

Изкуствен интелект

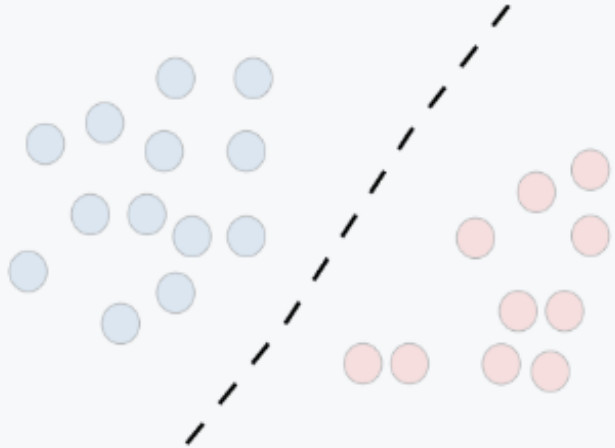
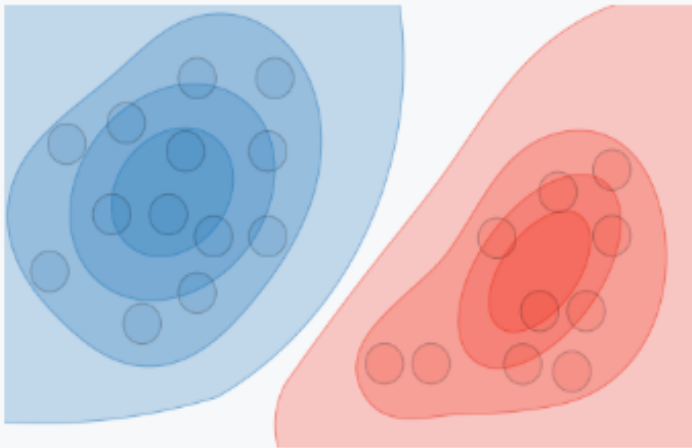
***Тема 15: Методи за машинно обучение.
Статистически методи. Класификация с
опорни вектори (SVM)***



Съдържание

- ☐ Обща класификация на методите
- ☐ Класификация с невронни мрежи. Архитектури.
MLPNN и PNN.
- ☐ Рекурентни невронни мрежи (RNN, LRPNN,
GLRPNN)
- ☐ Статистически методи (GMM, HMM)
- ☐ Класификация с опорни вектори (SVM)
- ☐ Самоорганизиращи се архитектури (SOM)
- ☐ Еволюционни и генетични методи за оптимизация
(DE, PSO)

Дискриминативни и генеративни методи

	Discriminative model	Generative model
Goal	Directly estimate $P(y x)$	Estimate $P(x y)$ to then deduce $P(y x)$
What's learned	Decision boundary	Probability distributions of the data
Illustration		
Examples	MLP NN, SVM, Decision trees, LDA...	PNN, GMM, Bayesian Networks ...

