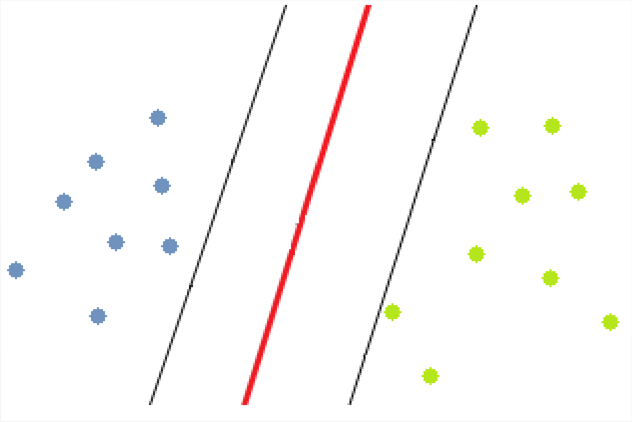
Линейные методы: SVM

Евгений Борисов

Линейные методы: SVM

Метод опорных векторов (SVM, support vector machine)

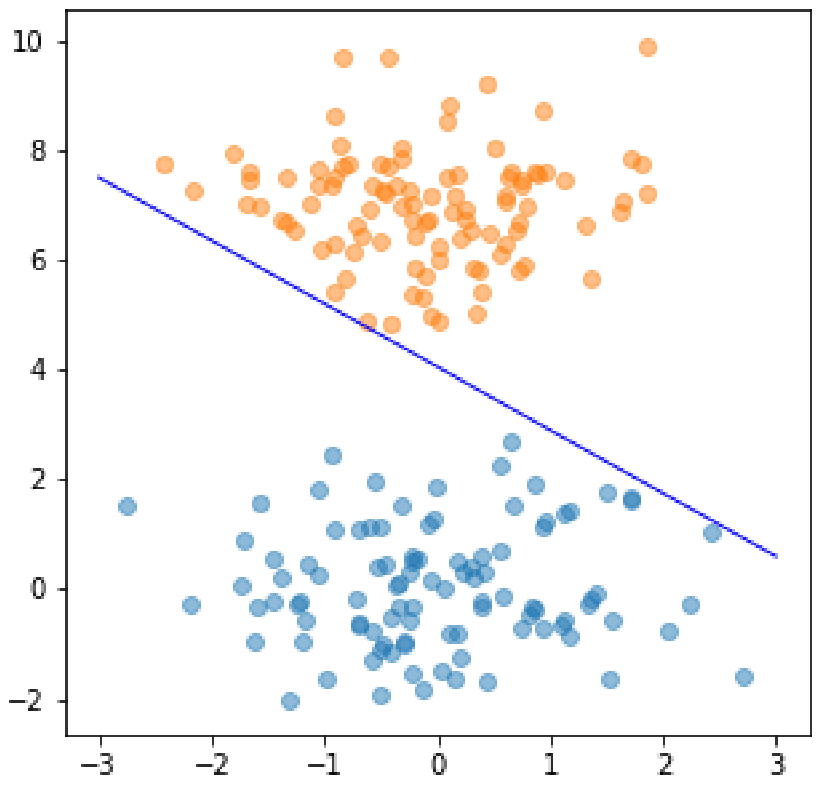
В.Н.Вапник, А.Я.Червоненкис, (1963)

