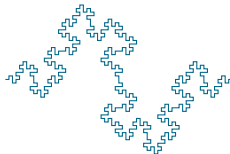


Введение в машинное обучение

Зуева Надежда

Московский Физико-Технический Институт



4 июля 2018 г.

План

Вводная часть

Историческая справка

Терминология и основные типы задач

 Что такое нейронная сеть и нейрон

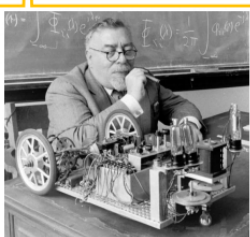
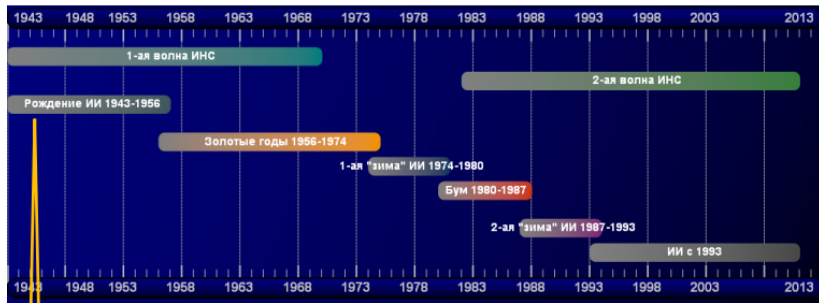
 Напоминания

 Терминология

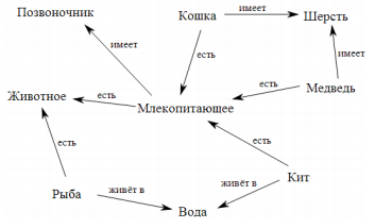
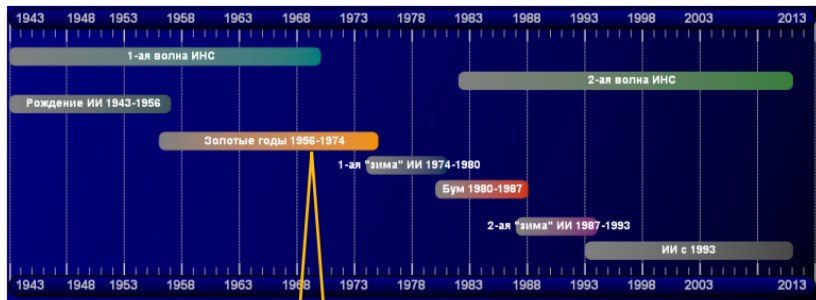
 Типы задач

Основные достижения и перспективы развития

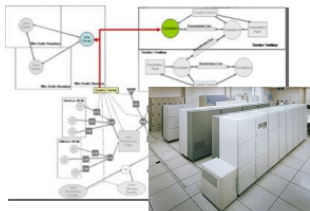
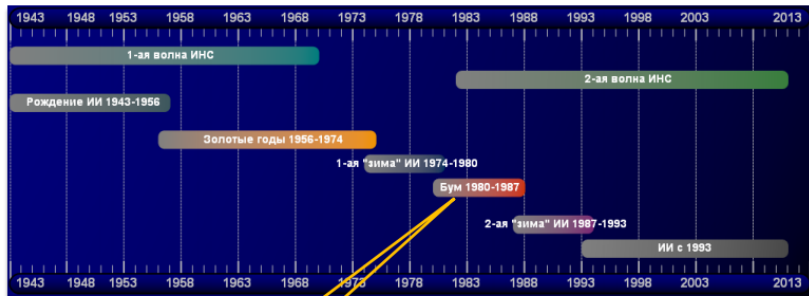
Рождение ИИ



Развитие ИИ

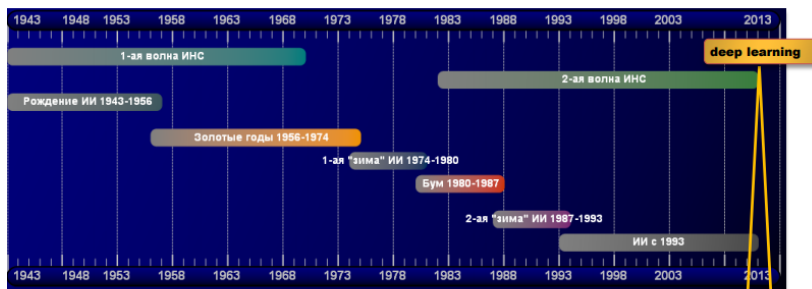


"Бум"

