

---

# Занятие №2

Классификация.  
Логистическая  
регрессия и SVM



---

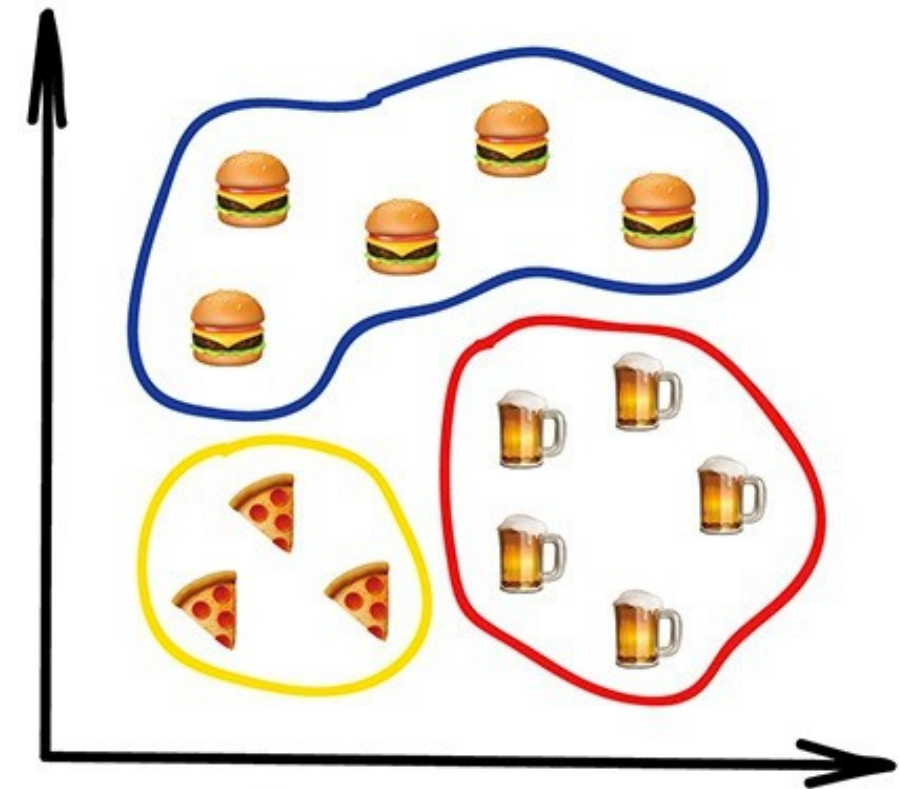
