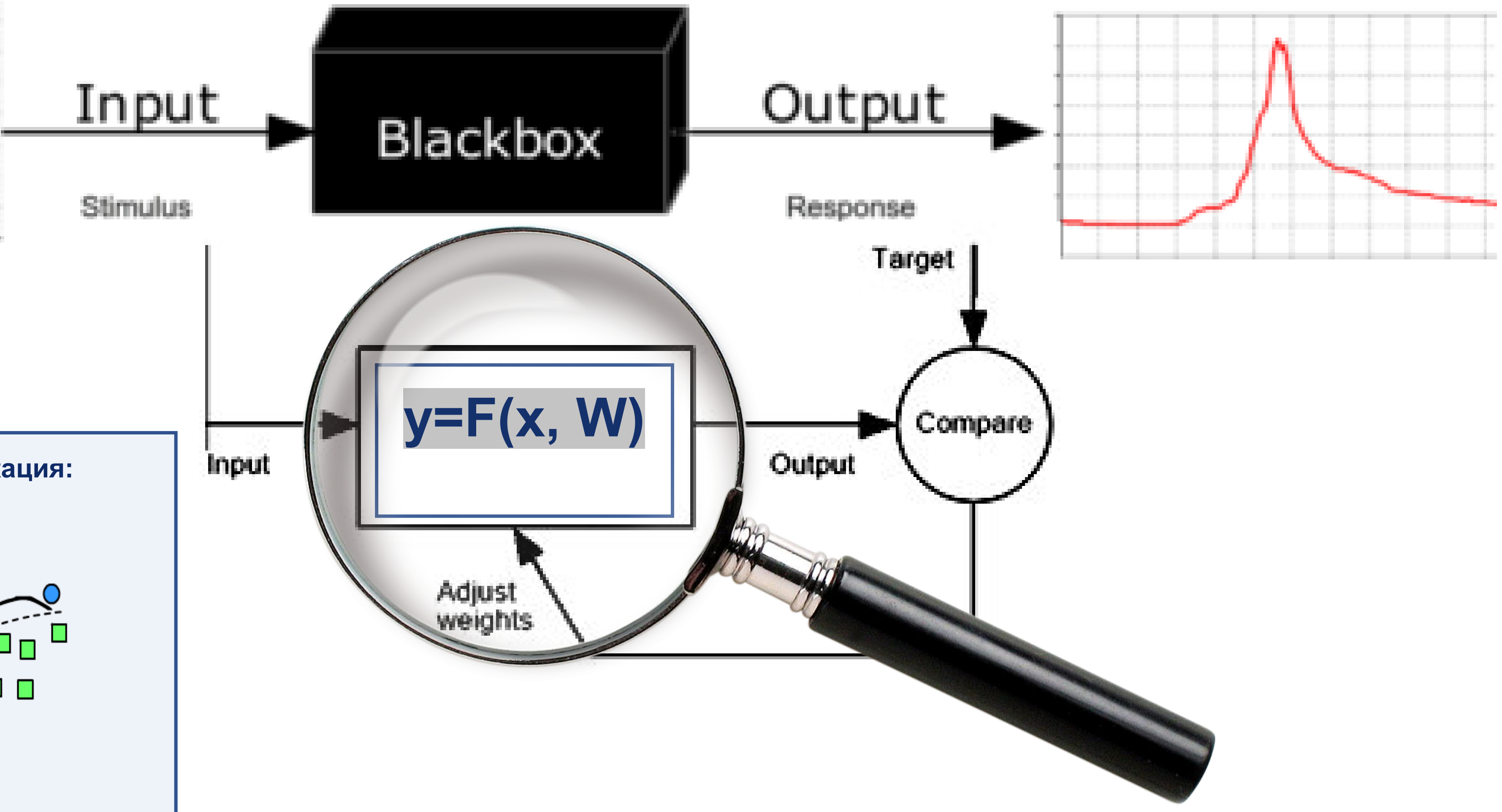
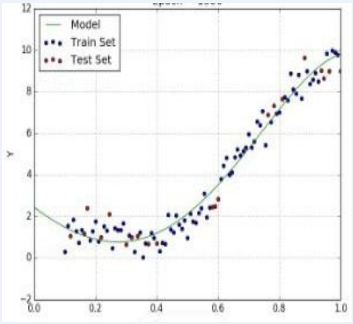
Лекция 05
Линейные модели

А. Линейная регрессия
Б. Регуляризация, Ridge, Lasso и ElasticNet
В. Логистическая регрессия, LogLoss
Г. Машины опорных векторов

«Черный ящик»



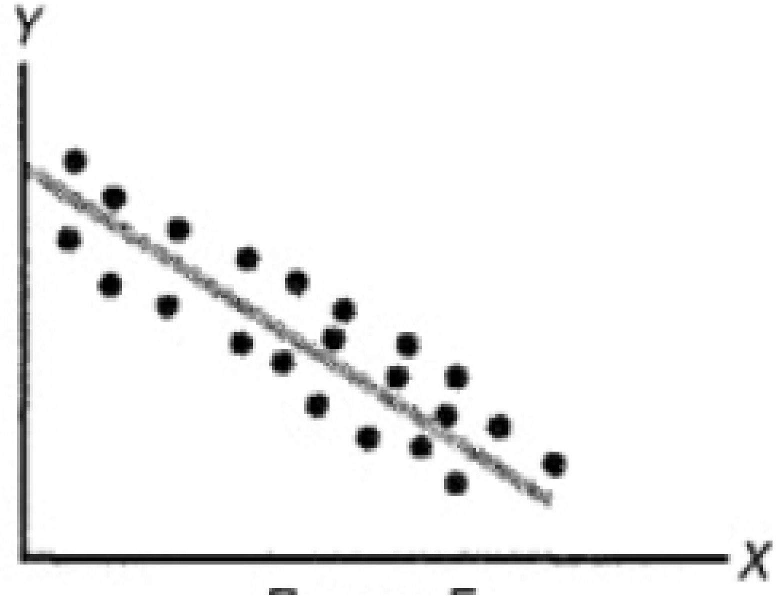
Регрессия:
непрерывные
значения



Классификация:
дискретные
значения

Линейная регрессия

- движение тел по инерции
- экономические расчеты, ценообразование
- расчеты металлоконструкций, тепловое удлинение рельс
- …



