

Краткое введение в машинное обучение

2 сентября 2017 г.

При изучении наук примеры полезнее правил
Исаак Ньютон

Нет царских путей к геометрии
Евклид

1 Задача оптимизации

2 Задача классификации

3 Состояние современной науки

Задача про домики

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

Небольшая компания производит домики для кошек. Фиксированные издержки составляют 20 тыс. руб. в месяц. Переменные издержки составляют 3 тыс. руб. на каждый проданный домик. В первый месяц по цене 6 тыс. руб. за домик было продано 100 домиков. Во второй месяц была установлена цена 8 тыс. руб. за домик, и не было продано ни одного домика. В предположении, что спрос линейно зависит от цены, определите оптимальную цену и объем продаж.

Задача про домики

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

Пусть x – количество проданных домиков, Q – прибыль компании

$$c(x) = 20 + 3x$$

$$x = 50(8 - p)$$

$$p(x) = 8 - 0.02x$$

$$Q(x) = p(x)x - c(x)$$

$$Q(x) \rightarrow \max$$

Производная и скорость

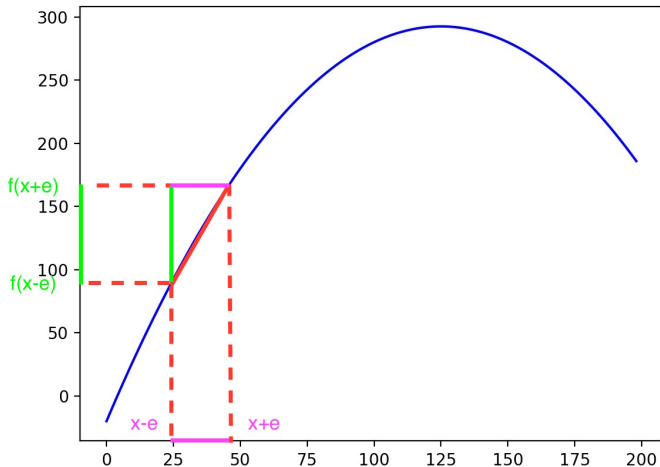
Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки



Производная и скорость

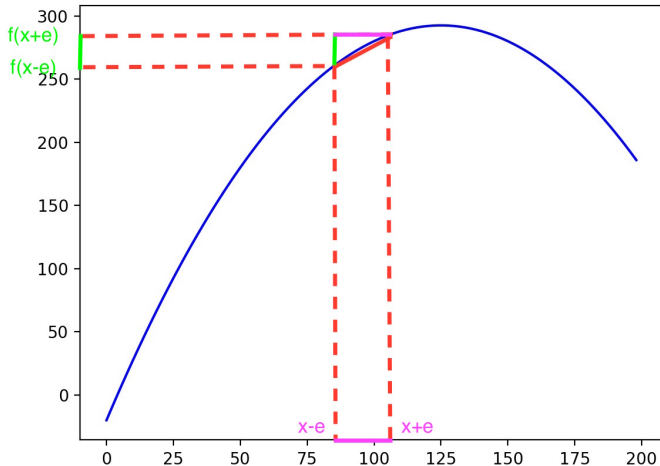
Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки



Градиентный спуск

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

$L(x)$ – функция потерь (Loss function)

$$L(x) \rightarrow \min$$

update step: $x = x - \alpha \nabla L(x)$

$$\nabla L(x) = \frac{L(x + \varepsilon) - L(x - \varepsilon)}{2\varepsilon}$$

Минимизация функции двух переменных

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

$L(x_1, x_2)$ – функция потерь

$$L(x_1, x_2) \rightarrow \min$$

$$x_1 = x_1 - \alpha \nabla_1 L(x_1, x_2)$$

$$x_2 = x_2 - \alpha \nabla_2 L(x_1, x_2)$$

$$\nabla_1 L(x_1, x_2) = \frac{L(x_1 + \varepsilon, x_2) - L(x_1 - \varepsilon, x_2)}{2\varepsilon}$$

$$\nabla_2 L(x_1, x_2) = \frac{L(x_1, x_2 + \varepsilon) - L(x_1, x_2 - \varepsilon)}{2\varepsilon}$$