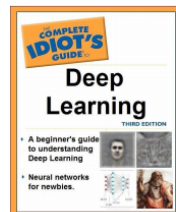
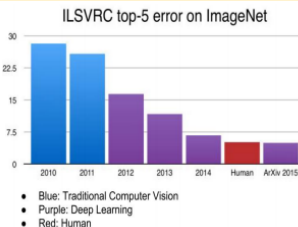


- Попытки смоделировать рассуждения человека в узкой предметной области – появление экспертных систем – систем основанных на знаниях
- Японское правительство выделяет \$850 млн. на программу «Компьютеры пятого поколения»

Нейронные сети

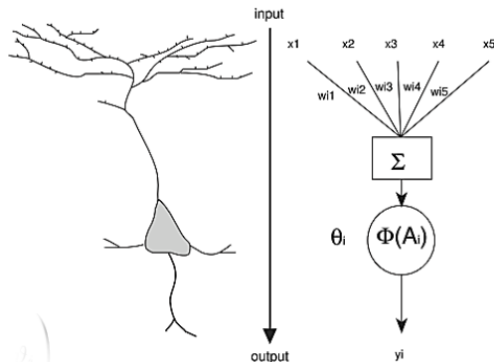


- Классификация огромных массивов изображений и видео.
- Google и Facebook охотятся за головами ученых занимающихся нейронными сетями.
- Нейросети учатся играть в игры по картинке.



Искусственный нейрон

1. Активность нейрона определяется преобразованием взвешенного суммарного воздействия на него
2. Воздействия могут быть положительными или отрицательными (активирующими либо тормозящими)



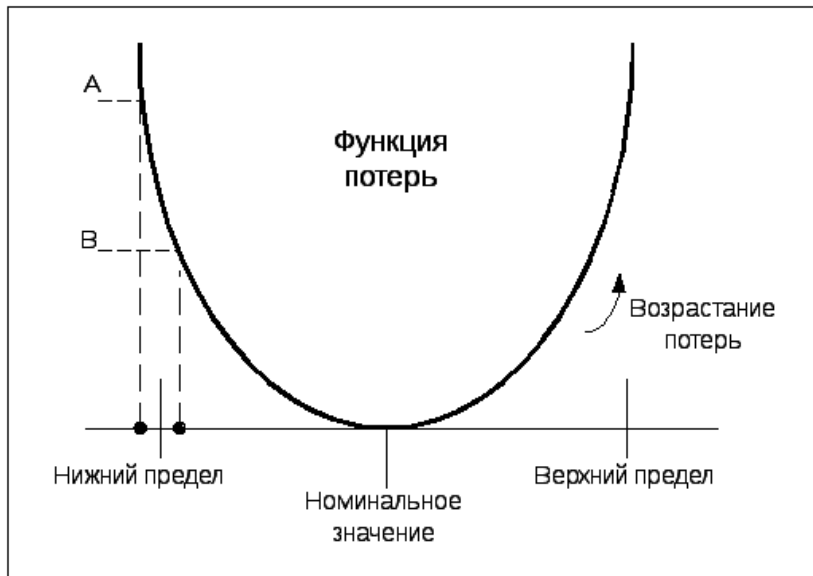
$$a_i = \sum_{j=1}^N w_{ij} x_j$$

$$y_i = \Phi(a_i) = \Phi \left(\sum_{j=1}^N w_{ij} x_j - \vartheta_i \right)$$

Функция потерь

Функция потерь — функция, которая в теории статистических решений характеризует потери при неправильном принятии решений на основе наблюдаемых данных. Если решается задача оценки параметра сигнала на фоне помех, то функция потерь является мерой расхождения между истинным значением оцениваемого параметра и оценкой параметра.

Loss

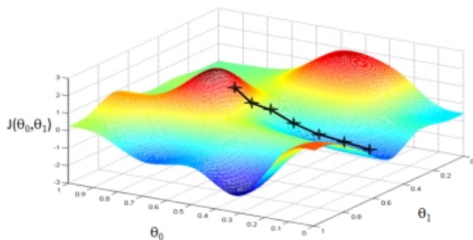
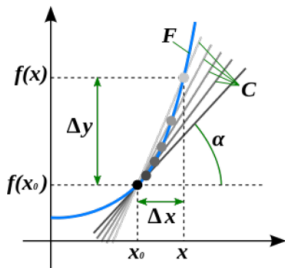


Производная

Производная функция — понятие дифференциального исчисления, характеризующее скорость изменения функции в данной точке.