$$y = a * x + b$$

$$\vec{y} = A * \vec{x} + \vec{b}$$

$\vec{x}$ – вектор размерности N

$\vec{y}$ – вектор размерности M

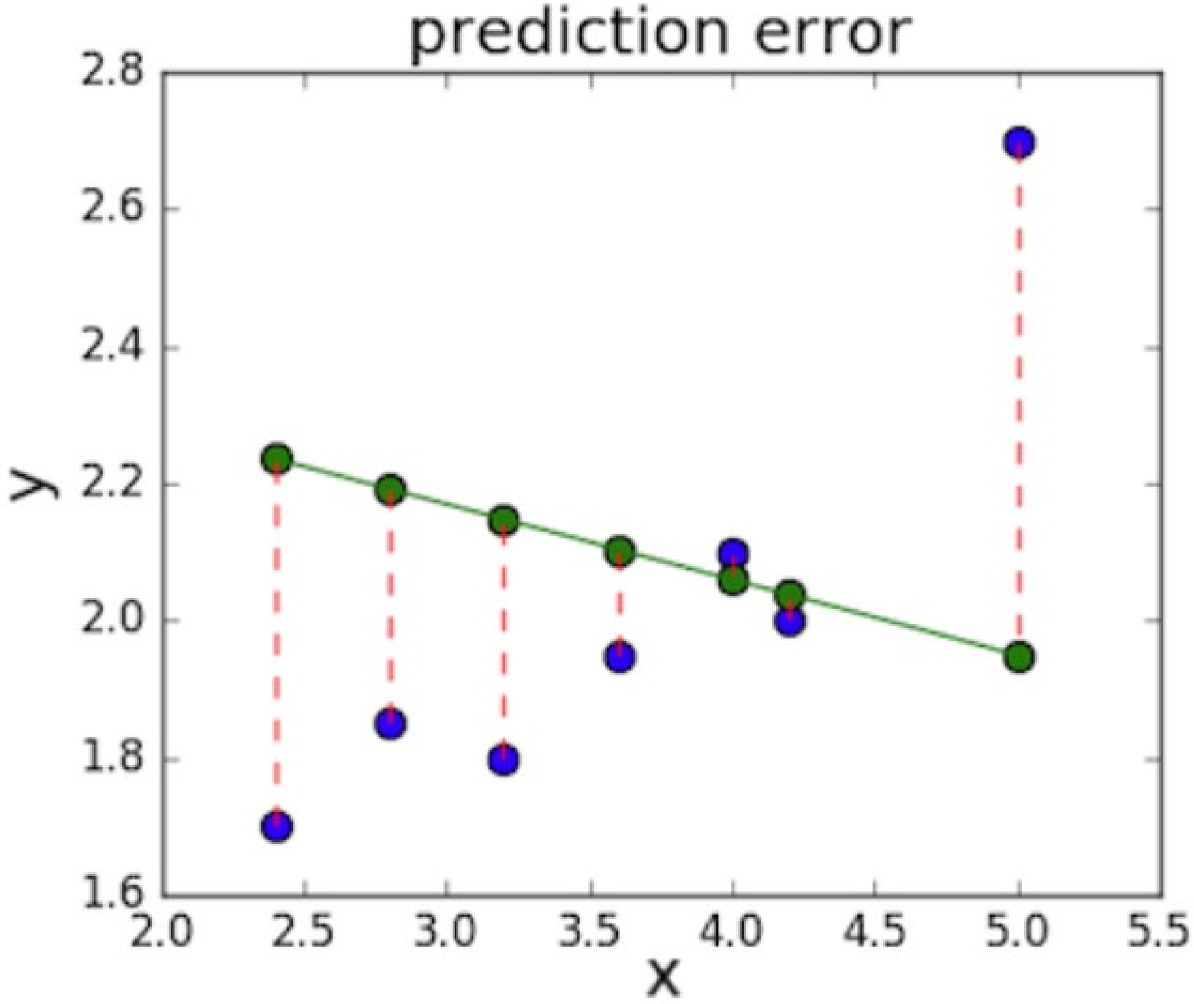
$\vec{b}$ – вектор размерности M

A – матрица размерности $M * N$

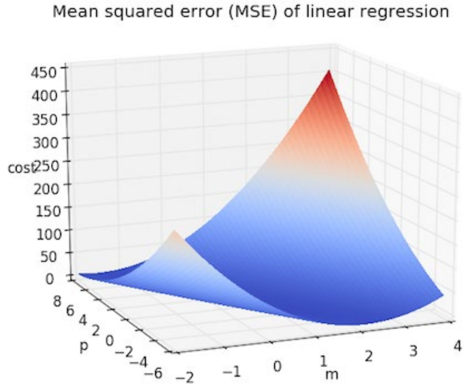
«Линейная» = линейная по параметрам

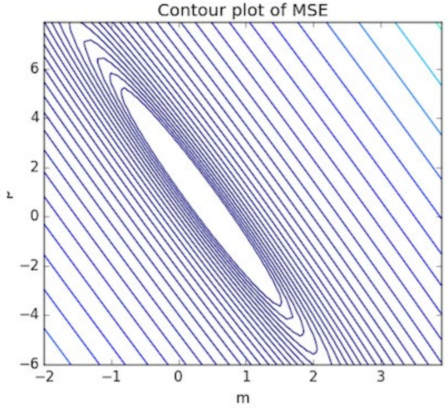
Simple Linear Regression	$y = b_0 + b_1 x_1$
Multiple Linear Regression	$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n$
Polynomial Linear Regression	$y = b_0 + b_1 x_1 + b_2 x_1^2 + \ldots + b_n x_1^n$

Линейная регрессия. Расчет ошибки



$$e(x) = y(x) - t(x)$$





$$E = \sum_{i=1}^{P} e(x_i)^2 = \sum_{i=1}^{P} (y(x_i) - t(x_i))^2$$

x_i : i – ый пример входа

$t(x_i)$: указание учителя для него

$y(x_i)$: выход модели для него

* https://neurohive.io/ru/osnovy-data-science/linejnaja-regressija/

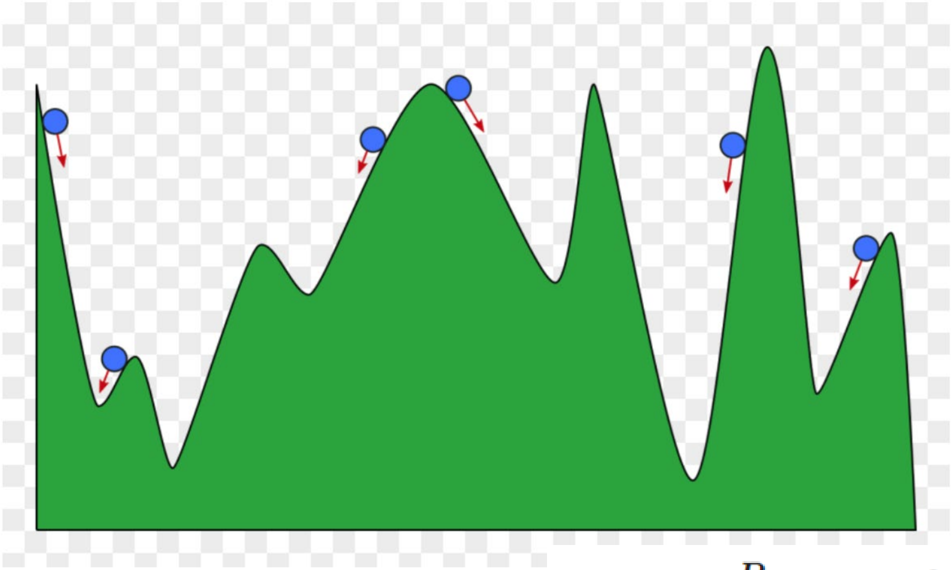
Линейная регрессия. Обучение

Не итерационный МНК

$$\begin{cases} a\sum_{i=1}^{n} x + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i \\ na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \end{cases}.$$

$$\boldsymbol{\theta} = (\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{X}^\top \mathbf{y}$$