Минимизация функции многих переменных

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

$$L(\mathbf{x}) \rightarrow \min$$

$$\text{update step: } \mathbf{x} = \mathbf{x} - \alpha \nabla L(\mathbf{x})$$

$$\nabla_i L(\mathbf{x}) = \frac{L(\mathbf{x}_+) - L(\mathbf{x}_-)}{2\varepsilon}$$

$$\mathbf{x}_+ = [x_1, x_2, \dots, x_i + \varepsilon, \dots, x_n]$$

$$\mathbf{x}_- = [x_1, x_2, \dots, x_i - \varepsilon, \dots, x_n]$$

Суммы и вектора

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

Знак суммы

$$\sum_{i=1}^n a_i = a_1 + a_2 + a_3 + \dots + a_n$$

$$\sum_i a_i = a_1 + a_2 + a_3 + \dots + a_n$$

Вектор – набор пронумерованных чисел (список чисел).

$\mathbf{x} = [x_1, x_2]$ – двумерный вектор (список из двух чисел).

Пример 1: Дан трехмерный вектор $\mathbf{w} = [2.3, 3.1, 7.0]$.

Рассчитать $\sum_i w_i$

Пример 2: $\mathbf{f} = [0.0, 1.0, 0.0]$. Рассчитать $\sum_i w_i f_i$

Классификация

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

\mathbf{f}_k – вектора признаков (features) объектов (samples), y_k – номера классов объектов, \mathbf{w} – параметры модели.

Функция потерь:

$$L(\mathbf{w}) = \sum_k L_k(\mathbf{w}) \rightarrow \min$$

Функция потерь на k -м объекте:

$$L_k(\mathbf{w}) = |g(\mathbf{w}, \mathbf{f}_k) - y_k|$$

$y = g(\mathbf{w}, \mathbf{f})$ – классифицирующая функция.

Стохастический градиентный спуск

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

Функция потерь:

$$L(\mathbf{w}) = \sum_k L_k(\mathbf{w}) \rightarrow \min$$

Будем обновлять параметры модели отдельно для каждого объекта (примера), а не для всей суммы сразу

$$\mathbf{w} = \mathbf{w} - \alpha \nabla L_k(\mathbf{w})$$

Линейная функция

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

\mathbf{w} , b – коэффициенты линейной функции

$$f(\mathbf{x}) = b + \sum_i w_i x_i$$

Двумерный случай

$$f(\mathbf{x}) = b + \sum_{i=0}^1 w_i x_i = b + w_0 x_0 + w_1 x_1$$

Сигмоид

Краткое
введение в
машинное
обучение

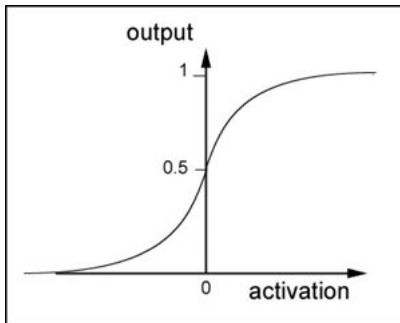
План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

$$\sigma(x) = \frac{e^x}{e^x + 1}$$



Персептрон

Краткое
введение в
машинное
обучение

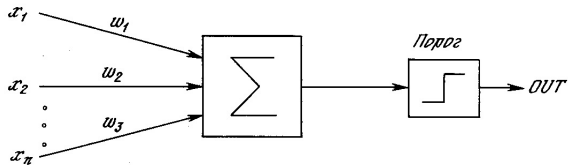
План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

$$g(\mathbf{x}) = \sigma(b + \sum_i w_i x_i)$$



Персептрон

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

$$g(\mathbf{x}) = \sigma(b + \sum_i w_i x_i)$$

Уравнение

$$\begin{aligned} g(\mathbf{x}) &= 0.5 \\ b + \sum_i w_i x_i &= 0 \end{aligned}$$

задает поверхность, разделяющую классы.