Класификатори

Discriminative approaches

LDA

Polynomial
classifier

TDNN
and RNN

FFNN

SVM

Decision
trees

Non-discriminative approaches

k-NN

LVQ

SOM

PNN

GMM

HMM

Generative approaches

Combined methods

GMM/SVM

HMM/ANN

RBF

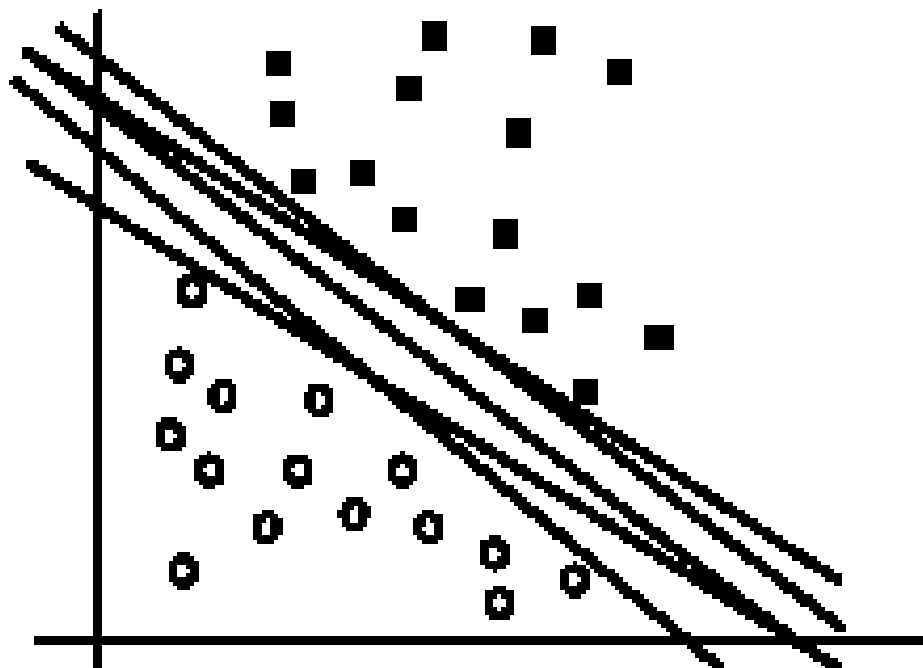
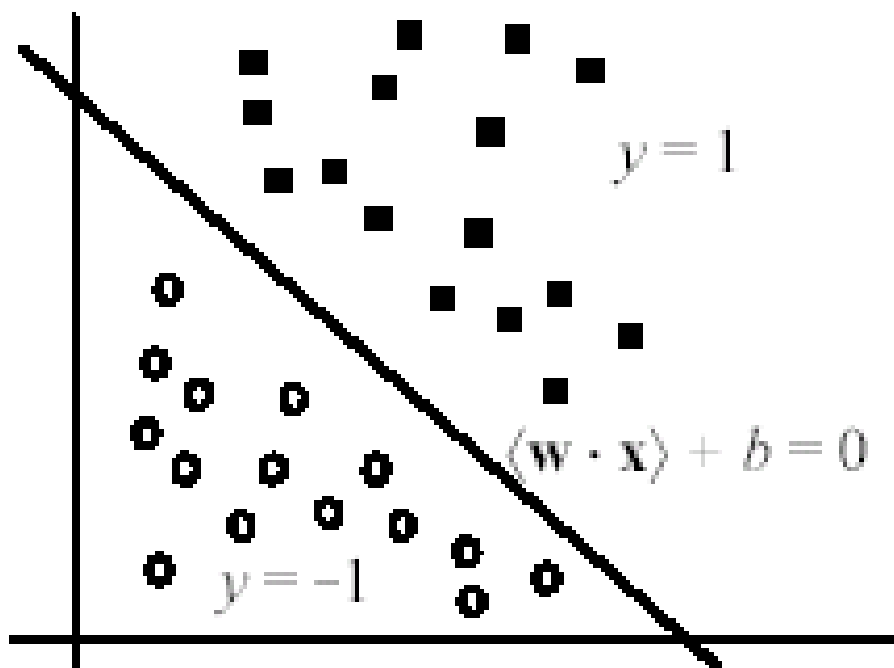
HMMs trained
discriminatively

PNN-RNN

Класификатори за два класа

Дискриминативни методи за класификация:

- Използват хипер-равнина разделяща двата класа

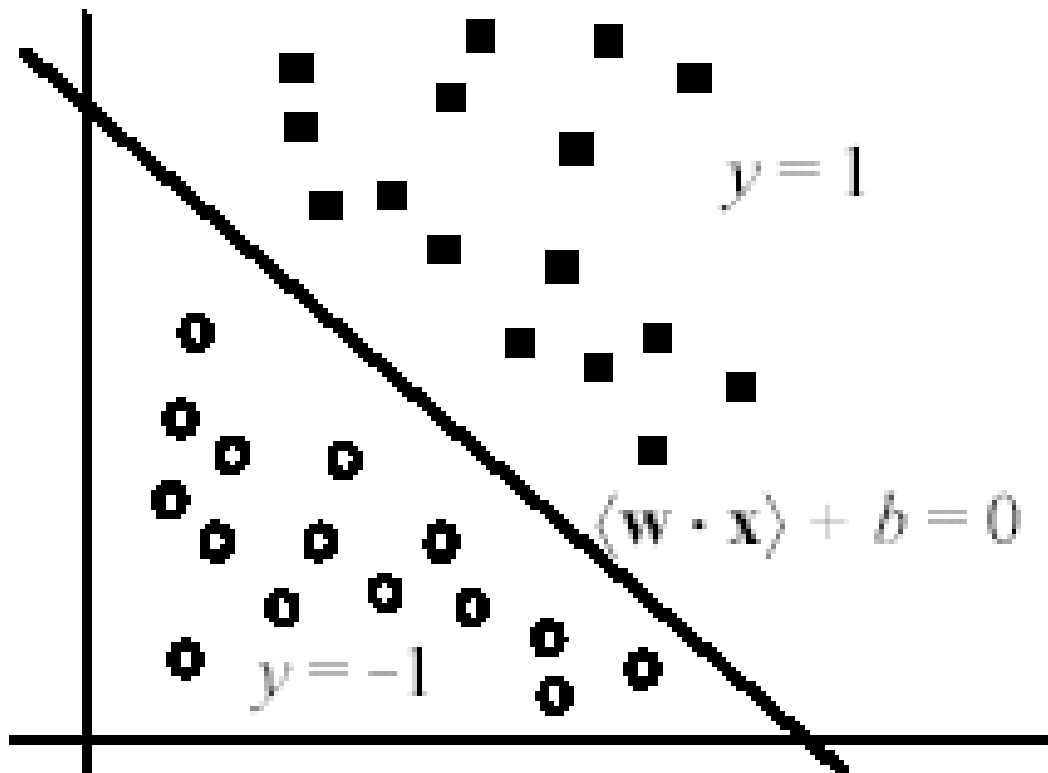


Класификатор с опорни вектори Support Vector Machine (SVM)