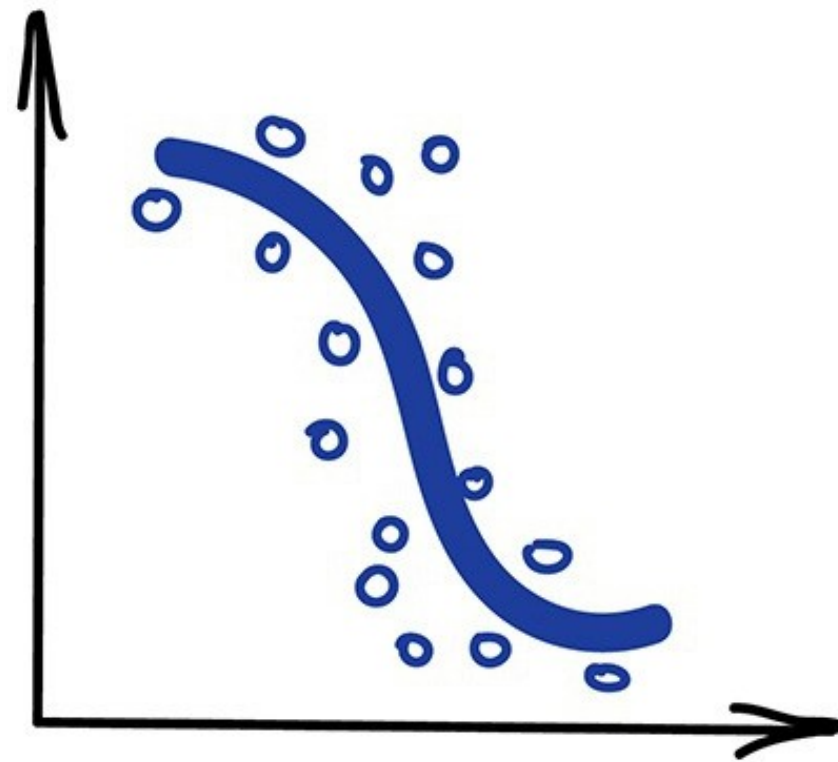
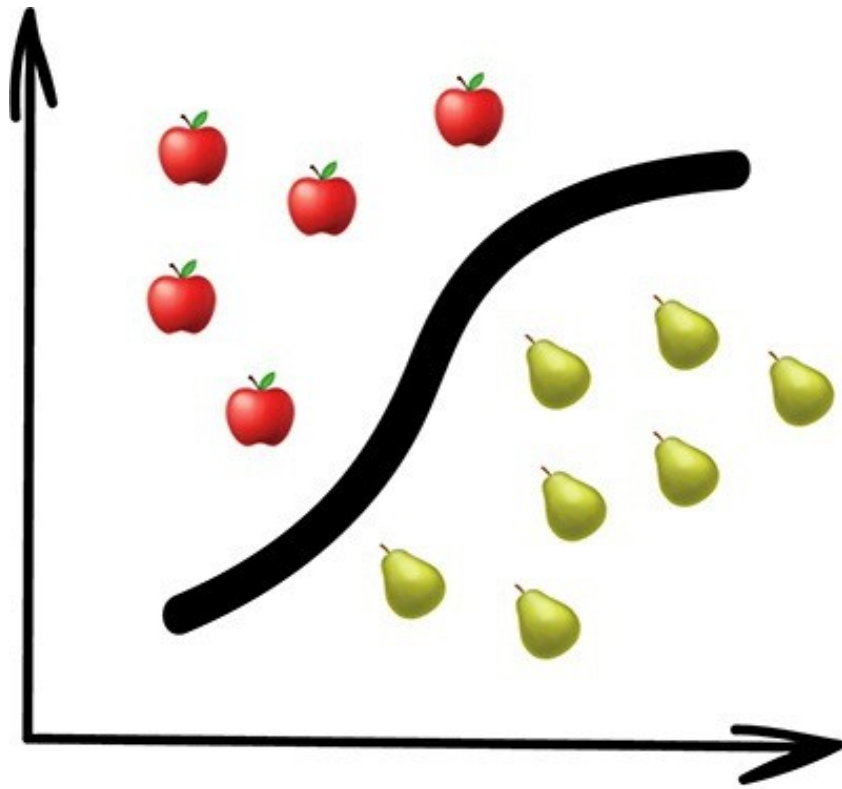
# Содержание