Итерационный МНК
Градиентный спуск

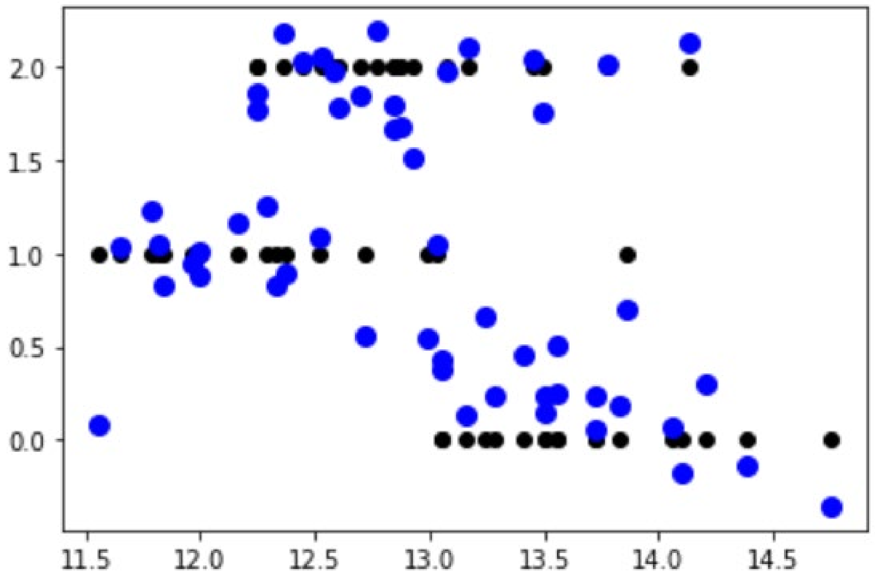


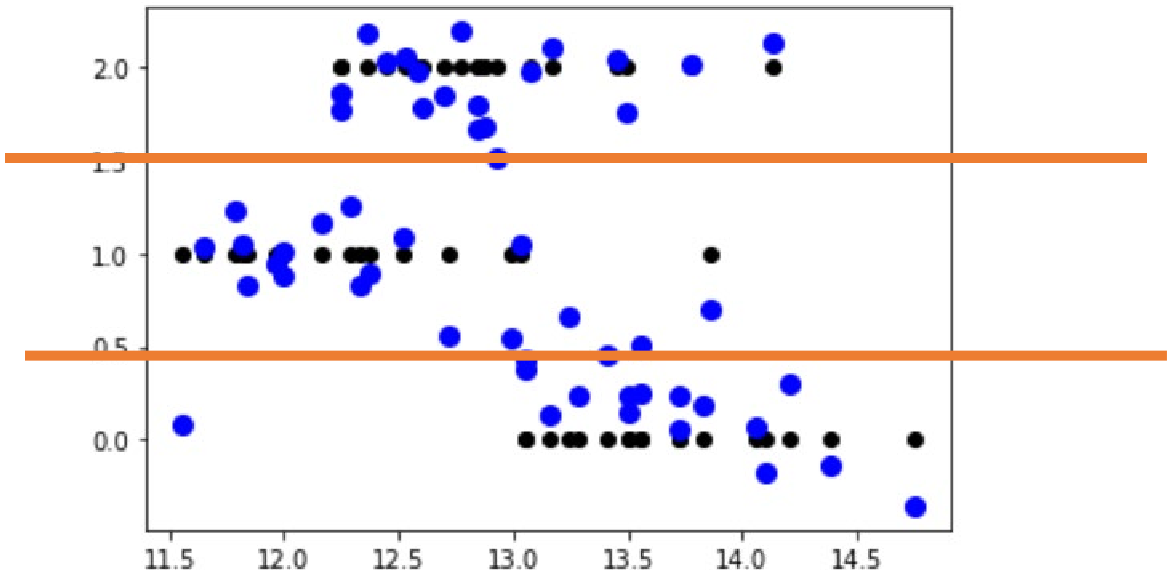
$$E = \sum_{i=1}^{P} e(x_i)^2 = \sum_{i=1}^{P} (y(x_i) - t(x_i))^2$$

for n = 1 : N

$$w_i^n := w_i^{n-1} - \eta \frac{\partial E_n(\mathrm{w})}{\partial w_i}$$

Линейная регрессия для классификации

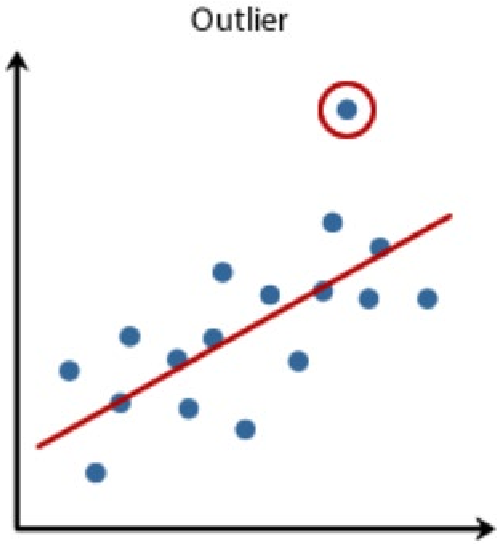




Accuracy = 0.94

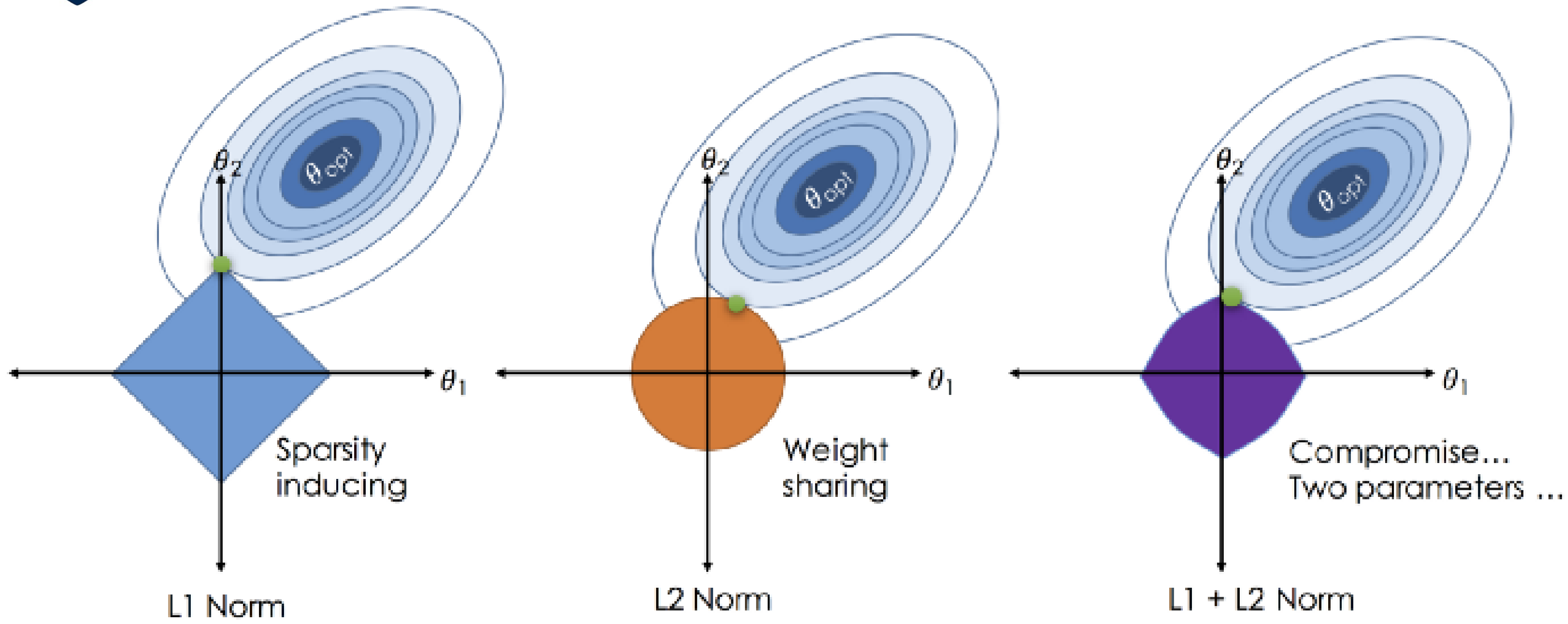
Линейная регрессия. Проблемы

- Восстанавливает только линейные зависимости
- Не очень удобна для классификации
- Чувствительна к выбросам!



* https://statistics.laerd.com/spss-tutorials/linear-regression-using-spss-statistics.php

