidewalk	Building	Fence
Pole	Vegetation	Vehicle	Unlabel

Что такое Машинное Обучение?

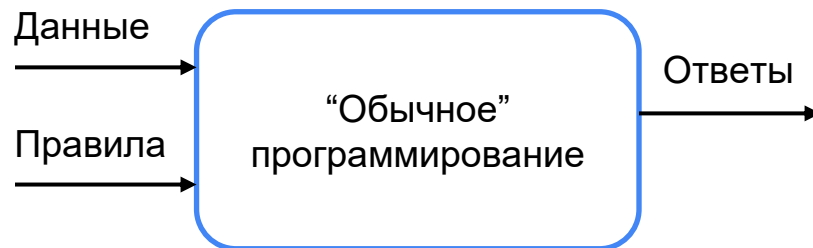
- Существует множество прикладных задач, где накоплено огромное количество данных (информация о пациентах, о клиентах банка, экспериментальные данные и т.д.)
- С этими данными можно, например, строить аналитику: выявлять закономерности, тенденции
- Однако, зачастую требуется решать задачи предсказания, принятия решения (какой диагноз поставить, выдавать ли кредит человеку или нет, предсказать распространение вируса и т.д.)
- **Машинное обучение** - совокупность методов, направленных на построение моделей на основе имеющихся данных.

Что такое Машинное Обучение?

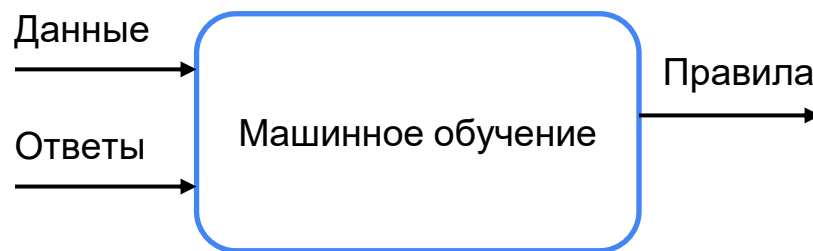


```
x = int(input())  
res = 3*x+2  
print(res)
```

Что такое Машинное Обучение?



```
x = int(input())  
res = 3*x+2  
print(res)
```



Постановка задачи

Обучение с учителем (supervised learning):

Есть входные данные x и
лейблы/метки y

Цель: обучить функцию $f: x \rightarrow y$

Классификация, регрессия, детекция
объектов, семантическая сегментация
и т.д.

Постановка задачи

Обучение с учителем (supervised learning):

Есть входные данные x и
лейблы/метки y

Цель: обучить функцию $f: x \rightarrow y$

Классификация, регрессия, детекция
объектов, семантическая сегментация
и т.д.

Обучение без учителя (unsupervised learning):

Есть только входные данные x

Цель: найти некоторые особенности,
структуру данных

Кластеризация, понижение
размерности и т.д.

Постановка задачи

Для начала рассмотрим обучение с учителем

Обучение с учителем (supervised learning):

Есть входные данные x и
лейблы/метки y

Цель: обучить функцию $f: x \rightarrow y$

Классификация, регрессия, детекция
объектов, семантическая сегментация
и т.д.

Обучение без учителя (unsupervised learning):

Есть только входные данные x

Цель: найти некоторые особенности,
структуру данных

Кластеризация, понижение
размерности и т.д.

Виды признаков при работе с табличными данными

Описывать объекты можно с помощью некоторых характеристик - **признаков** (features). Вектор признаков будем называть признаковым описанием.

Какие бывают признаки?

- Бинарные (1/0, болен/здоров)
- Вещественные (возраст, площадь квартиры)
- Категориальные (цвет глаз, гражданство)
- Порядковые (тип населенного пункта, место на пьедестале, образование)
- Множественные признаки (в какие игры играл пользователь)

Виды признаков при работе с табличными данными

В основном будем работать с табличными данными.
Ниже приведена часть данных из Titanic датасета.