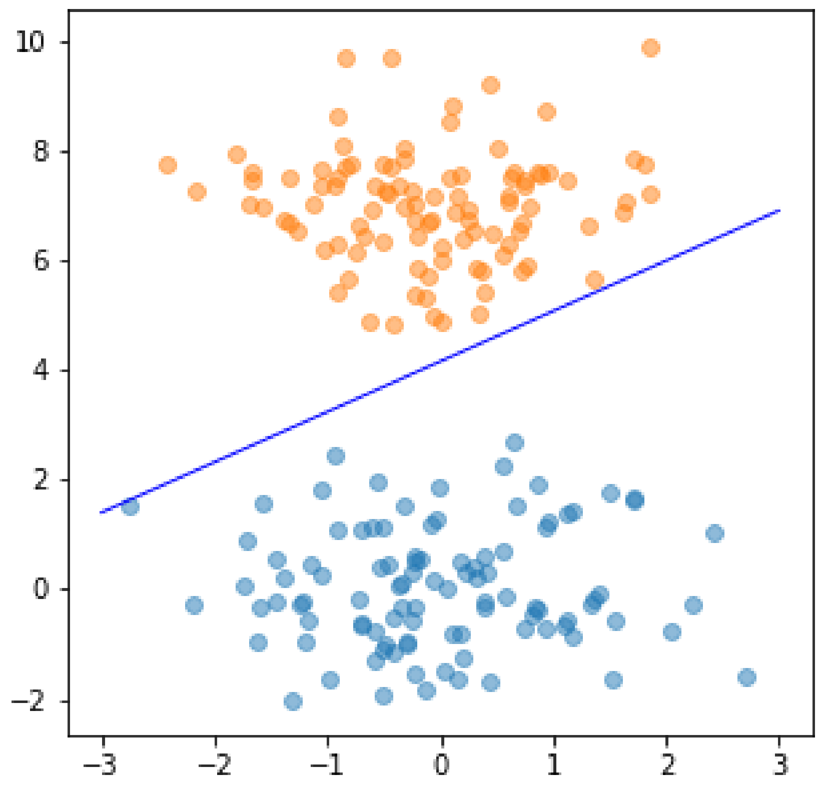


Линейные методы: SVM

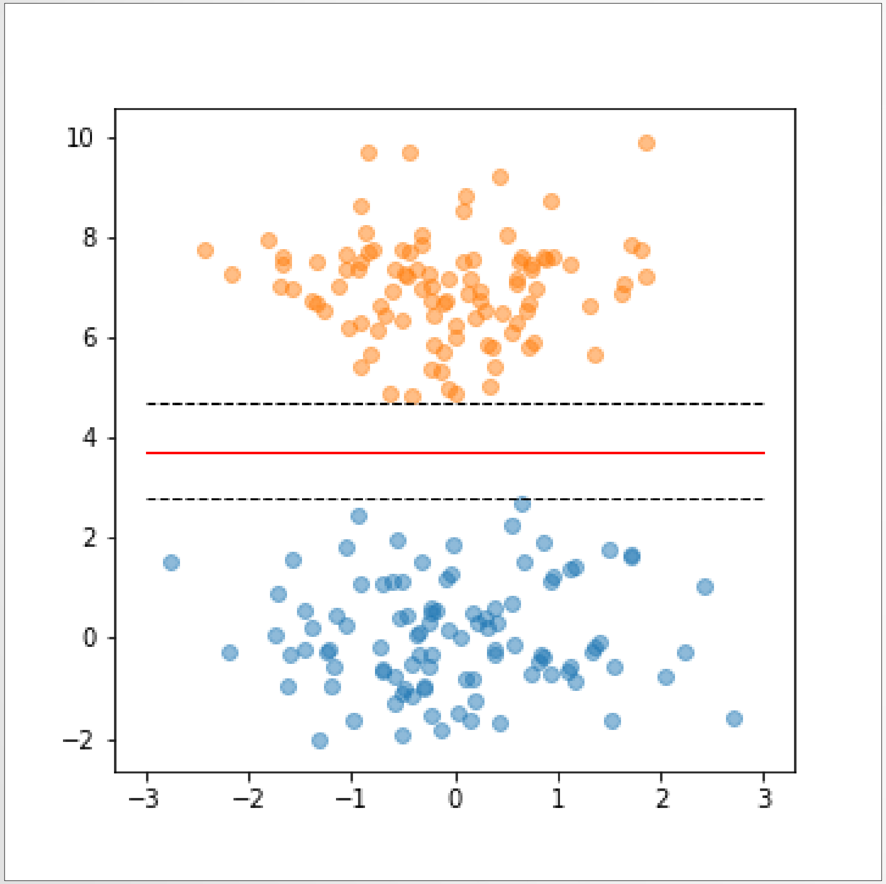
линейно разделимый набор

много разделяющих гиперплоскостей





Линейные методы: SVM

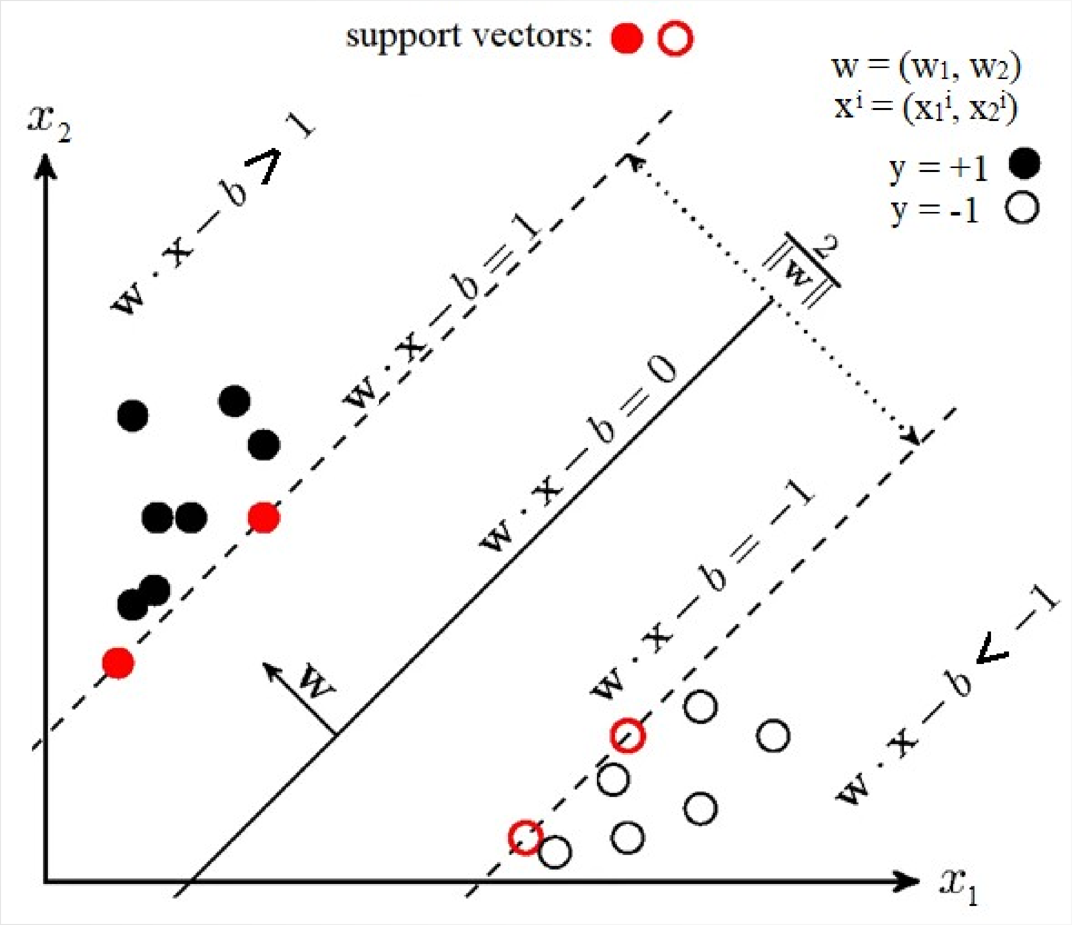


<u>**разделительная полоса**</u>

опорные (support) - объекты ограничивающие полосу

цель - полоса максимальной ширины

Линейные методы: SVM



<u>метки классов</u>

$$y \in \{-1, +1\}$$

<u>модель классификатора</u>

$$a(x) = sign(x \cdot w - b)$$

<u>отступ (margine)</u>

$$M = y \cdot (x \cdot w - b)$$

<u>разделяющая гиперплоскость</u>

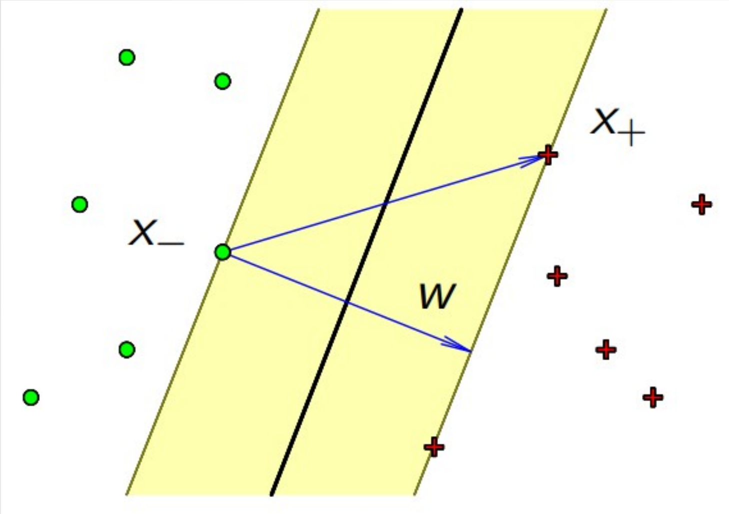
$$\langle x, w \rangle - b = 0$$

$$\langle x^+, w \rangle - b > 0$$ позитивные

$$\langle x^-, w \rangle - b < 0$$ негативные

https://habr.com/ru/company/ods/blog/484148/

Линейные методы: SVM



для определенности для опорных точек
назначим отступ в единицу (нормировка)

$$M(x_s) = y \cdot (x_s \cdot w - b) = 1$$

тогда для положительных и отрицательных
опорных точек имеем

$$\langle x_s^+, w \rangle - b = 1 \qquad \langle x_s^-, w \rangle - b = -1$$

ширина полосы

$$\left\langle x_s^+ - x_s^-, \frac{w}{\|w\|} \right\rangle = \frac{\langle x_s^+, w \rangle - \langle x_s^-, w \rangle}{\|w\|} = \frac{(b+1) - (b-1)}{\|w\|} = \frac{2}{\|w\|}$$

http://www.machinelearning.ru/wiki/images/archive/a/a0/20150316172222%21Voron-ML-Lin-SVM.pdf