Регуляризация



Lasso

Ridge

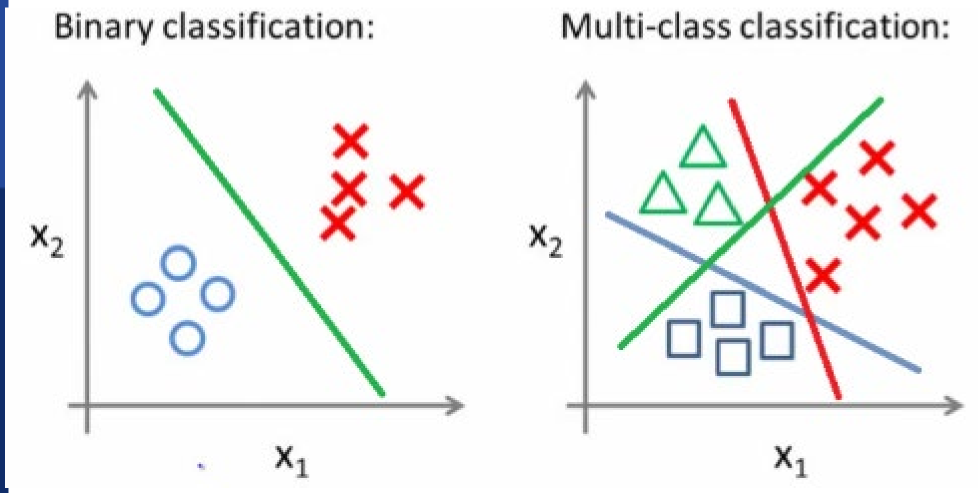
ElasticNet

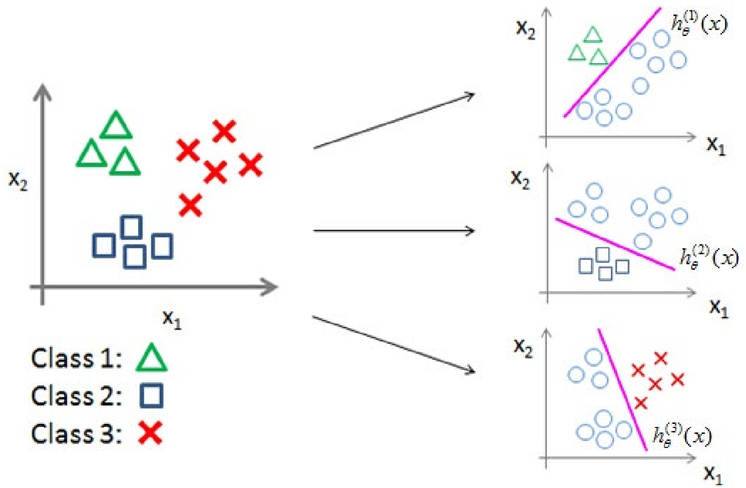
$$L_1 = \sum_i |W_i|$$

$$L = L_{\text{задача}} + \alpha * L_{\text{регуляризация}}$$

$$L_2 = \sum_i W_i^2$$

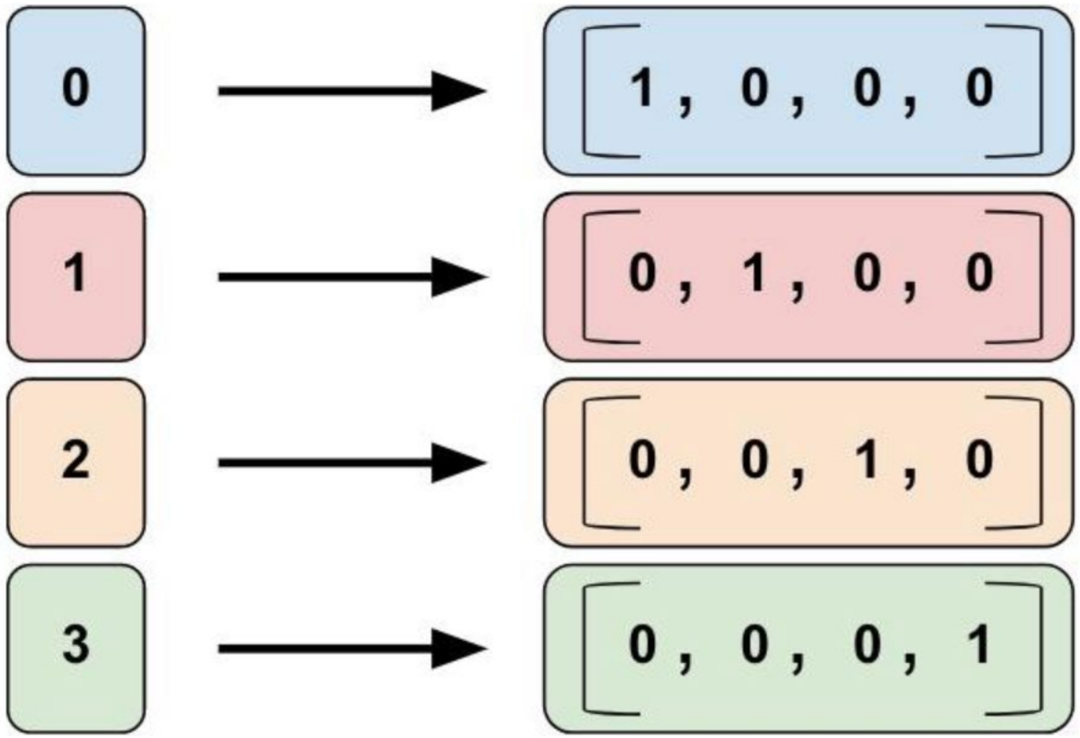
Логистическая регрессия. Задача классификации





https://codeantenna.com/a/bZcyRpm1ic

One-hot encoding



* https://colab.research.google.com/github/dpressel/dlss-tutorial/blob/master/1_pretrained_vectors.ipynb

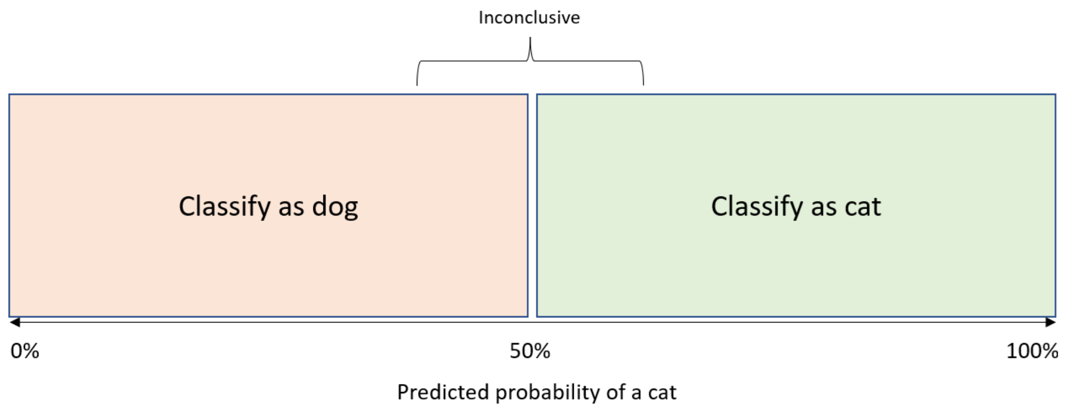
Логистическая регрессия. Уровень уверенности

Выход 0.7 – Цель 1
Выход 0.9 - Цель 1
Выход 0.3 – Цель 0
Выход 0.4 – Цель 0

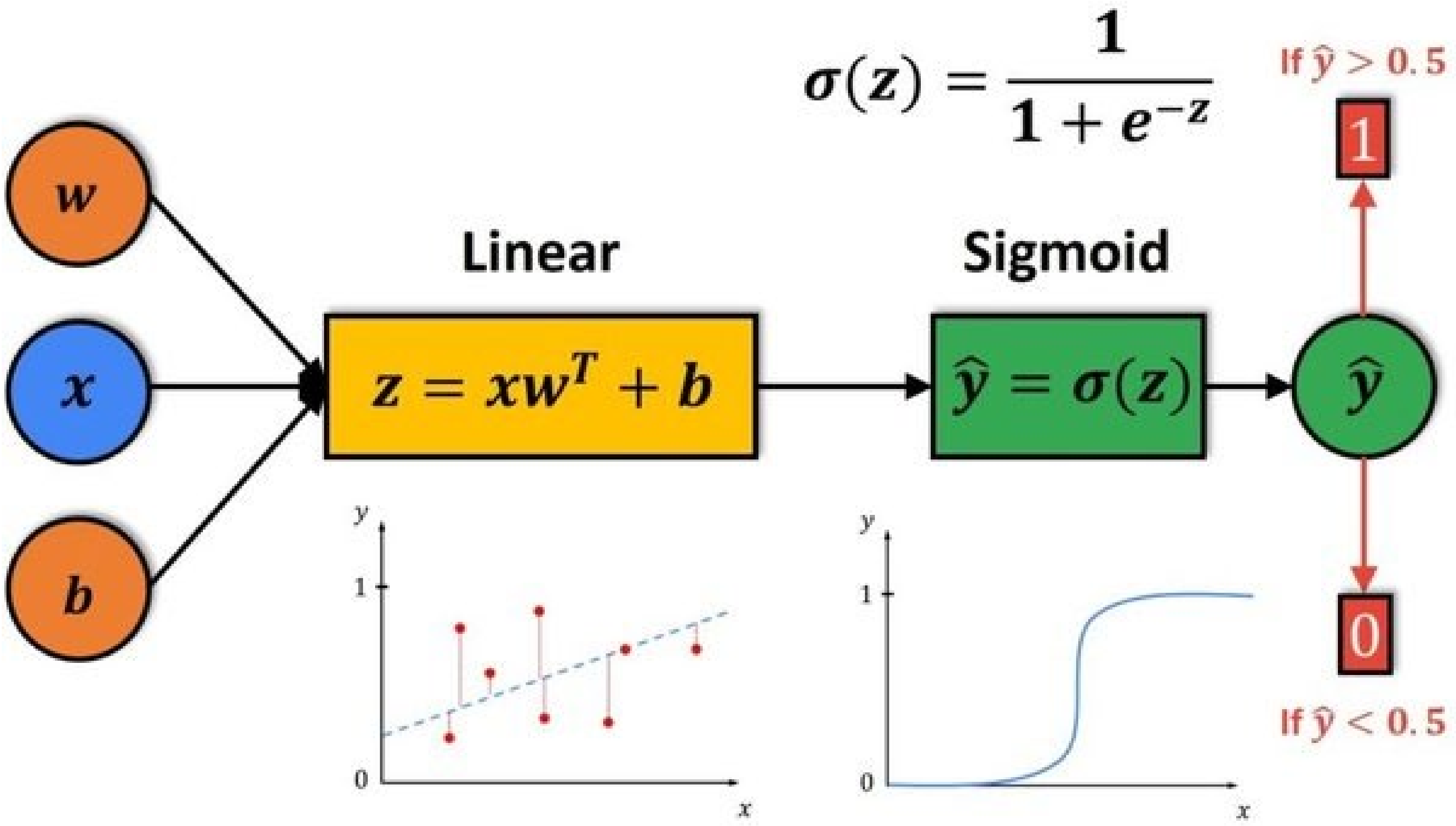
Выход 0.5 – ??