Определяется как предел отношения приращения функции к приращению её аргумента при стремлении приращения аргумента к нулю, если такой предел существует.

Если производная равна нулю в некоторой точке, то эта точка — **экстремум** (локальный или глобальный максимум или минимум функции). $f'(x_0) = \frac{df}{dx}(x_0)$



Линейная алгебра

Вектор — в линейной алгебре вектором называется элемент линейного пространства. Векторы могут иметь различную природу: направленные отрезки, матрицы, числа, функции и другие, однако все линейные пространства одной размерности изоморфны между собой.

Матрица — математический объект, записываемый в виде прямоугольной таблицы элементов

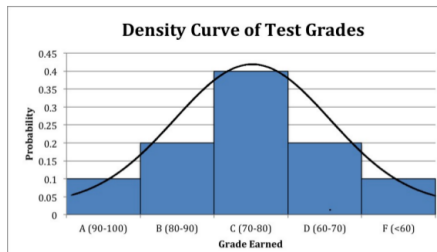
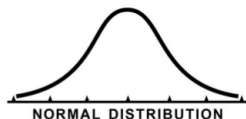
$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Теория вероятностей

Случайная величина — это переменная, значения которой представляют собой исходы какого-нибудь случайного феномена или эксперимента.

Выборка — набор независимых между собой случайных величин

Распределение вероятностей — это закон, описывающий область значений случайной величины и вероятности их исхода (появления).



Основные понятия

X — множество **объектов**

Y — множество **допустимых ответов**

y^* — целевая функция, $y^*: X \rightarrow Y$, $y_i = y^*(x_i)$ известны только на **конечном** подмножестве объектов x_1, \dots, x_m из X

Пары (x_i, y_i) — прецеденты

Совокупность пар таких пар при i из $1, \dots, m$ — **обучающая выборка** (X_{train})

a — **решающая функция** (алгоритм), которая любому объекту из X ставит в соответствие допустимый ответ из Y и приближает целевую функцию y^*

X_{test} — **выборка прецедентов** для тестирования построенного алгоритма a

Для решения задачи обучения по прецедентам в первую очередь фиксируется восстанавливаемой зависимости.

Основные понятия

Признак (feature) f объекта x — это результат измерения некоторой характеристики объекта. Формально признаком называется отображение $f : X \rightarrow D_f$, где D_f — множество допустимых значений признака. В частности, любой алгоритм $a : X \rightarrow Y$ также можно рассматривать как признак

Пусть дан набор признаков $f_1(x), \dots, f_n(x)$.

Признаковое описание объекта x — вектор (одномерный массив) (f_1, \dots, f_n) . Совокупность признаковых описаний всех объектов выборки длины m , записанную в виде таблицы размера mn , называют матрицей объектов–признаков.

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ≡ ≡ ≡ ↺ 🔍 ↻

Кредитный скоринг

Задание

Данные — информация о выданных кредитах, требуется предсказать вероятность успешного погашения кредита.

X_{train} содержится в файле train.csv, X_{test} — test.csv.

Информация о значениях признаков содержится в файле featureDescr.csv

Целевой признак — $loan_{status}$, бинарный признак. 1 означает, что кредит успешно погашен.

Задача

Предсказать, кому стоит выдавать кредит?

Рекомендательная система



4 620 р

Автомобильная
шина MICHELIN...



16 р

Форма для кулича
Жостовская фабри...



4 000 р

Матрас Аскона
Balance Forma...



405 р

Matrix шампунь Total
Results So Long...



9 999 р

Диван Hoff Хаген



6 430 р

Кроватка
Mimi 7 в 1

Женская парфюмерия

Все товары



amp;
Dolce



1 506 р

LACOSTE Lacoste
pour Femme



2 745 р

Christian Dior J'adore
Eau de Parfum



2 145 р

Dolce & Gabbana 3...