---

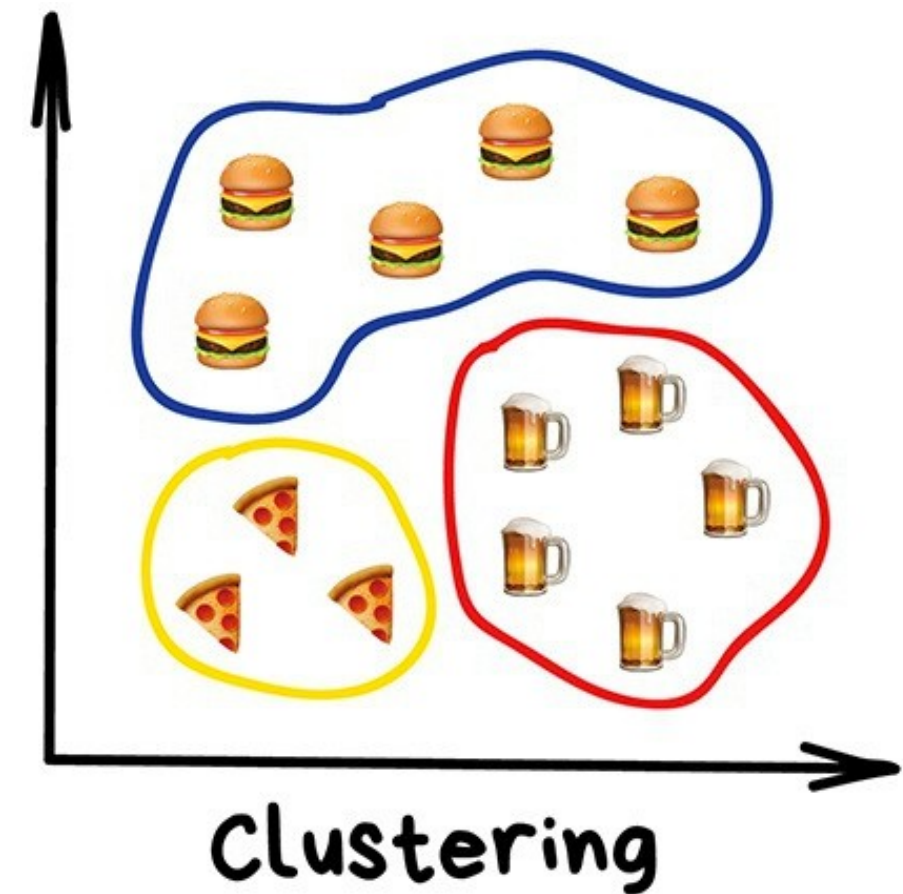
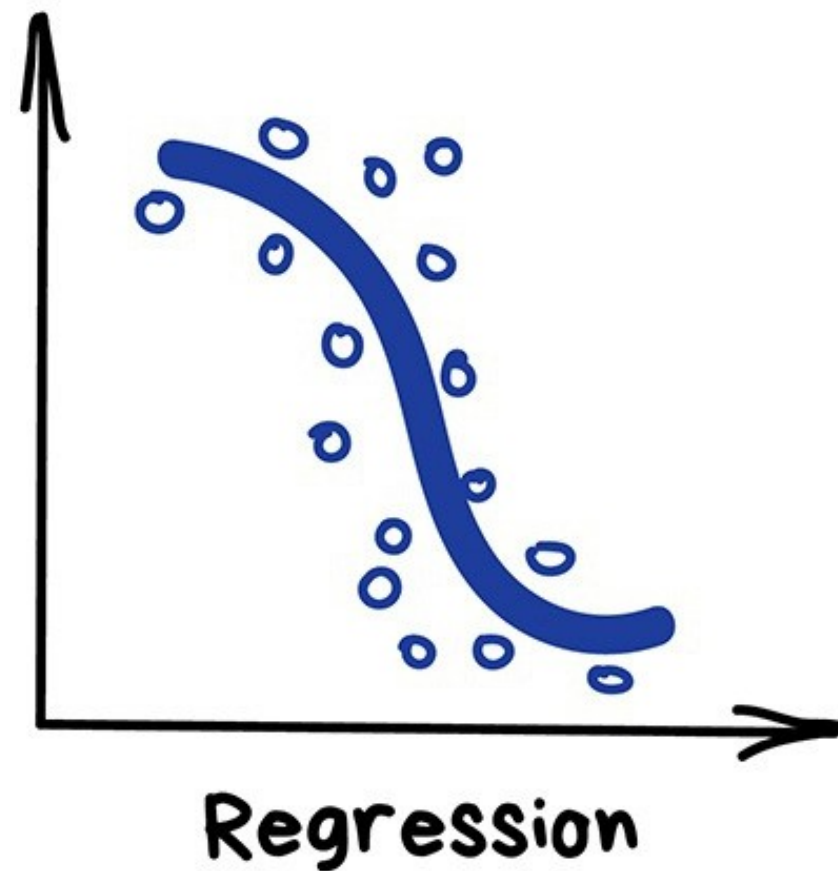
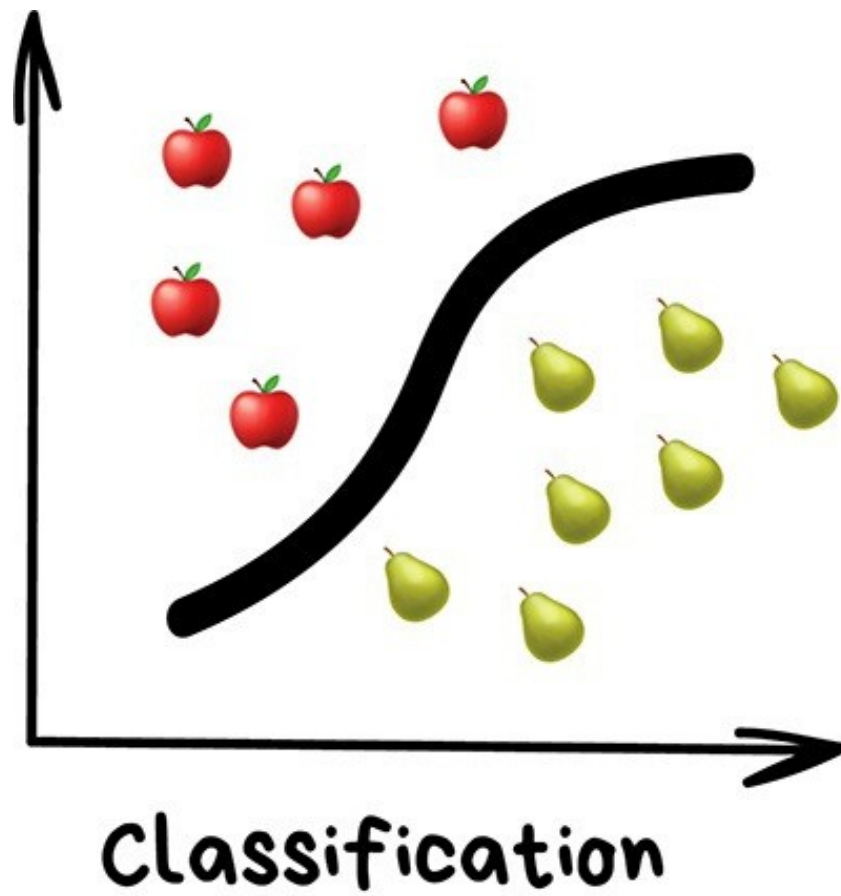
- 1 Введение. Задачи машинного обучения.  
Классификация
- 2 Логистическая регрессия
- 3 SVM. Kernel trick.
- 4 Практика.