Выход 0.7 – Цель 0
Выход 0.9 - Цель 0
Выход 0.3 – Цель 1
Выход 0.4 – Цель 1

Выход 0.5 – ??



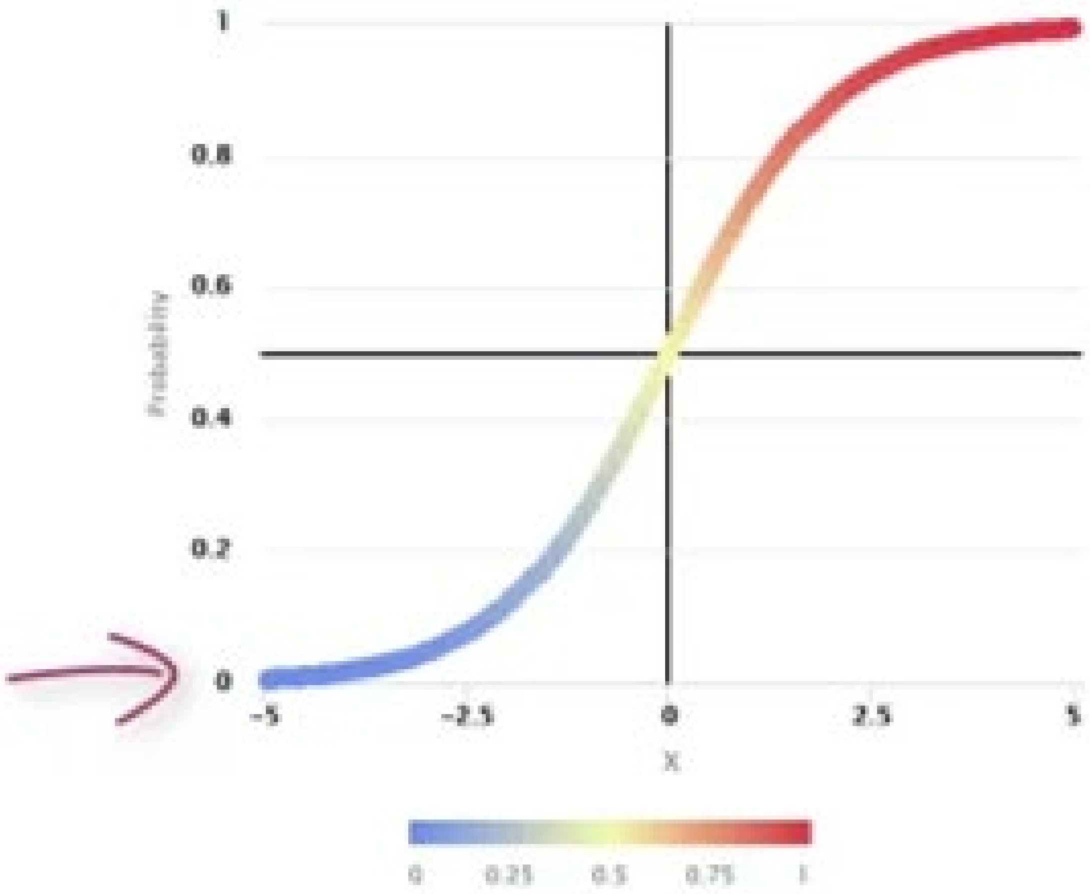
Логистическая регрессия. Модель



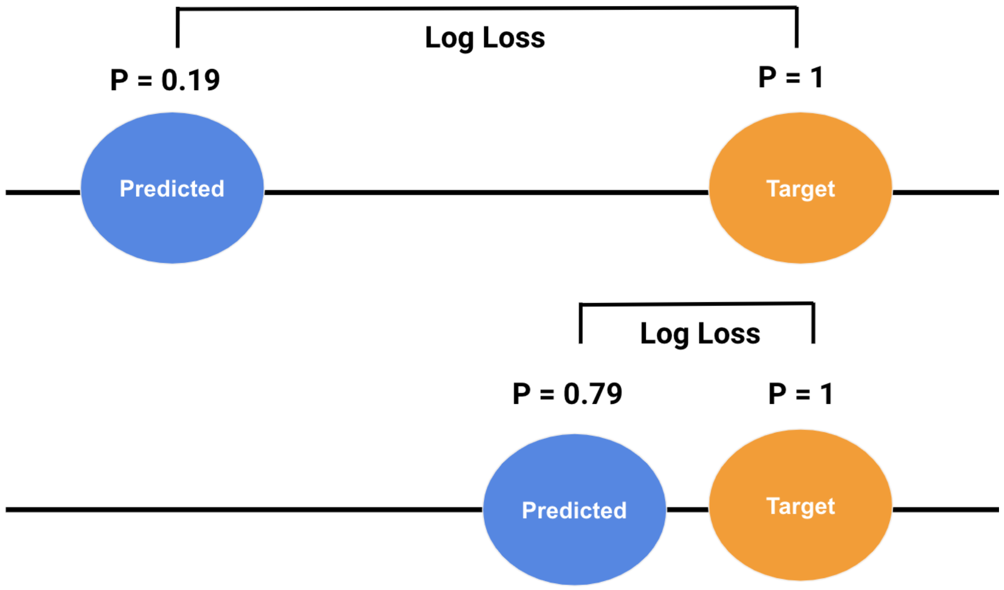
https://www.quora.com/Can-we-use-logistic-regression-for-continuous-variables

Логистическая регрессия. Сигмоида

- Logit: $\ln[p/(1-p)] = a + BX$
- Logistic: $p = \frac{e^{a+BX}}{1+e^{a+BX}}$



Логистическая регрессия.
Ошибка классификации LogLoss



Один пример:
Выход **Y** | Цель t=**1** → уверенность E=**Y**
Выход **Y** | Цель t=**0** → уверенность E=**1-Y**

Совместим: **E=(1-Y)*(1-t)+t*Y**

https://emilyswebber.github.io/LogLoss/

Логистическая регрессия. Ошибка классификации LogLoss

Много примеров:

0	1	1
0.1	0.9	0.8

(1-0.1) * 0.9 * 0.8 = 0.648

0	1	1
0.4	0.6	0.5

(1-0.4) * 0.6 * 0.5 = 0.18

0	1	1
0.7	0.2	0.1

(1-0.7) * 0.2 * 0.1 = 0.006

0	1	1
0.2	0.7	0.2

(1-0.2) * 0.7 * 0.2 = 0.112

- Простая сумма не подойдет
- Произведение уверенностей!
- Произведение сложно считать ⍰ сумма логарифмов (со знаком минус)

$$E = \sum_{i=1}^{P}(t(x_i) * ln(y(x_i)) + (1 - t(x_i)) * ln(1 - y(x_i)))$$

x_i : i – ый пример входа

$t(x_i)$: указание учителя для него

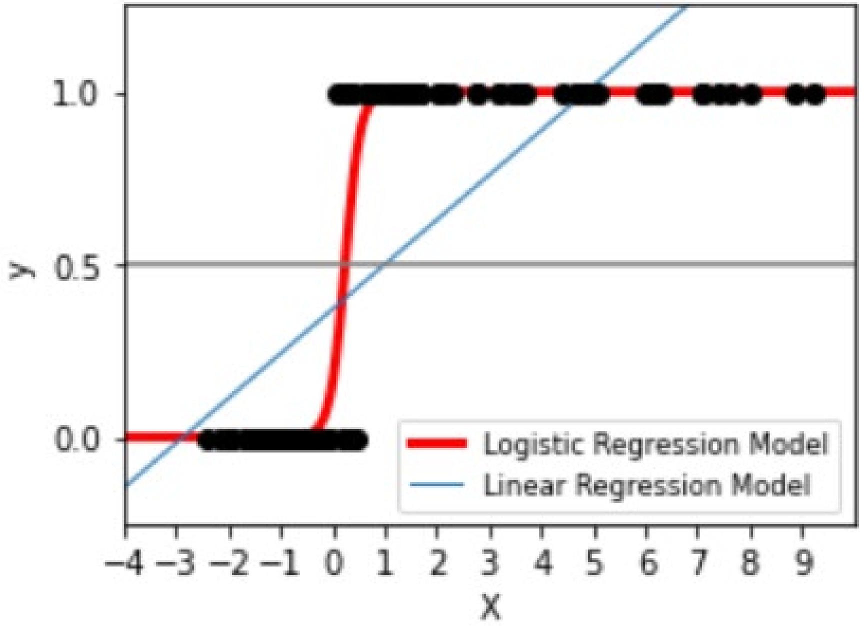
$y(x_i)$: выход модели для него

Логистическая регрессия.
Реализация sklearn

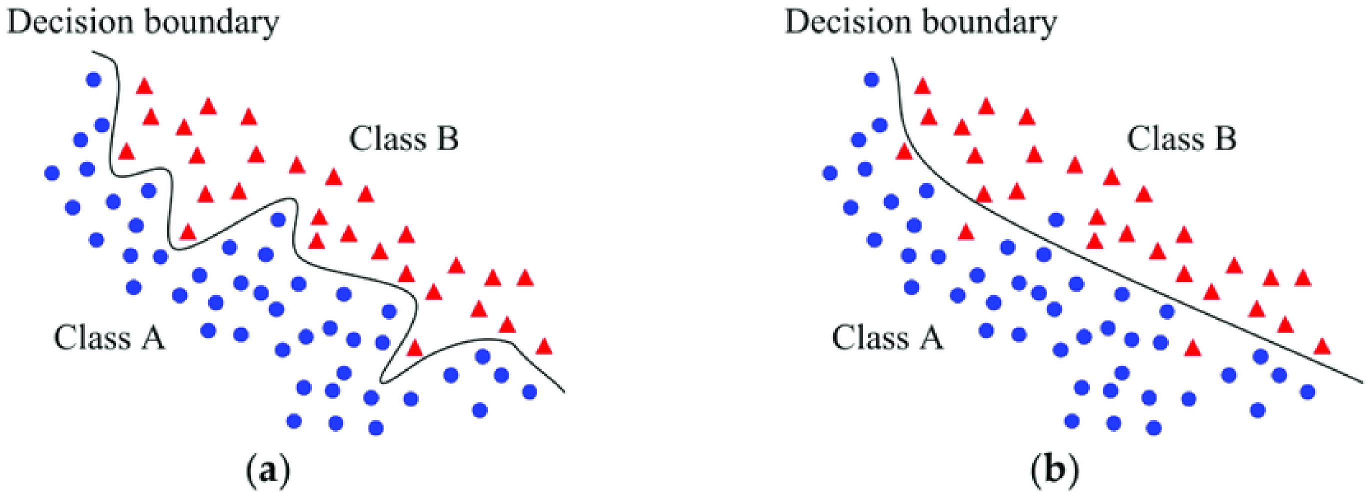
sklearn.linear_model.LogisticRegression¶

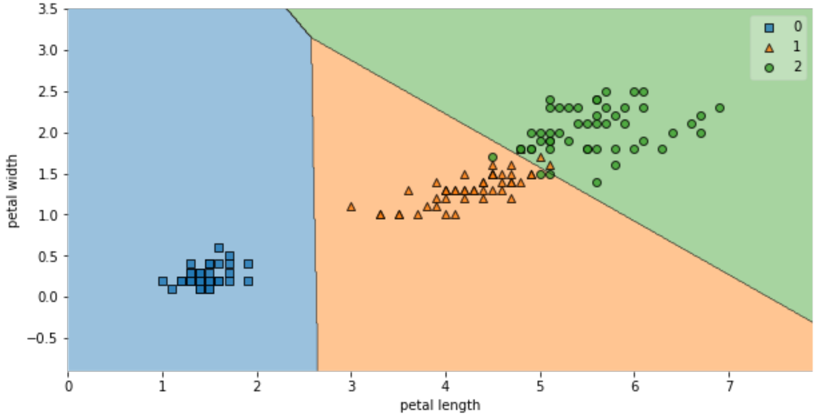
L+1/C * reg

class sklearn.linear_model.**LogisticRegression**(*penalty='l2', *, dual=False, tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1, class_weight=None, random_state=None, solver='lbfgs', max_iter=100, multi_class='auto', verbose=0, warm_start=False, n_jobs=None, l1_ratio=None*) [source]



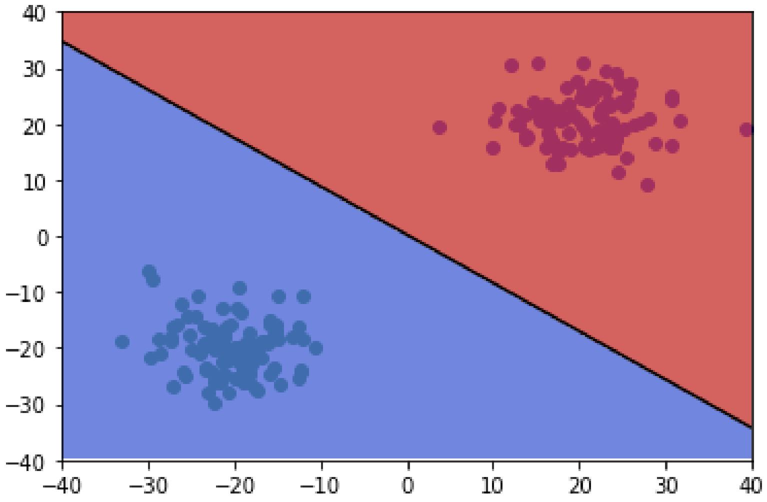
Логистическая регрессия. Разделяющая поверхность



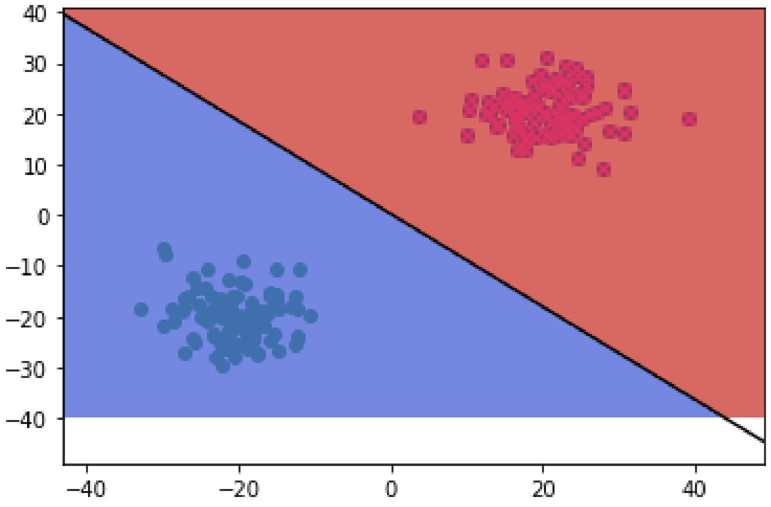


https://www.researchgate.net/figure/Example-of-overfitting-in-classification-a-Decision-boundary-that-best-fits-training_fig1_349186066