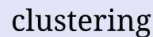
Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	
PassengerId											
1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

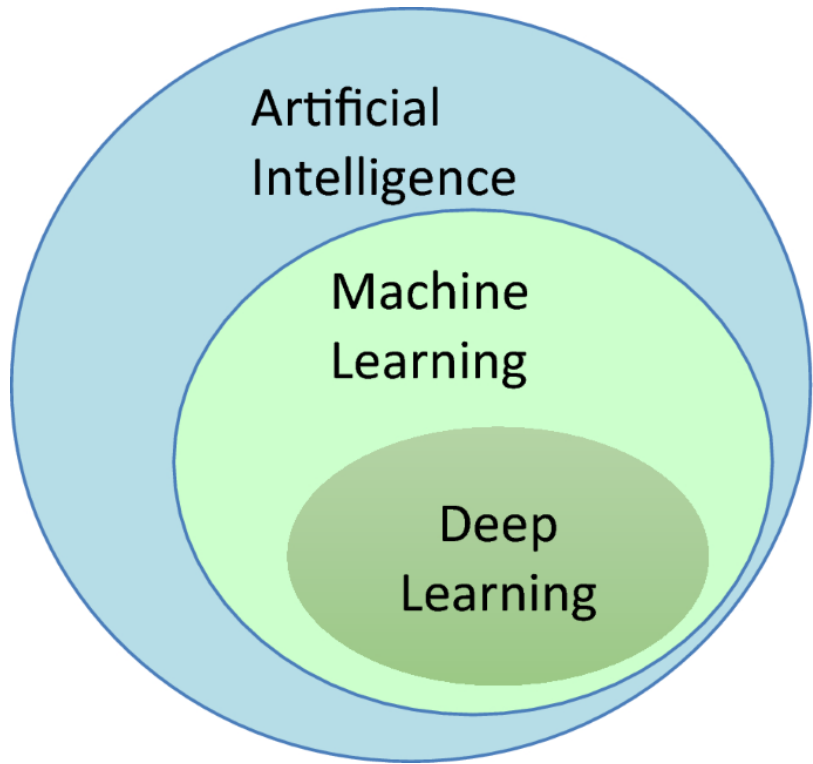
Виды ответов (target value)

- **Бинарная классификация** $Y = \{0, 1\}$ (вернет ли клиент кредит или нет)
- **Многоклассовая классификация** $Y = \{1, \dots, K\}$ (определение предметной области статьи)
- **Многоклассовая с пересекающимися классами** $Y = \{0, 1\}^M$ (задача диагностики, у одного пациента несколько заболеваний)
- **Задача регрессии.** Ответы - действительные числа. (инженерные и научные приложения, оценка стоимости недвижимости)

Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
PassengerId	Y	X								
1	0	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
2	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
3	1	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
4	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
5	0	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

classification

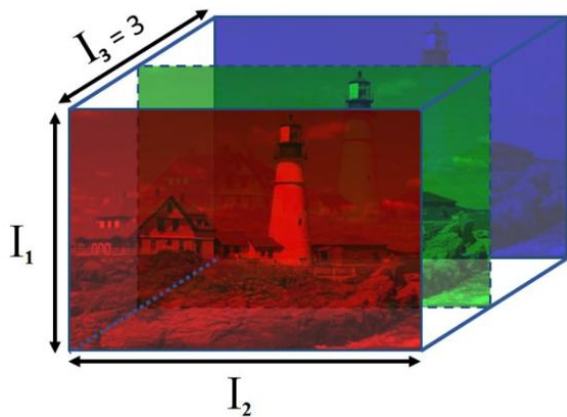




Применительно к обработке изображений мы будем большое внимание уделять глубокому обучению

Компьютерное зрение (Computer Vision)

Чаще всего изображения представляют собой тензор размерности (H, W, C) . $C = 3$ для RGB, $C = 1$ для grayscale



Представление
изображения в виде
тензора 3-го порядка



RGB (слева) и grayscale
(справа) изображения

Компьютерное зрение: задачи

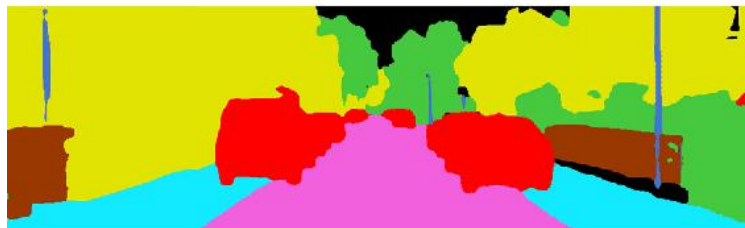
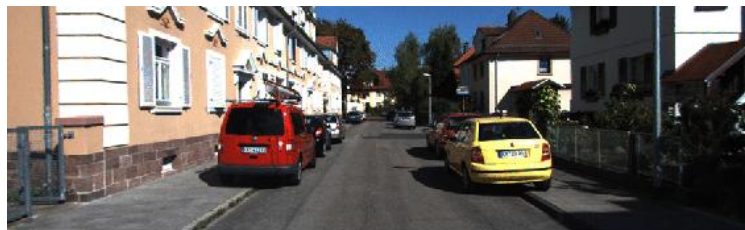
Классификация (classification). Необходимо правильно определить, к какому классу принадлежит изображение


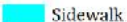








Изображения из датасета ImageNet с предсказанными классами

Компьютерное зрение: задачи

Семантическая сегментация (semantic segmentation). Присваивание лейбла/класса определенным областям изображения.