Линейные методы: SVM

Задача оптимизации:

Условие 1. максимизация ширины полосы

$$\frac{2}{\|w\|} \rightarrow max \qquad (w^T \cdot w)/2 \rightarrow min$$

Условие 2. никакому объекту не разрешается попадать на полосу разделения (hard-margin SVM)

$$M(x) = y \cdot (x \cdot w - b) \geq 1$$

Решение аналитическое:

задача выпуклой квадратичной оптимизации, имеет единственное решение, эквивалентна задаче поиска седловой точки функции Лагранжа,

$$a(x) = sign\left(\sum_i \lambda_i y_i \langle x_i, x \rangle - w_0\right)$$

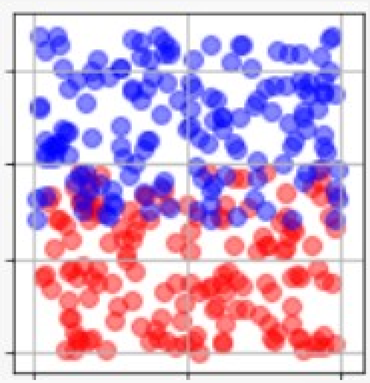
$\lambda_i \neq 0$ только для опорных объектов $\boldsymbol{x_i}$

для определения λ_i применяется алгоритм SMO (sequential minimal optimization)

Линейные методы: SVM

для **линейно НЕразделимых** данных
система будет **несовместна**

$$\begin{cases} M(x) = y \cdot (x \cdot w - b) \geq 1 \\ (w^T \cdot w)/2 \rightarrow min \end{cases}$$



ослабим условия - разрешим ошибки на учебном наборе, (soft-margin SVM)

$$\begin{cases} M(x_i) = y_i \cdot (x_i \cdot w - b) \geq 1 - \xi_i \\ (w^T \cdot w)/2 + \alpha \sum \xi_i \rightarrow min \\ \xi_i \geq 0 \end{cases}$$

для каждого примера x_i определим величину допустимой ошибки ξ_i

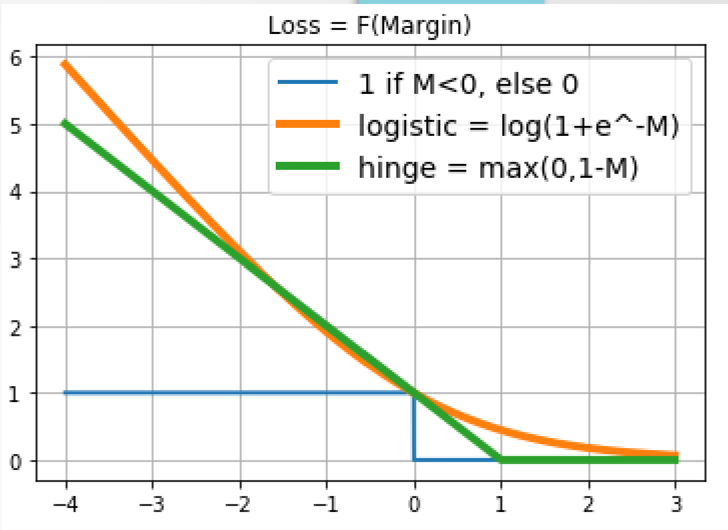
Линейные методы: SVM

будем считать количество ошибок классификатора $\sum [M_i < 0]$

и определим «штраф» за ошибку для классификатора

$$\sum [M_i < 0] \le \sum (1 - M_i)_+ = \sum max(0, 1 - M_i)$$

чем больше отступ М уходит в минус
тем больше «штраф»



Определим функцию потери на основе «штрафа» **Hinge loss**

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

Для минимизации **Hinge loss** будем применять метод градиентного спуска

$$a(x) = sign(x \cdot w - b) \qquad w := w - \eta \cdot \nabla Q \qquad \nabla Q = \begin{cases} \alpha w - w x; y x w < 1 \\ \alpha w; y x w \ge 1 \end{cases}$$

Линейные методы: литература

git clone https://github.com/mechanoid5/ml_lectorium.git

Воронцов К.В. Метод опорных векторов.
http://www.ccas.ru/voron/download/SVM.pdf

Машинное обучение для людей
https://vas3k.ru/blog/machine_learning/