# Введение. Задачи машинного обучения



# Введение. Задачи машинного обучения



# Введение. Классификация

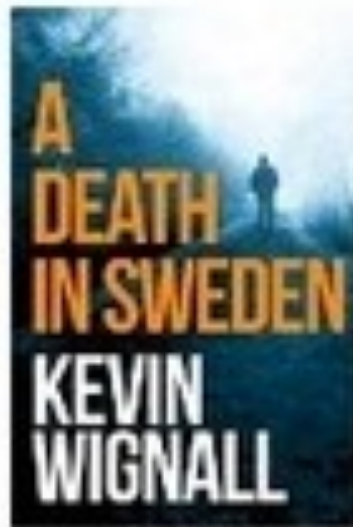




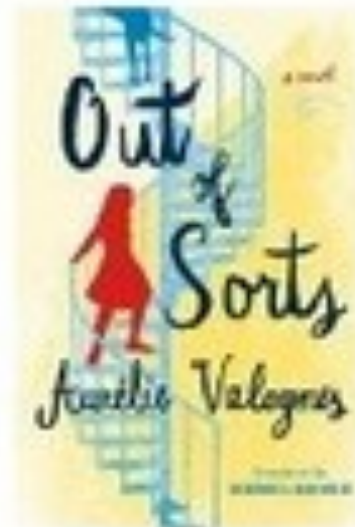
# Введение. Классификация

Your Recently Viewed Items and Featured Recommendations

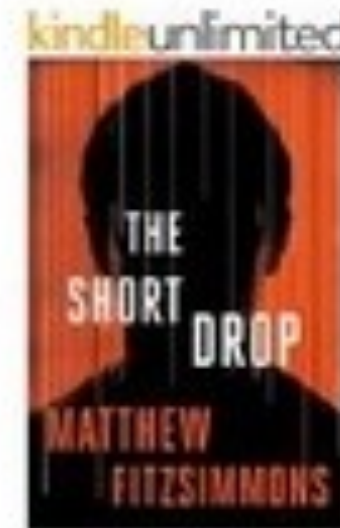
Best Sellers



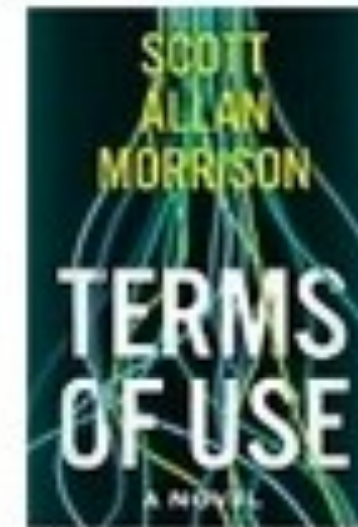
A Death in Sweden  
> Kevin Wignall  
★★★★★ 219  
Kindle Edition  
\$5.99



Out of Sorts  
> Aurélie Valognes  
★★★★★ 118  
Kindle Edition  
\$5.99



The Short Drop  
> Matthew FitzSimmons  
★★★★★ 2,217  
Kindle Edition  
\$5.99



Terms of Use  
> Scott Allan Morrison  
★★★★★ 138  
Kindle Edition  
\$5.99





# Введение. Классификация





# Введение. Классификация

The screenshot displays a video analysis application interface. The main window shows a surveillance camera feed of a hallway. A yellow bounding box highlights a man's face in the foreground. To the right, a sidebar provides detailed information about the identified person:

- Иванов Иван**
- Степень совпадения: **100%**
- Время прохода: **11 сен 18:36:02**
- Канал: **FC2D NC-K 10.0.15.11[10.0.15.11]**
- Снимок: **1 из 18**

Below this information is a row of small thumbnail images showing different frames of the person. Further down, a section titled "Результат идентификации" (Identification Result) shows a small image of the person and the text "Иванов 100%".

At the bottom of the interface, there is a timeline of events with several small thumbnail images and timestamps, such as "11 сен 18:35:47", "11 сен 18:35:51", "11 сен 18:35:54", "11 сен 18:35:55", "11 сен 18:35:57", "11 сен 18:35:59", "11 сен 18:36:00", and "11 сен 18:36:02".