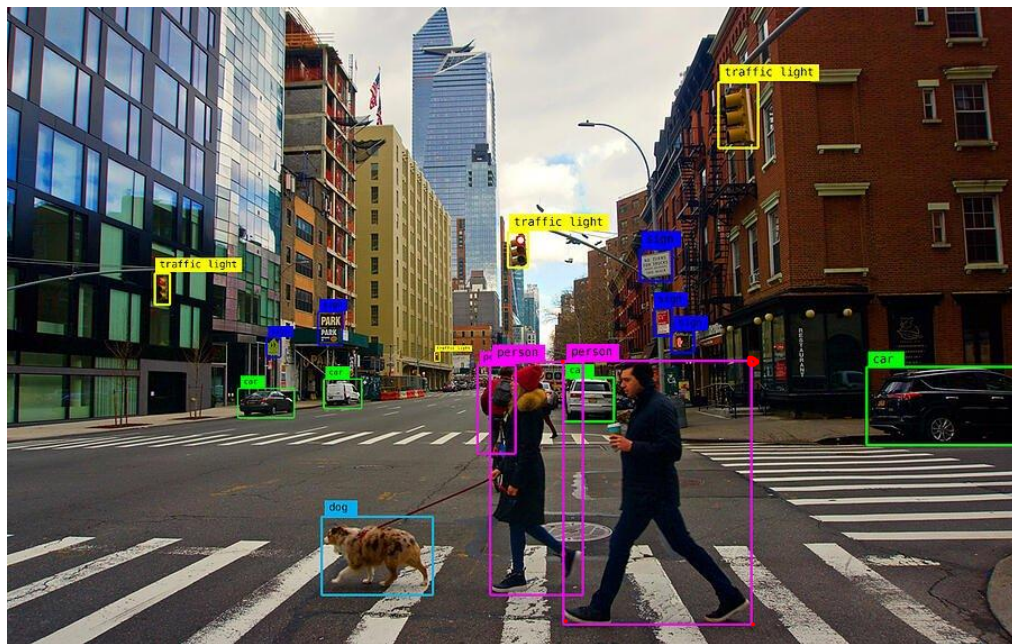


 Road	 Sidewalk	 Building	 Fence
 Pole	 Vegetation	 Vehicle	 Unlabel

Пример семантической сегментации. Разным областям на картинке присвоен свой класс.

Компьютерное зрение: задачи

Детекция объектов (object detection). Необходимо найти объект на картинке и присвоить ему класс



При детекции объект выделяется алгоритмом с помощью бокса (bounding box) с соответствующим лейблом.

Компьютерное зрение: задачи

Перенос стиля (style transfer). Применение визуального стиля картины на исходное изображение.



Компьютерное зрение: задачи

Глубокие порождающие (генеративные) модели (Deep Generative Models).

Относятся к парадигме обучения без учителя; могут генерировать новые данные



Запрос (**prompt**), который был передан в модель:

The interior of a colossal starship's command center, panoramic views of cosmic vistas through a transparent dome, holographic star maps, crew members operating high-tech consoles, blue and white aesthetics::2 interstellar navigation, artificial gravity, warp drive controls, serene spacefaring environment --ar 16:9 --v 6.0