Вапник 1970s (и след 1992 и за западния свят)

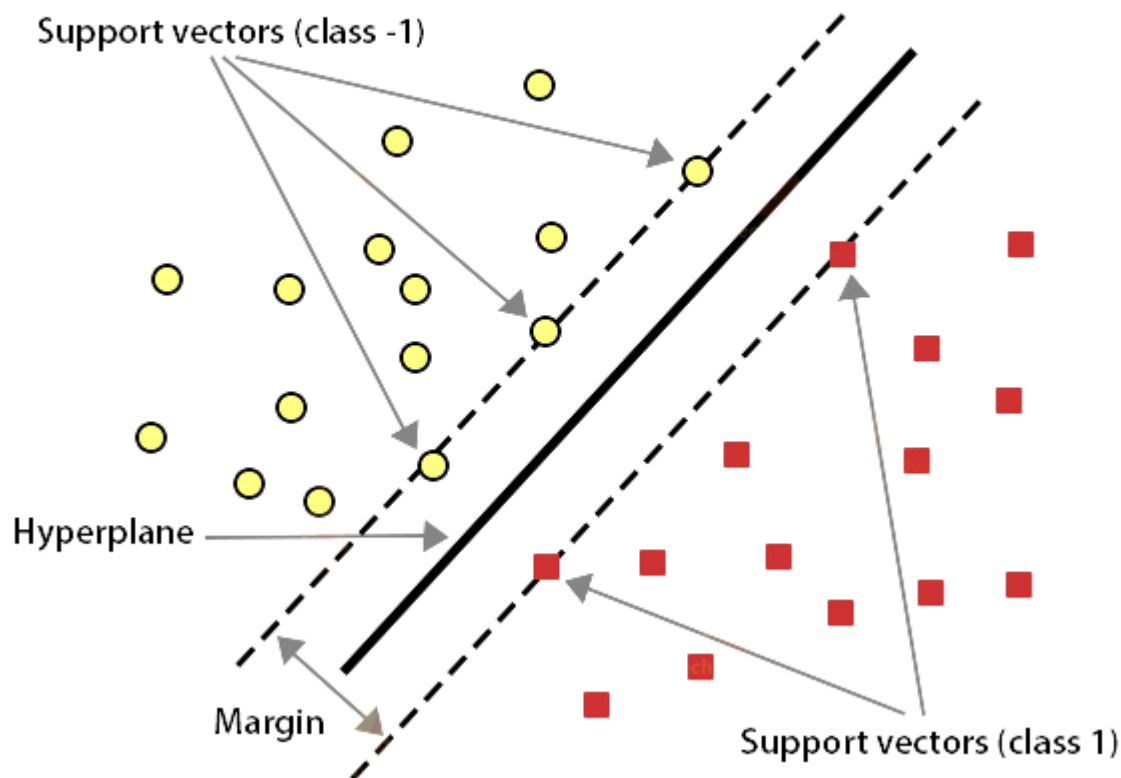
- Оптимална хипер-равнина разделяща два класа