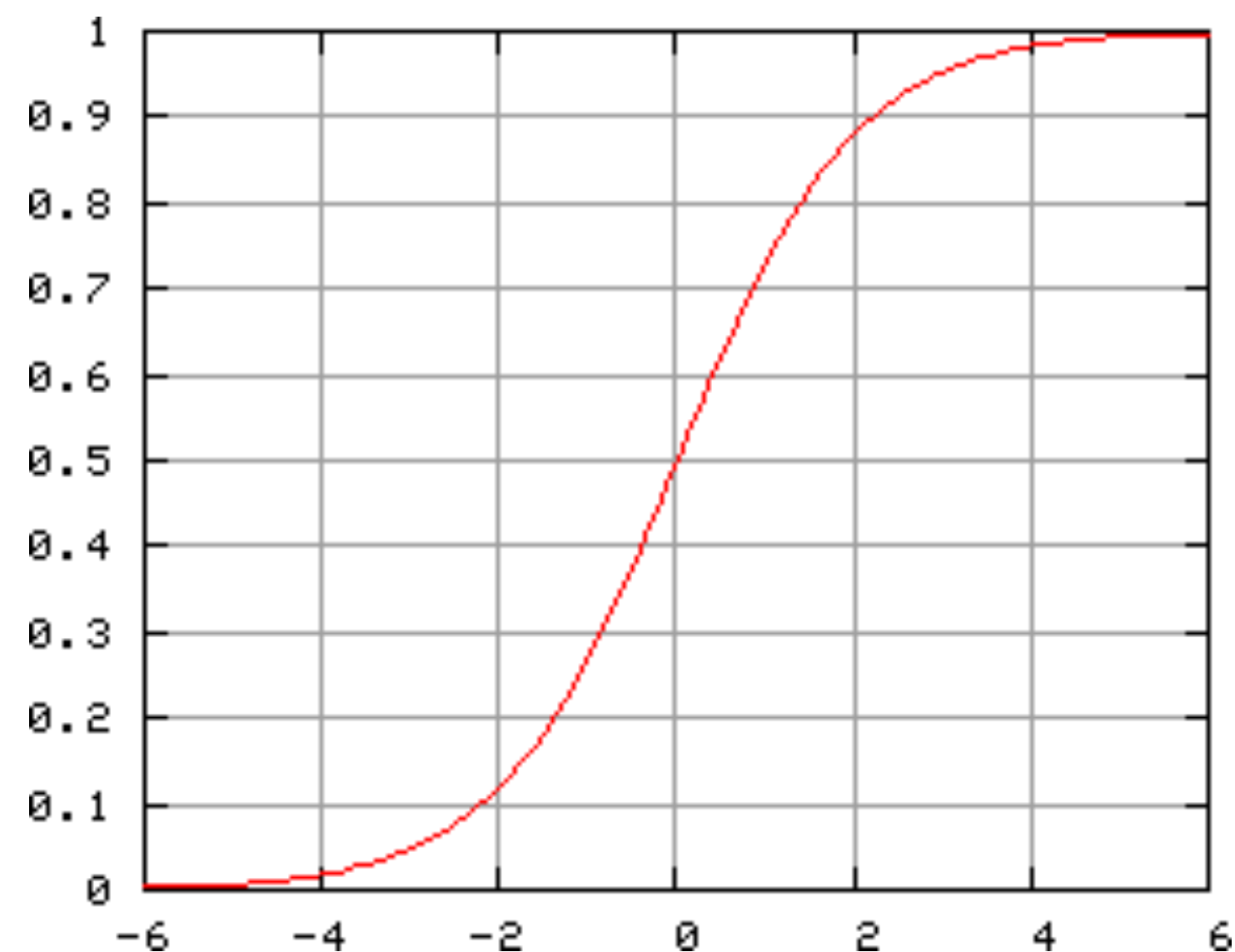




# Логистическая регрессия

применяется для прогнозирования вероятности возникновения некоторого события по значениям множества признаков

$$f(z) = \frac{1}{1 + e^{-z}}$$



# Логистическая регрессия

Знакомо? ;)

$$z = \theta^T x = \theta_0 + \theta_1 x_1 + \dots + \theta_n x_n$$

$$f(z) = \frac{1}{1 + e^{-z}},$$





# Логистическая регрессия. Максимизация правдоподобия

$$\hat{\theta} = \operatorname{argmax}_{\theta} L(\theta) = \operatorname{argmax}_{\theta} \prod_{i=1}^m \mathbb{P}\{y = y^{(i)} \mid x = x^{(i)}\}$$

Чтобы модель могла обучаться, ей необходимо получать «штраф» за то, что она ошибается.