1 468 р

Versace Bright Crystal

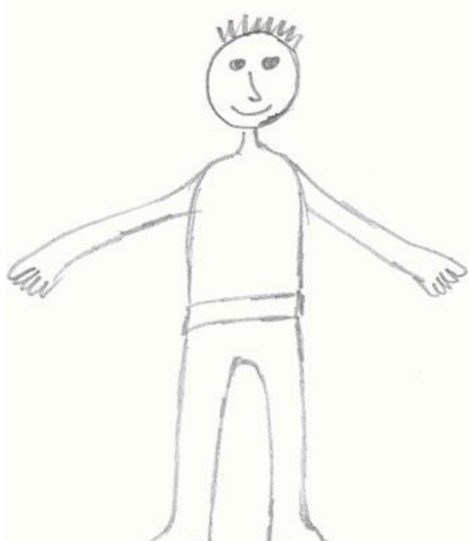


3 620 р

Guerlain Mon Guer...

Пример

Знакомьтесь — это **Вася**. Кем он может быть в нашей терминологии? Какие признаки могут быть у Васи?

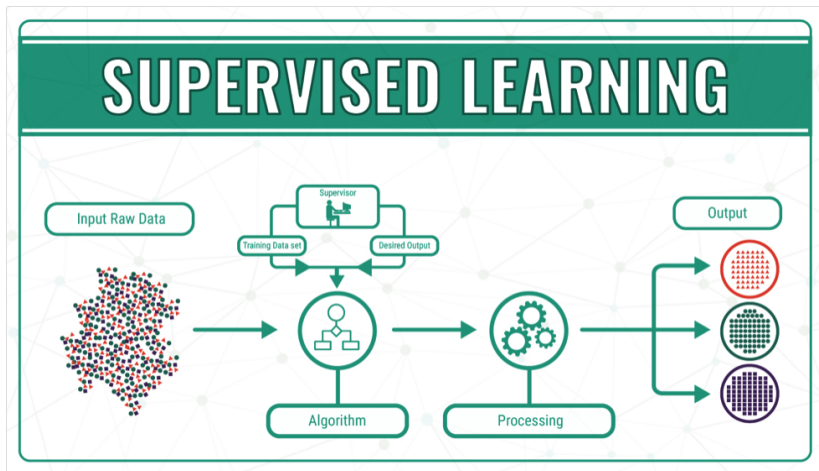


Задача обучения по прецедентам

По выборке X_{train} построить решающую функцию (*decisionfunction*) $a : X \rightarrow Y$, которая приближает целевую функцию y^* , причём не только на объектах **обучающей выборки, но и на всём множестве X** .

Решающая функция a должна быть вычислимой.

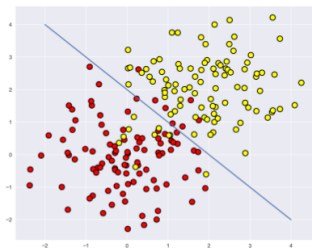
Задача обучения по прецедентам



Классификация

Классификация

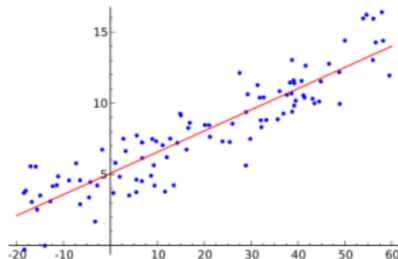
Множество допустимых ответов конечно. Их называют метками классов (class label). Класс — это множество всех объектов с данным значением метки.



Регрессия

Регрессия

Отличается тем, что допустимым ответом является действительное число или числовой вектор.



Типы задач по методу обучения

1 Обучение с учителем

Каждый прецедент представляет собой пару «объект, ответ».

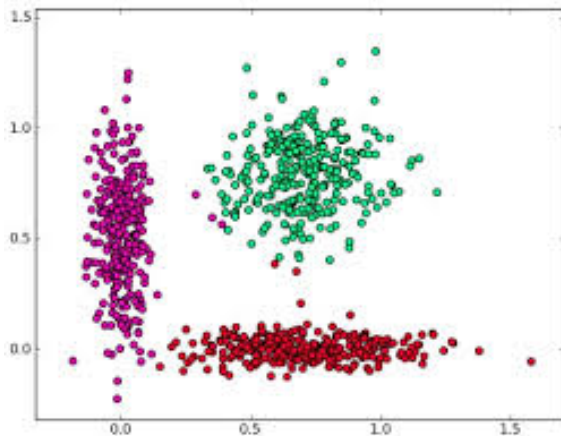
Требуется найти функциональную зависимость ответов от описаний объектов и построить алгоритм, принимающий на входе описание объекта и выдающий на выходе ответ.

Функционал качества обычно определяется как средняя ошибка ответов, выданных алгоритмом, по всем объектам выборки.

2 Обучение без учителя

В этом случае ответы не задаются, и требуется искать зависимости между объектами.