$$\langle \mathbf{w} \cdot \mathbf{x} \rangle + b = 0$$

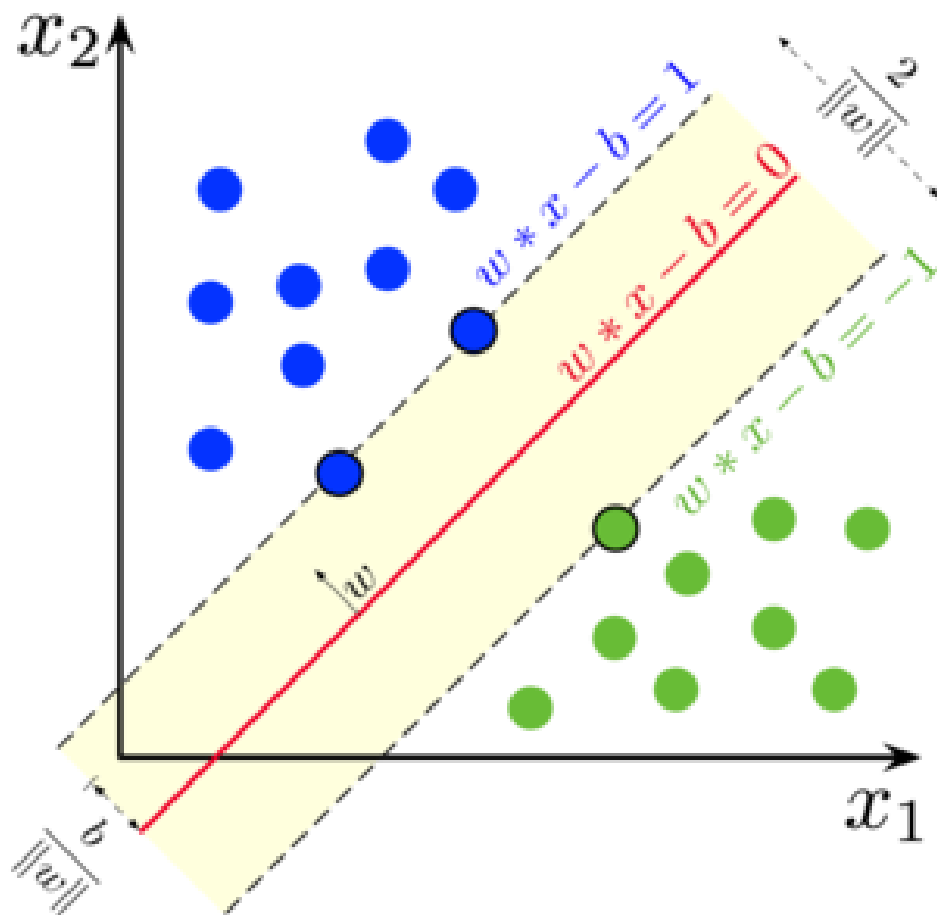


Класификатор с опорни вектори Support Vector Machine (SVM)

Support Vector Machines



Класификатор с опорни вектори Support Vector Machine (SVM)



Даден е набор от N вектора, x_k ,
 $k=1 \dots N$

Търсим възможно най-големият резерв на границата м/у двата класа.

Концепцията за SVM е да се минимизира риска от грешка за всеки от класовете ...

Класификатор с опорни вектори Support Vector Machine (SVM)

$$y(x) = \text{sign} \left[\sum_{k=1}^N a_k y_k \psi(x, x_k) + b \right]$$

Допускаме, че

$$\begin{cases} w^T \phi(x_k) + b \geq +1, y_k = +1 \\ w^T \phi(x_k) + b \leq -1, y_k = -1 \end{cases}$$