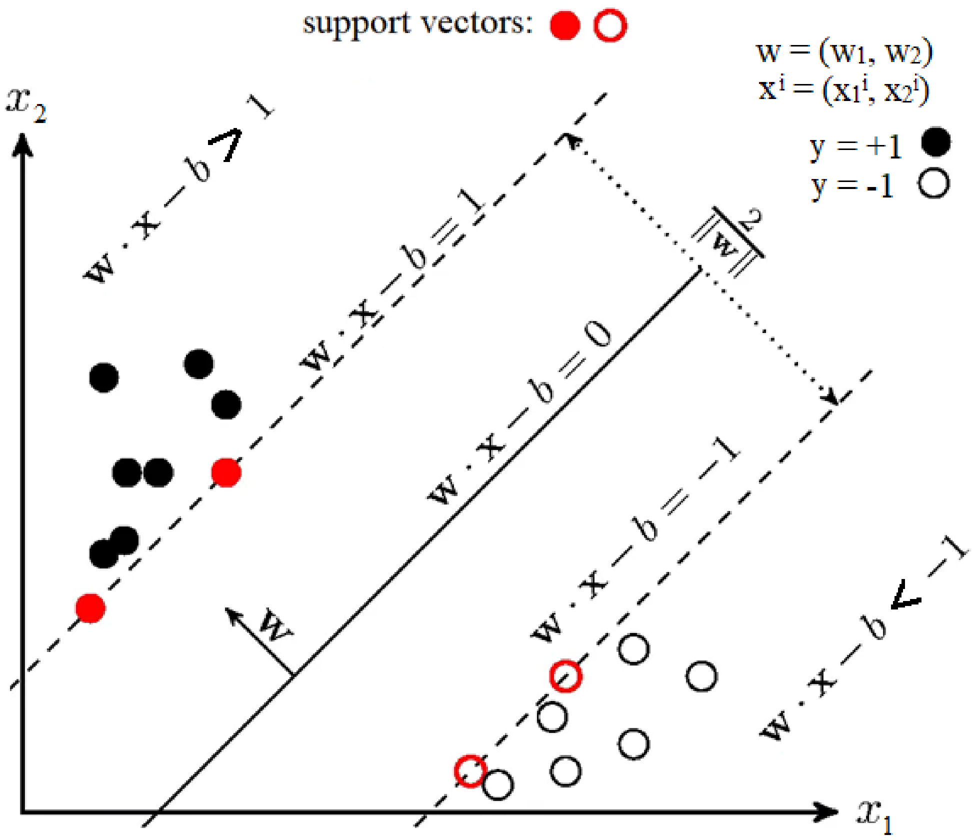
Линейная регрессия y=0

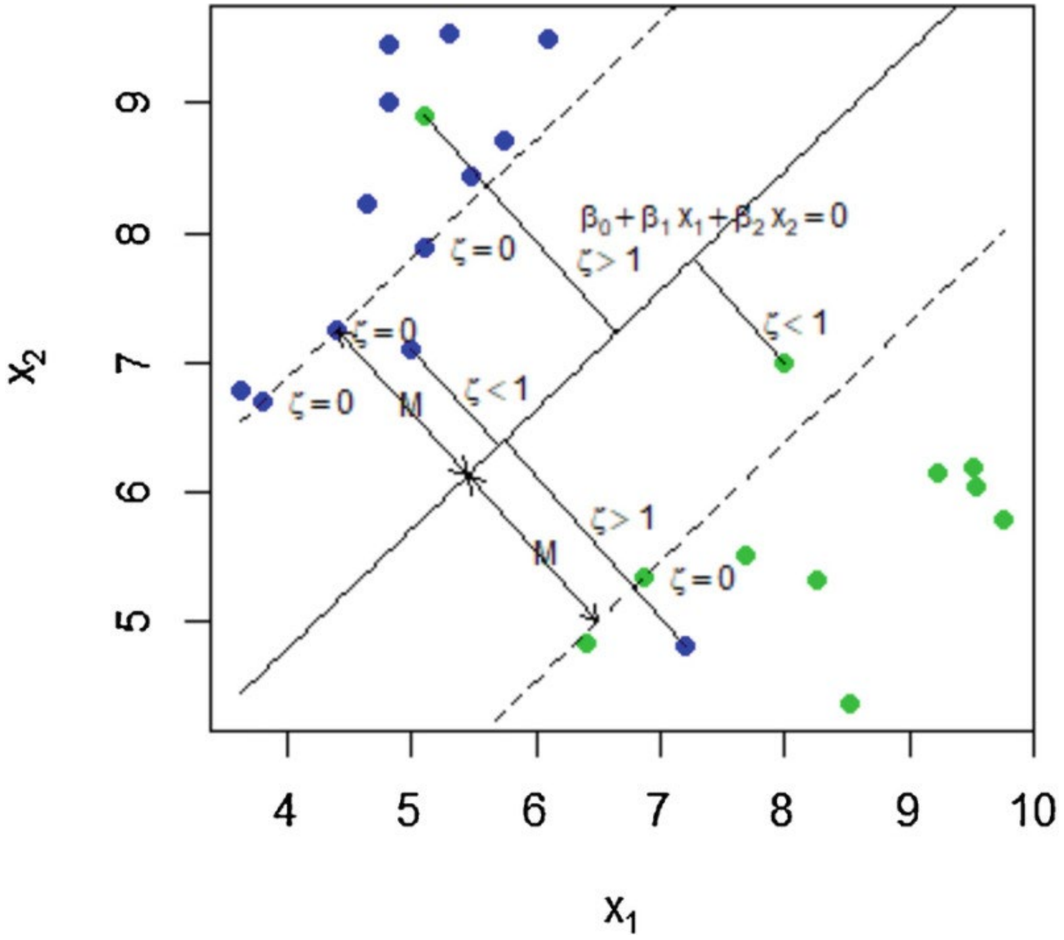


Логистическая регрессия y=0.5



Машины опорных векторов





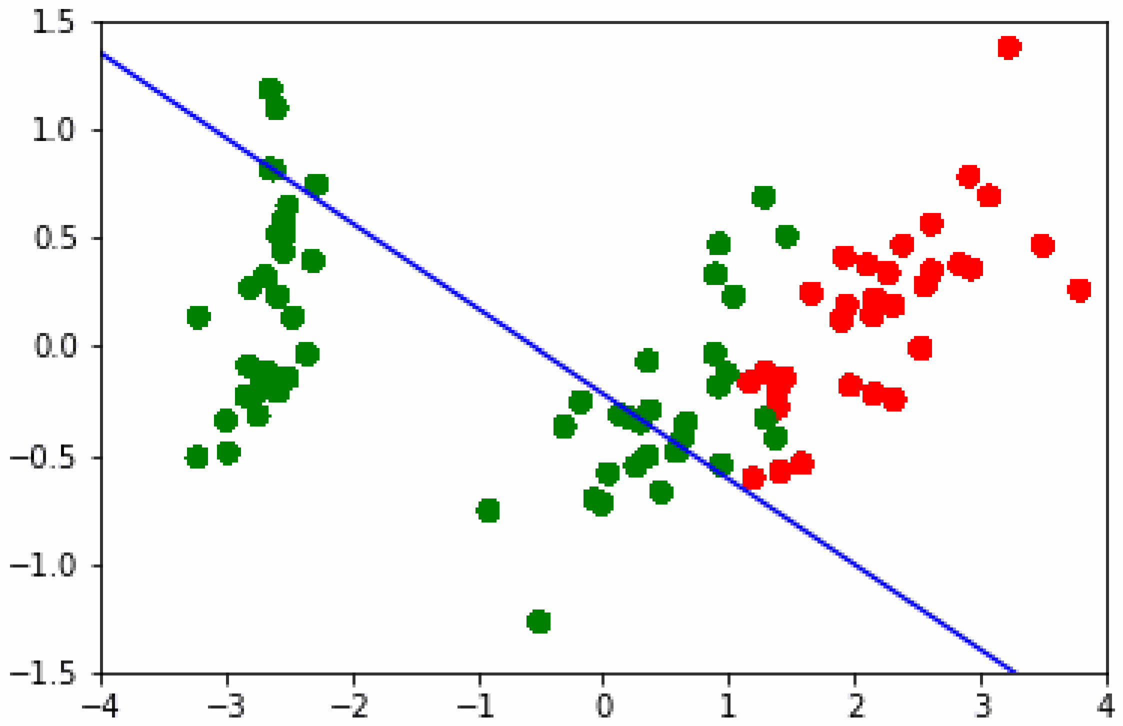
$$\begin{cases} (w^T w)/2 \to min \\ y(w^T x - b) \geqslant 1 \end{cases}$$