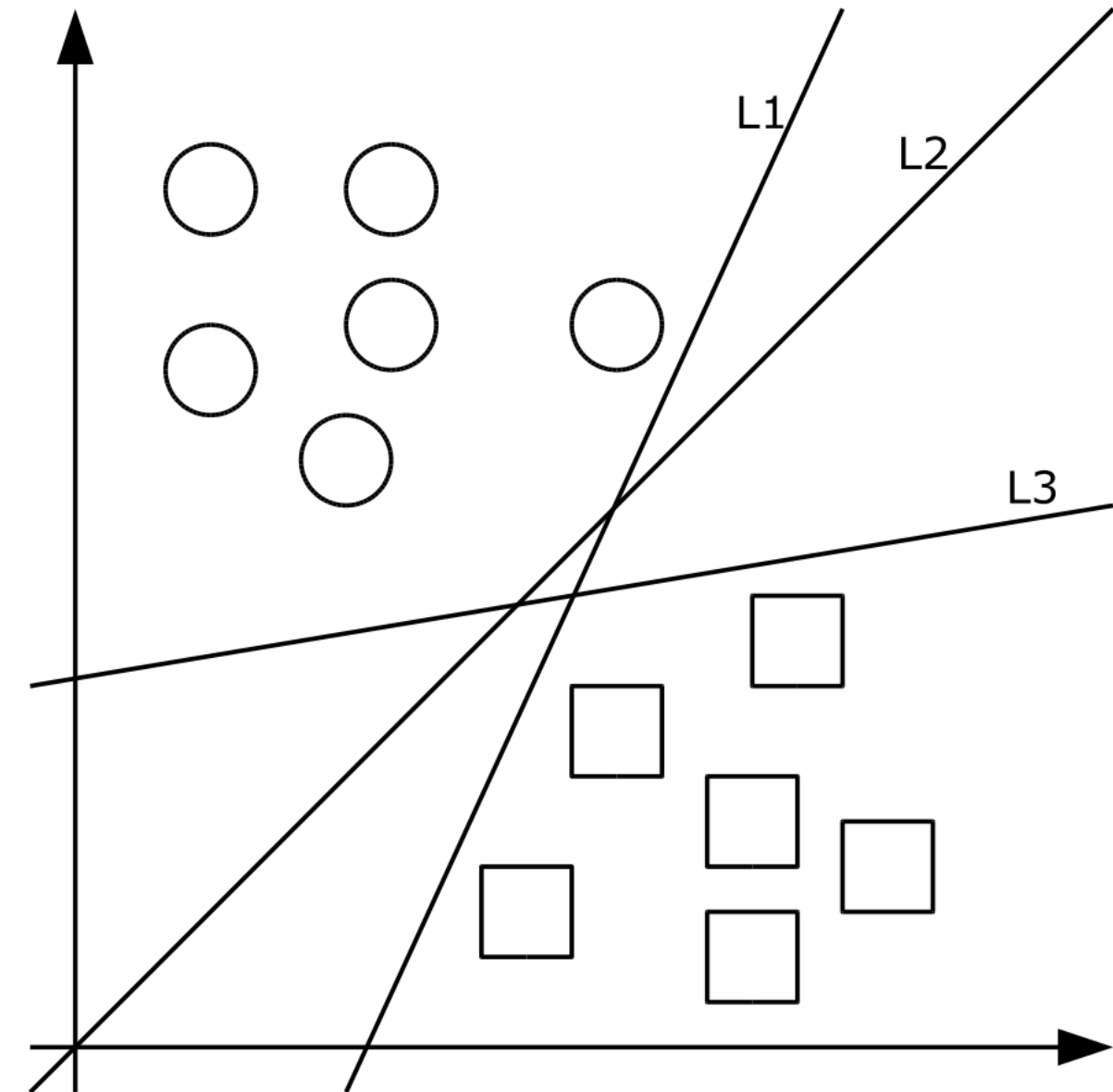
$$\ln L(\theta) = \sum_{i=1}^m \log \mathbb{P}\{y = y^{(i)} \mid x = x^{(i)}\} = \sum_{i=1}^m y^{(i)} \ln f(\theta^T x^{(i)}) + (1 - y^{(i)}) \ln(1 - f(\theta^T x^{(i)})),$$



# Метод опорных векторов. SVM

перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве.