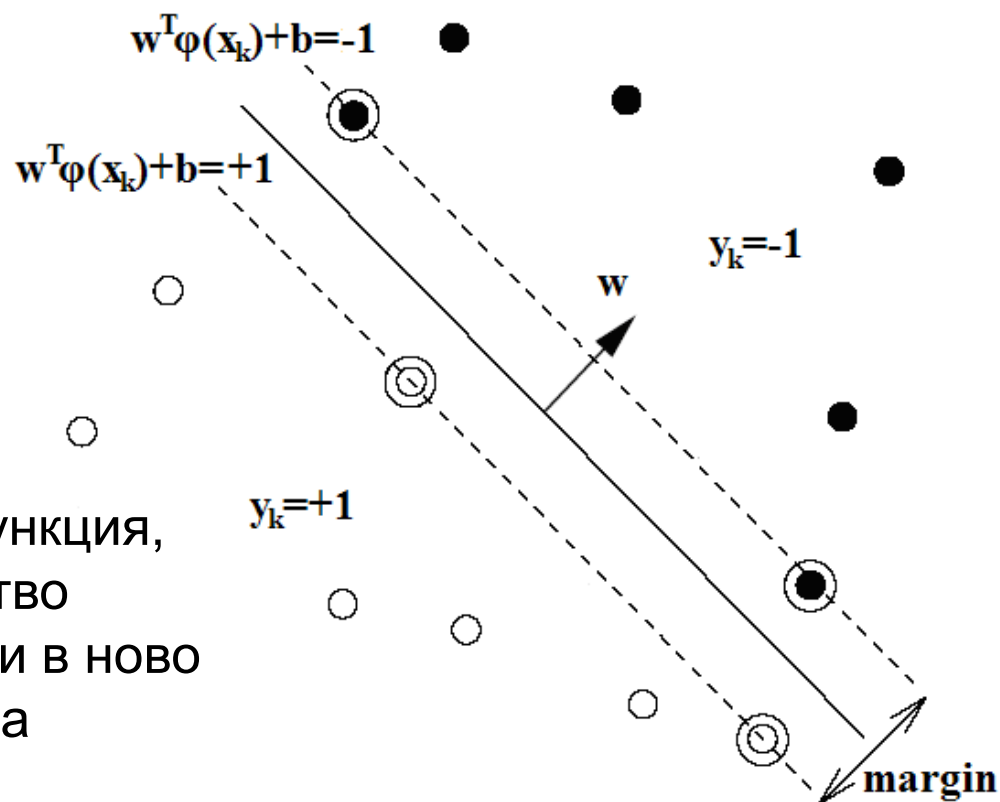
Коеето е еквивалентно на

$$y_k \left[w^T \phi(x_k) + b \right] \geq 1, k = 1, \dots, N$$

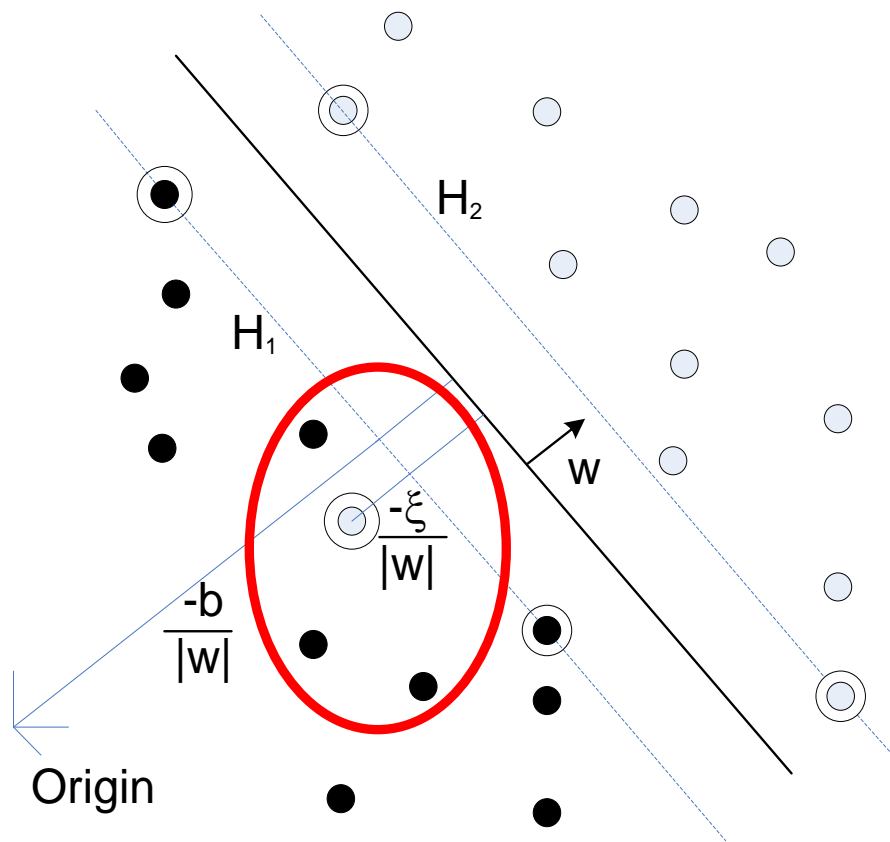
Тук $\phi(\cdot)$ е някаква нелинейна функция, която трансформира пространство дефинирано от входните вектори в ново пространство с (много) по-висока размерност

$\{y_k, x_k\}_{k=1}^N$ – k -тият изход

a_k са положителни коефициенти
 b е отместването



Класификатор с опорни вектори Support Vector Machine (SVM)



За случаите, когато двата класа не могат да бъдат разделени в пространството на описателите, към уравнението се добавят променливите ξ_k

При което формулировката на задачата придобива вида:

$$\begin{cases} y_k \left[w^T \phi(x_k) + b \right] \geq 1 - \xi_k, k = 1, \dots, N \\ \xi_k \geq 0, k = 1, \dots, N \end{cases}$$

Класификатор с опорни вектори Support Vector Machine (SVM)

Съгласно принципа за намаляване на структурния риск, границата на риска може да се минимизира посредством следната формулировка на оптимизационната задача

$$\min_{w, \xi_k} J_1(w, \xi_k) = \frac{1}{2} w^T w + c \sum_{k=1}^N \xi_k$$

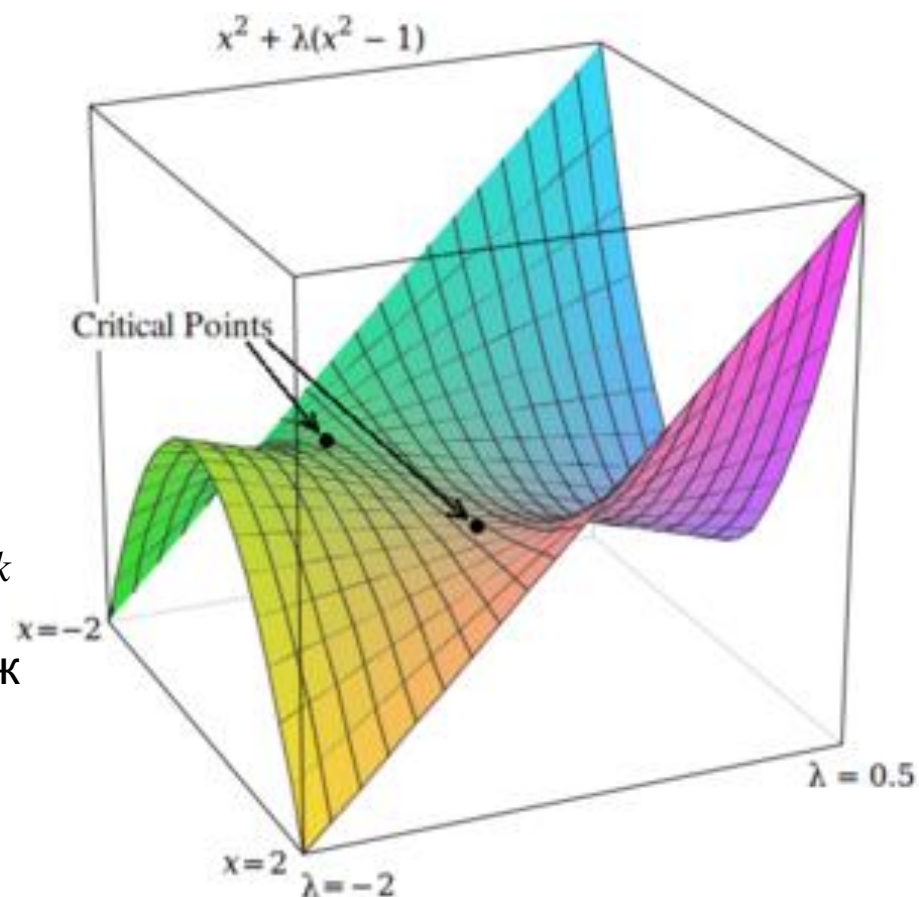
Конструираме функция на Лагранж като използваме множители на Лагранж $a_k \geq 0$, $v_k \geq 0$ ($k=1, \dots, N$).