$$M = y(w^T x - b)$$

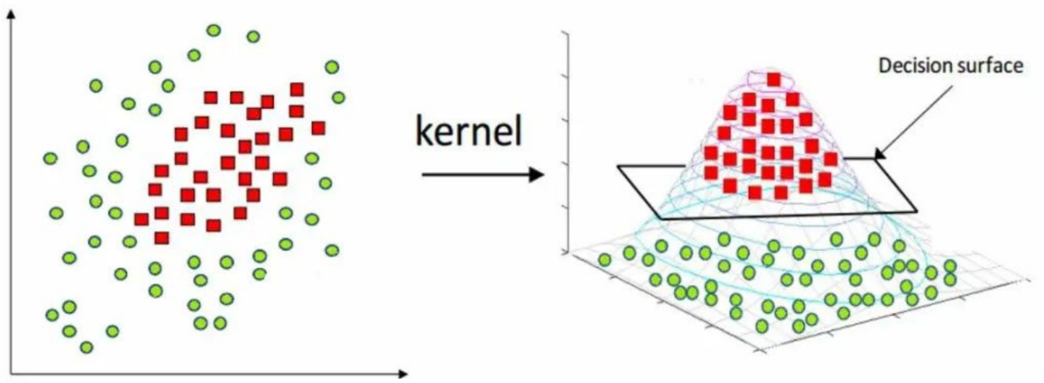
$$Q = max(0, 1 - M_i) + \alpha(w^T w)/2$$

https://habr.com/ru/companies/ods/articles/484148/

Машины опорных векторов



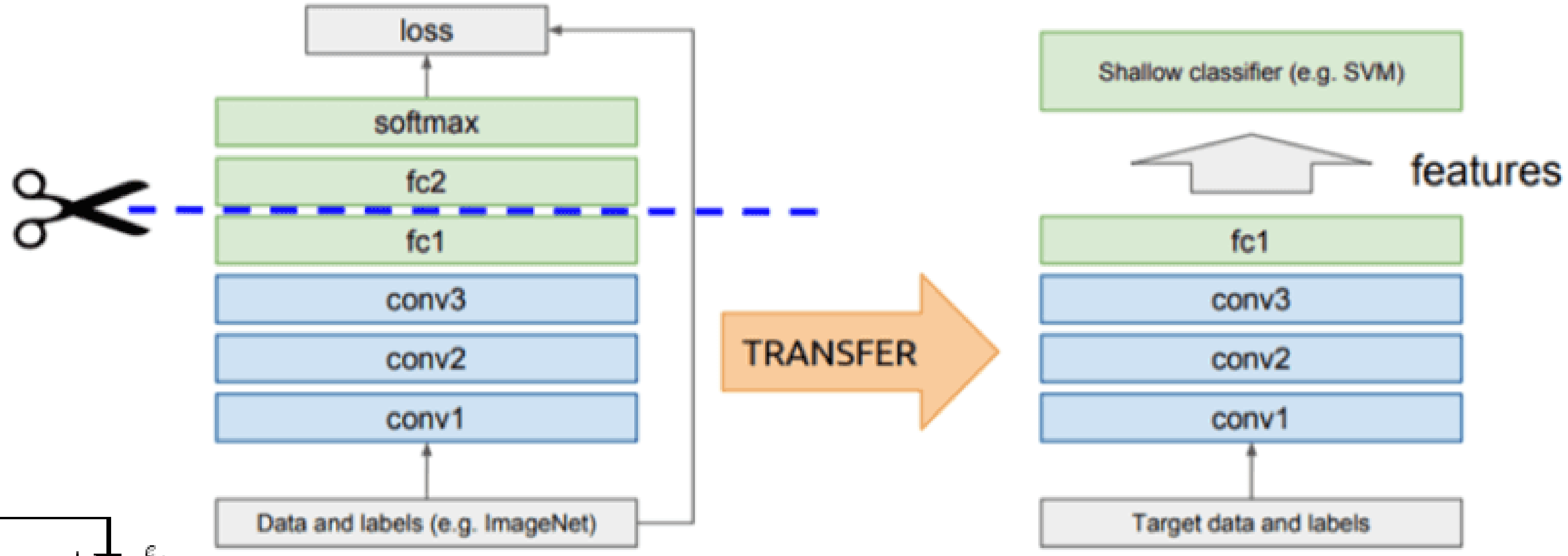
Data Transformed with SVM

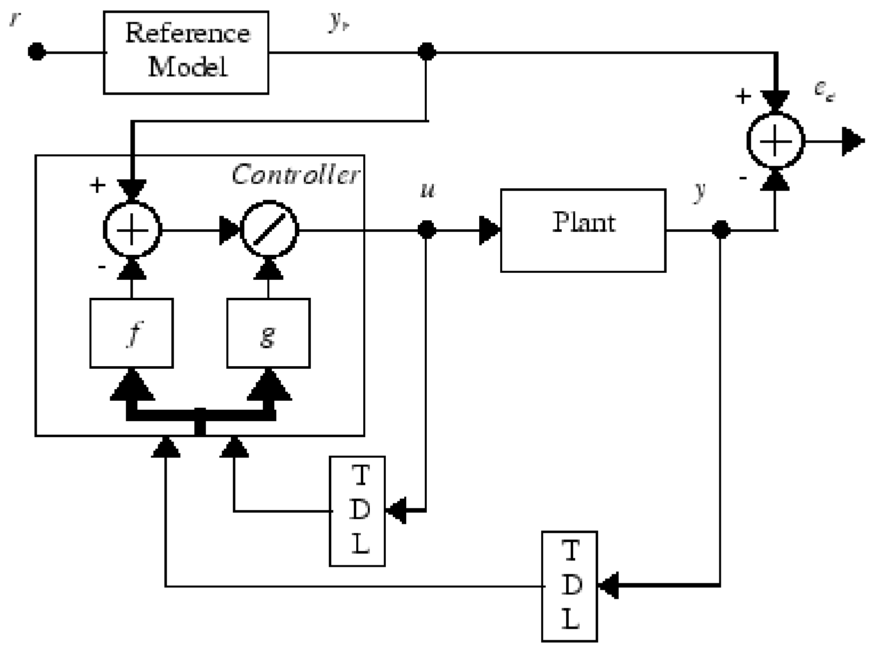


$$y(x) = w^T\phi(x) + b$$

- Common kernel functions for SVM
  - linear $k(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1 \cdot \mathbf{x}_2$
  - polynomial $k(\mathbf{x}_1, \mathbf{x}_2) = (\gamma\, \mathbf{x}_1 \cdot \mathbf{x}_2 + c)^d$
  - Gaussian or radial basis $k(\mathbf{x}_1, \mathbf{x}_2) = \exp\left(-\gamma \|\mathbf{x}_1 - \mathbf{x}_2\|^2\right)$
  - sigmoid $k(\mathbf{x}_1, \mathbf{x}_2) = \tanh(\gamma\, \mathbf{x}_1 \cdot \mathbf{x}_2 + c)$

Линеаризация и перенос обучения





Q&A и контакты

Группа по дисциплине:

https://t.me/+8dShF1tFSDg0ZmJi