Обучение без учителя



Обучение без учителя

Результатом является перекрашивание исходных пикселей, где каждому пикселю присваивается цвет его ближайшего центра кластера

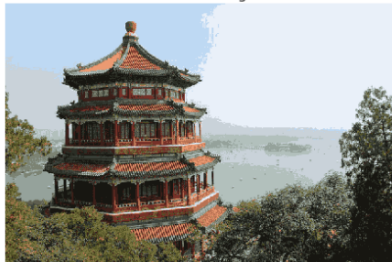
```
china_recolored = new_colors.reshape(china.shape)

fig, ax = plt.subplots(1, 2, figsize=(16, 6),
                        subplot_kw=dict(xticks=[], yticks=[]))
fig.subplots_adjust(wspace=0.05)
ax[0].imshow(china)
ax[0].set_title('Original Image', size=16)
ax[1].imshow(china_recolored)
ax[1].set_title('16-color Image', size=16);
```

Original Image



16-color Image



Google Neural Machine Translation

Почти год назад Google анонсировала запуск новой модели для Google Translate.

Основной результат: сокращение отставания от человека по точности перевода на 55—85 процентов (оценивали люди по 6-балльной шкале). Воспроизвести высокие результаты этой модели сложно без огромного датасета, который имеется у Google.

Бот для переговоров от Facebook

Вы могли слышать дурацкие новости о том, что Facebook выключила своего чат-бота, который вышел из-под контроля и выдумал свой язык. Этого чат-бота компания создала для переговоров. Его цель — вести текстовые переговоры с другим агентом и достичь сделки: как разделить на двоих предметы (книги, шляпы...). У каждого агента своя цель в переговорах, другой ее не знает. Просто уйти с переговоров без сделки нельзя.

Бот научился одной из стратегий реальных переговоров — показывать поддельный интерес к некоторым аспектам сделки, чтобы потом по ним уступить, получив выгоду по своим настоящим целям.

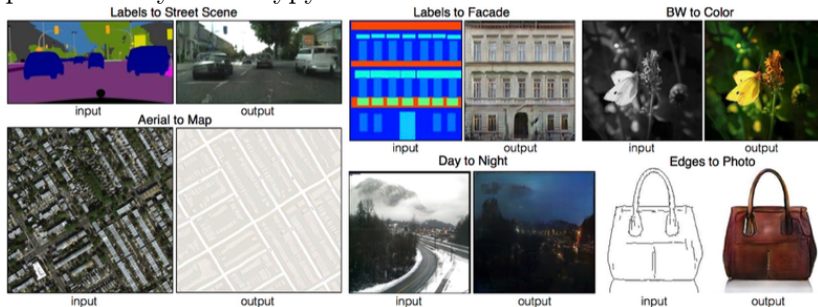
Чтение по губам от DeepMind

Google Deepmind создали модель, которая была обучена на телевизионном датасете и смогла превзойти профессионального lips reader'а с канала BBC.

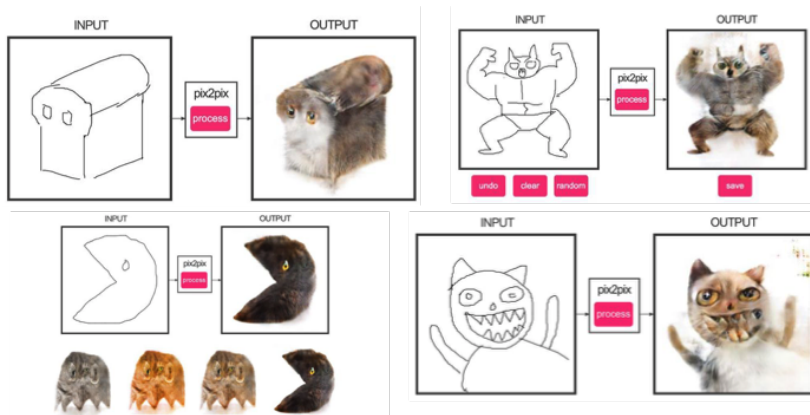


pix2pix

Исследователи решали проблему image-to-image генерирования, когда, например, требуется по снимку со спутника создать карту или по наброску предметов — их реалистичную текстуру.



pix2pix



Материалы

1. Терминология введена в соответствии с пособием К.В. Воронцова "Обучение по прецедентам"
2. Использованы материалы лаборатории IPavlov
3. Использованы материалы с ресурса Nabrahabr