$$L_1(w, b, \xi_k, a_k, v_k) = J_1(w, \xi_k) -$$

$$\sum_{k=1}^N a_k \{y_k [w^T \phi(x_k) + b] - 1 + \xi_k\} - \sum_{k=1}^N v_k \xi_k$$

Решението е седловината на Лагранж изчислена като

$$\max_{w, \xi_k} \min_{w, \xi_k} L_1(w, b, \xi_k; a_k, v_k)$$



Класификатор с опорни вектори Support Vector Machine (SVM)

Води до

$$\begin{cases} \frac{\partial L_1}{\partial w} = 0 \rightarrow w = \sum_{k=1}^N a_k y_k \phi(x_k) \\ \frac{\partial L_1}{\partial b} = 0 \rightarrow \sum_{k=1}^N a_k y_k = 0 \\ \frac{\partial L_1}{\partial \xi_k} = 0 \rightarrow 0 \leq a_k \leq c, k = 1, \dots, N \end{cases}$$

Което дава решението на проблем от типа *квадратично програмиране*

$$\max_{a_k} Q_1(a_k; \phi(x_k)) = -\frac{1}{2} \sum_{k,l=1}^N y_k y_l \phi(x_k)^T \phi(x_l) a_k a_l + \sum_{k=1}^N a_k$$

Така че

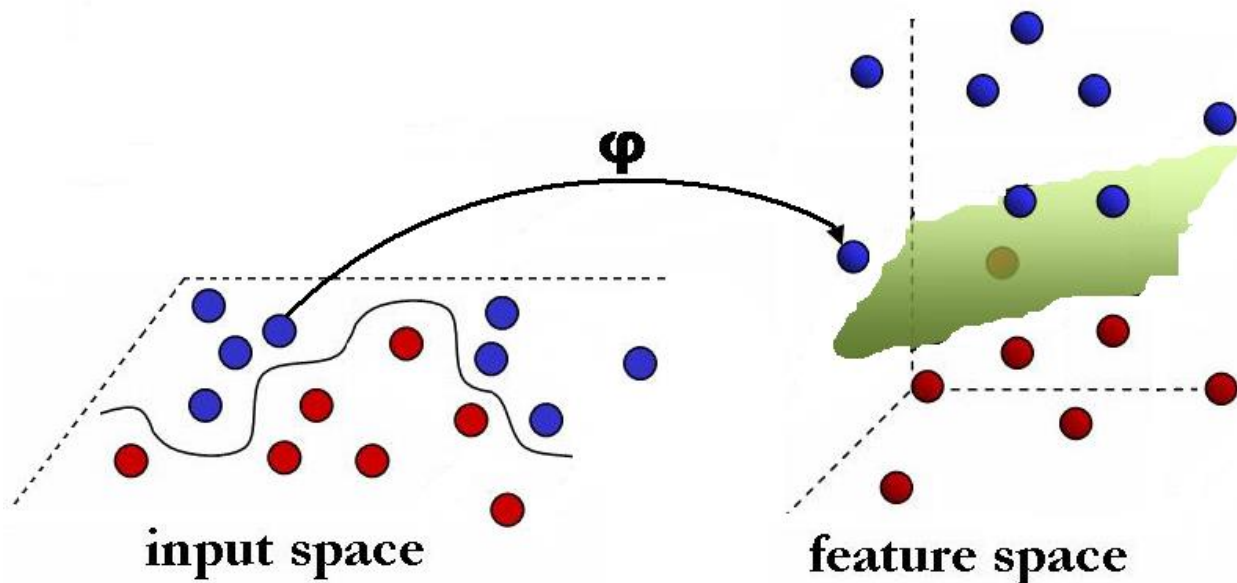
$$\begin{cases} \sum_{k=1}^N a_k y_k = 0 \\ 0 \leq a_k \leq c, k = 1, \dots, N \end{cases}$$