Как было раньше

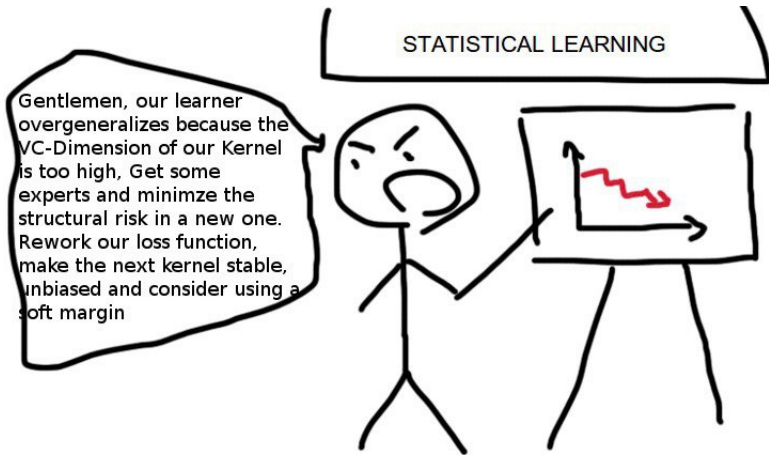
Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки



Как теперь

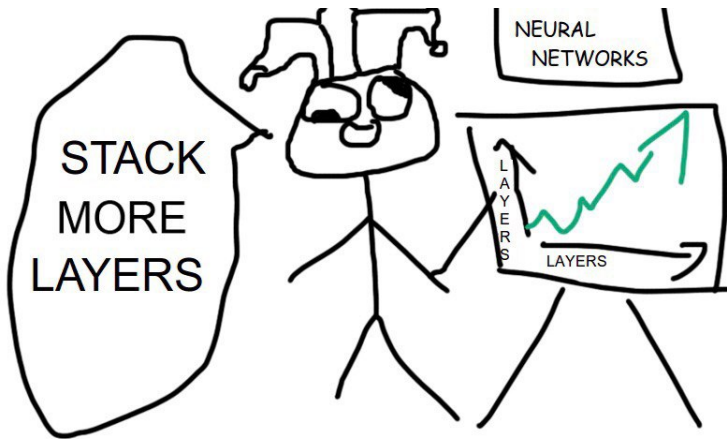
Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки



⁰<https://vk.com/weirdkerneltricks>

Нейронные сети и распознавание изображений

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

- 1936 Ирисы Фишера
- 1943 Модель МакКаллока–Питтса
- 1958 Нейронная сеть Марка Розенблатта
- 1969 Марвин Минский «Персептроны»
- 1995 Ян Ле Кун – LeNet и конволюционные нейронные сети (<http://yann.lecun.com/exdb/lenet/>)
- ImageNet – лучший результат с SVM 26%
- 2012 AlexNet 15%
- 2014 VGG-16 7.5%
- 2014 GoogLeNet 6.7%
- 2015 Inception-v3 4.2%
- 2015 ResNet 3.6%

Inception-v3

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки



¹Inception 5 (GoogLeNet)



Inception 7a

¹Going Deeper with Convolutions, [C. Szegedy et al, CVPR 2015]

Другие крупные области

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

Обучение с учителем (Supervised Learning)

- Распознавание звука
- Предсказание временных рядов
- Обработка естественного языка
- Экспертные системы

Обучение без учителя, кластеризация

- LSA
- 2008 Collobert and Weston – Word Embeddings
- 2013 Mikolov – Word2Vec

Обучение с подкреплением (Reinforcement Learning)

- Поведение в реальном мире
- 1990-1995 Шашки – Chinook
- 1995 Нарды – TD-Gammon
- 1996 Шахматы – Deep Blue
- 2011 Jeopardy – Watson
- 2015 Го – AlphaGo

Пакеты

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

- SciKit Learn <http://scikit-learn.org/stable/index.html>
- NLTK
- TensorFlow
https://www.tensorflow.org/get_started/get_started
- Keras
- Torch <http://torch.ch/>

Блоги и статьи

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

- <https://habrahabr.ru/post/303196/>
- <http://karpathy.github.io/2015/03/30/breaking-convnets/>
- <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>
- <http://colah.github.io/posts/2014-03-NN-Manifolds-Topology/>
- <http://colah.github.io/>

Лекции и ресурсы онлайн

Краткое
введение в
машинное
обучение

План

Задача
оптимизации

Задача
классификации

Состояние
современной
науки

- <https://www.coursera.org/learn/machine-learning>
- <https://www.kaggle.com/>
- ШАД и Воронцов – не стоит)

Что потребуется из матана?

- Производные, градиент, производная сложной функции
- Векторные пространства и матрицы – самое начало
- Основы статистики – вероятность, распределение вероятности для непрерывных переменных