Текстовые данные: задачи

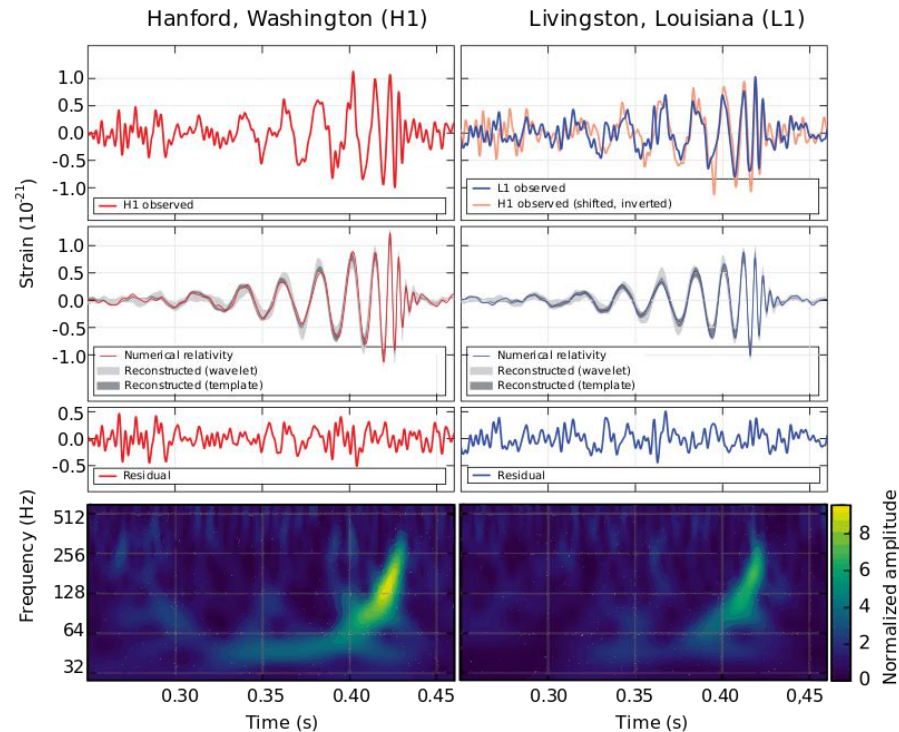
Название данной области - Обработка естественного языка (**Natural Language Processing, NLP**)

- Классификация текста
- Машинный перевод (machine translation)
- Генерация текста (text generation)
- Ответы на вопросы (question answering)
- Чат-боты
- Определение части речи (part-of-speech tagging, POS)
- Анализ эмоциональной окраски (sentiment analysis)



Обработка сигналов/временных рядов (time series): задачи

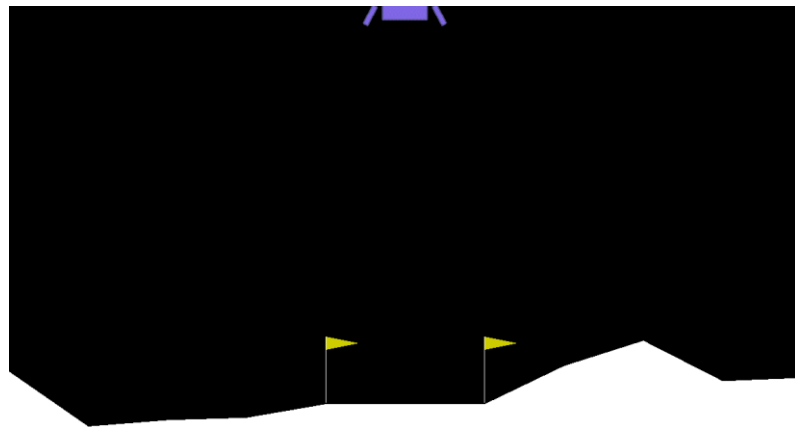
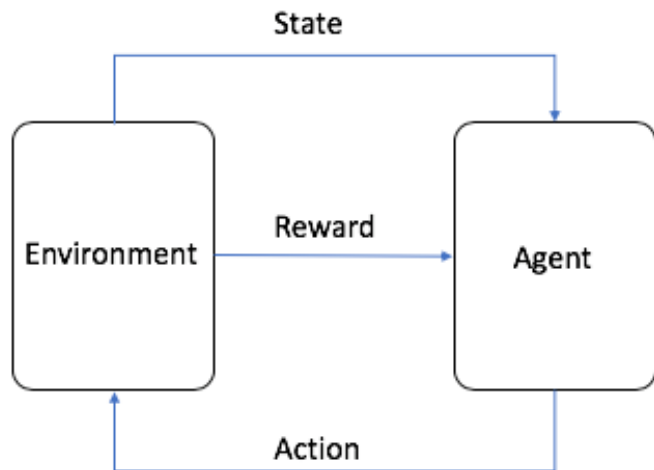
- Распознавание речи (automatic speech recognition, speech-to-text)
- Синтез речи
- Генерация музыки (music generation)
- Устранение шумов (denoising)
- Обнаружение сигнала



Сигнал с двух интерферометров и соответствующие спектрограммы

Обучение с подкреплением (Reinforcement Learning)

Метод машинного обучения, в котором система (**агент**) взаимодействует со средой (**environment**) и получает награду (**reward**) за выполненные действия.



В этом примере задача заключается в
нахождении оптимального способа
посадки лунохода