Класификатор с опорни вектори Support Vector Machine (SVM)

$$\phi(x)^T \phi(x_k) = \psi(x, x_k)$$

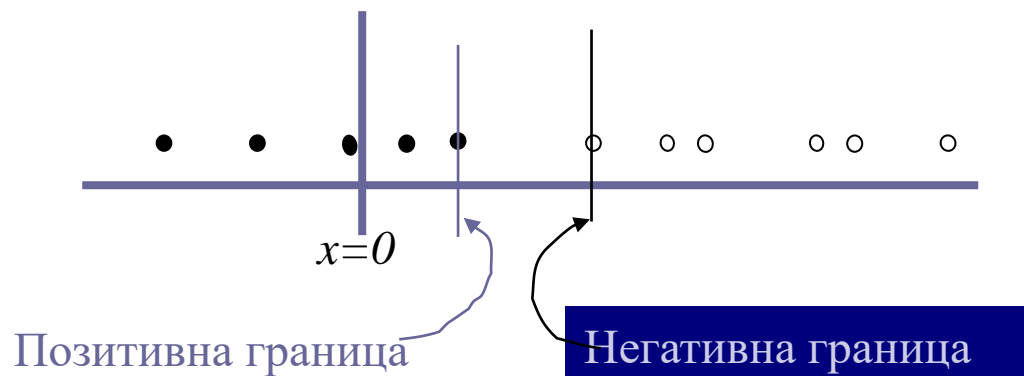
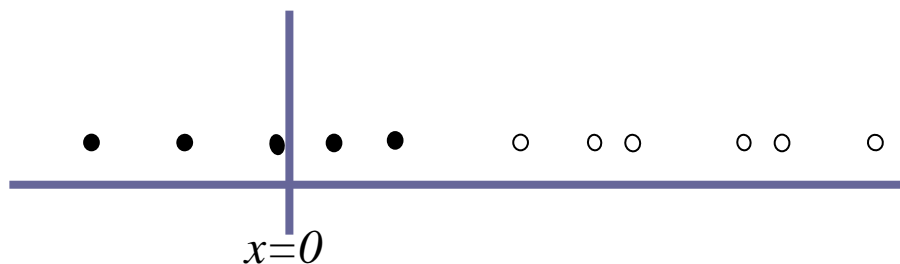
Класификатора се създава след изчисляване на

$$\max_{a_k} Q_1(a_k; \psi(x_k, x_l)) = -\frac{1}{2} \sum_{k,l=1}^N y_k y_l \psi(x_k, x_l) a_k a_l + \sum_{k=1}^N a_k$$



Класификатор с опорни вектори Support Vector Machine (SVM)

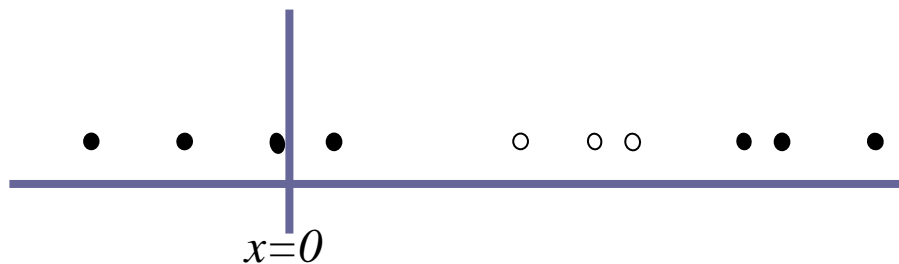
Лесен 1-D пример





Класификатор с опорни вектори Support Vector Machine (SVM)

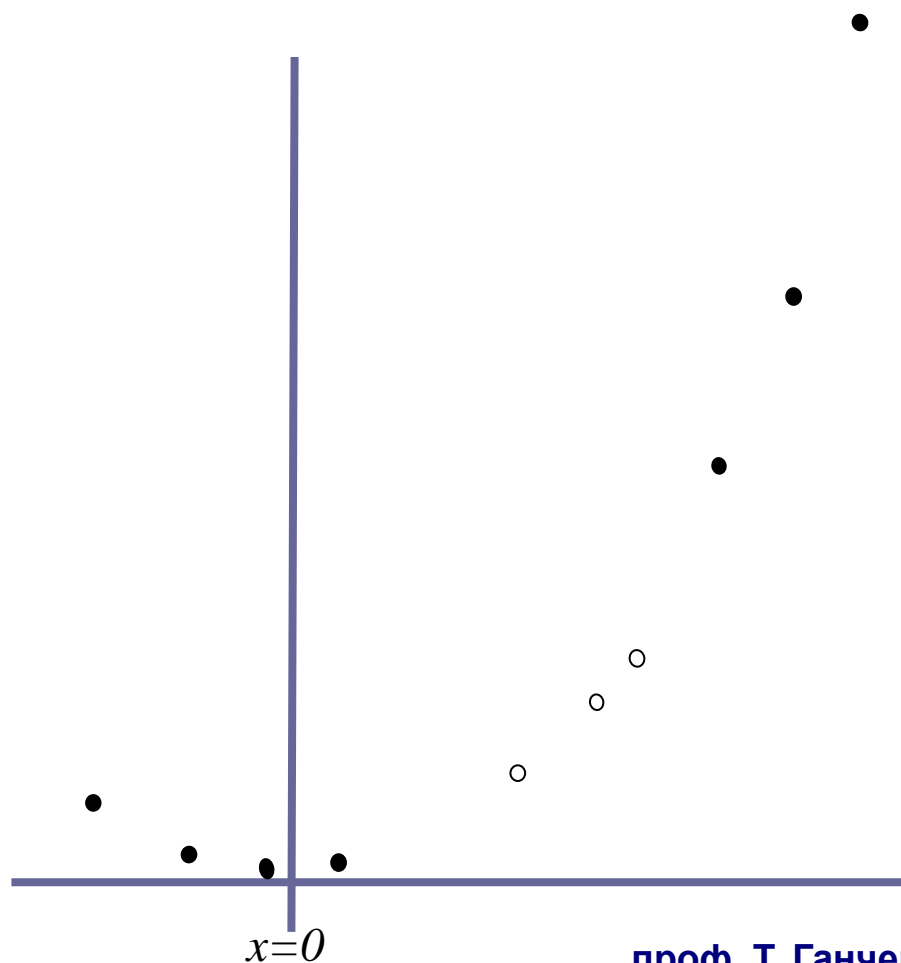
По-труден 1-D пример





Класификатор с опорни вектори Support Vector Machine (SVM)

По-труден 1-D пример

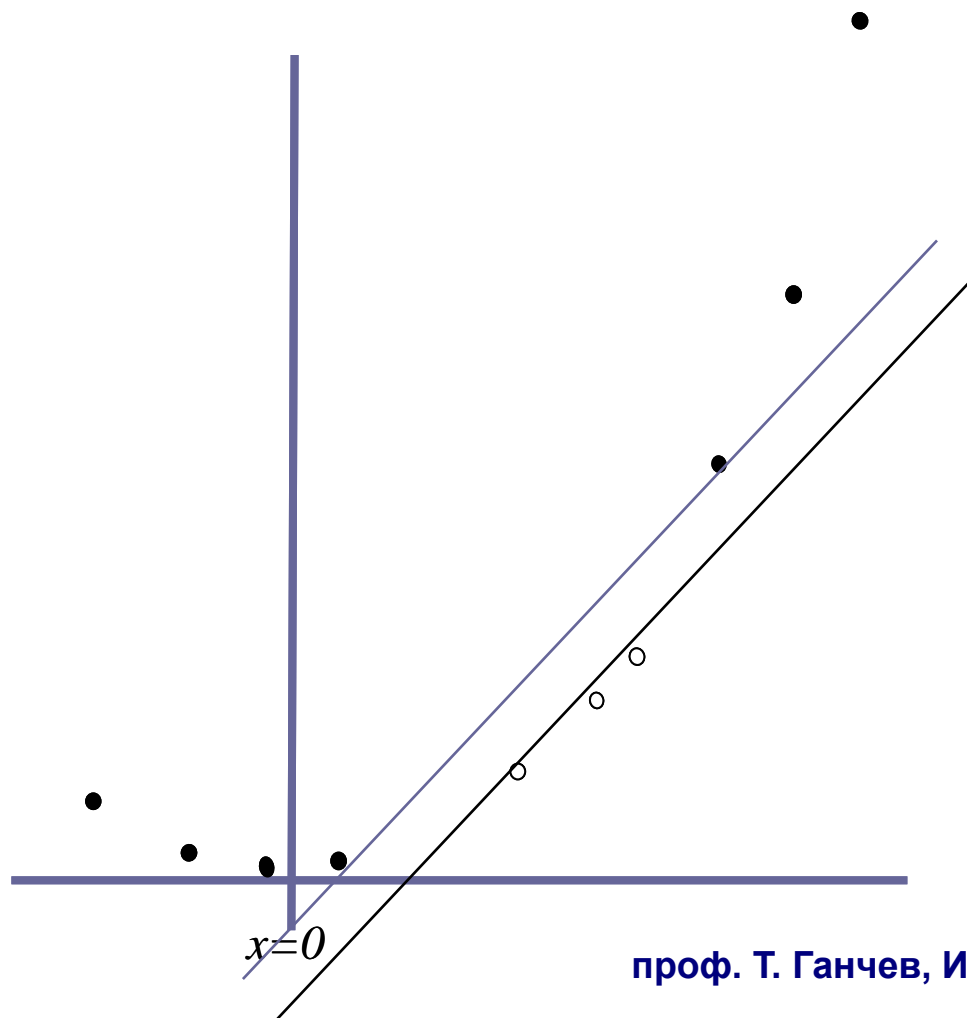


$$\mathbf{z}_k = (x_k, x_k^2)$$



Класификатор с опорни вектори Support Vector Machine (SVM)

По-труден 1-D пример



$$\mathbf{z}_k = (x_k, x_k^2)$$



Въпроси?