$$r = y \frac{\mathbf{w}^T \mathbf{x} + b}{\|\mathbf{w}\|}$$

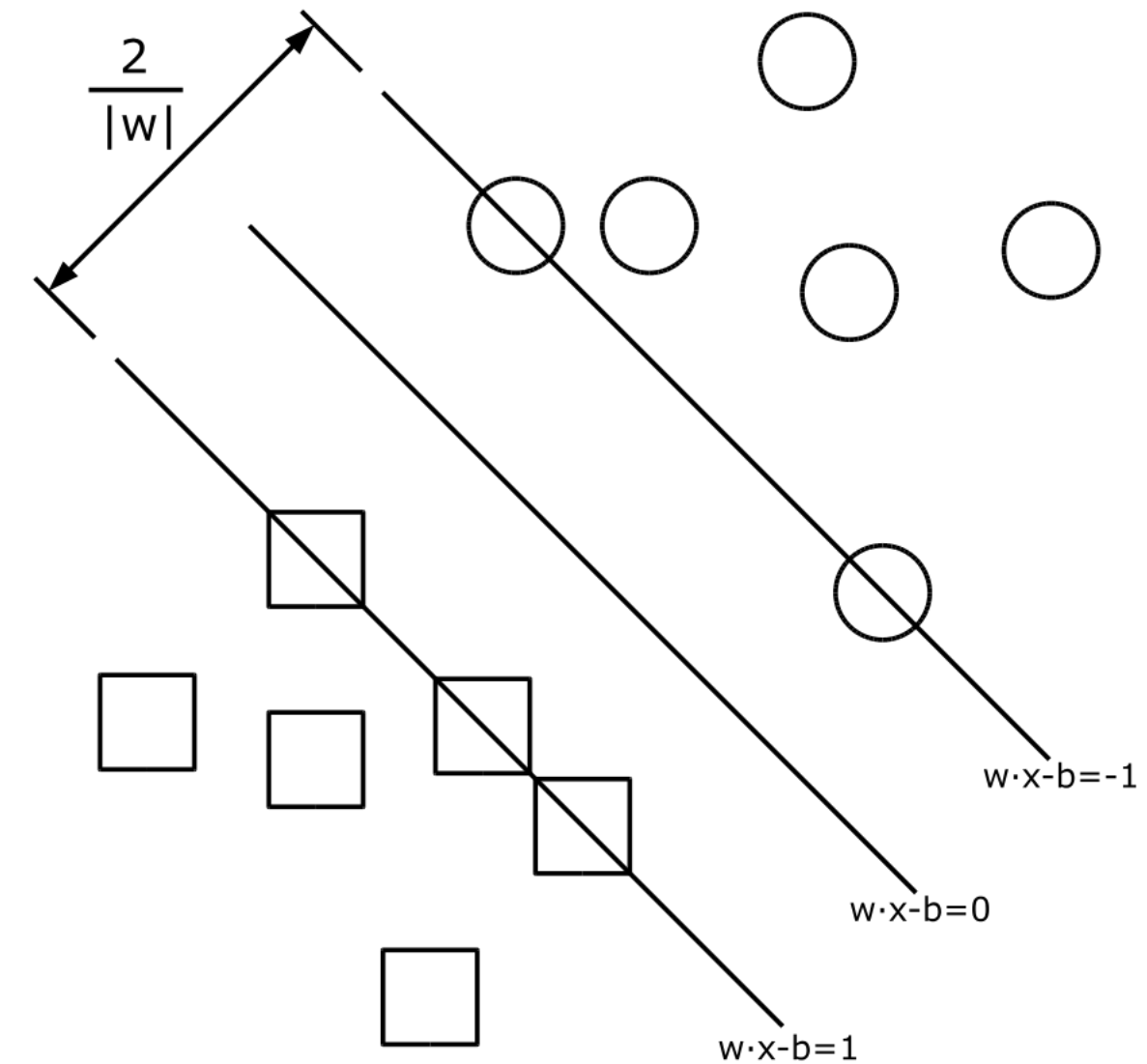




# Метод опорных векторов. SVM

Две параллельных гиперплоскости строятся по обеим сторонам гиперплоскости, разделяющей классы. Разделяющей гиперплоскостью будет гиперплоскость, максимизирующая расстояние до двух параллельных гиперплоскостей.

Алгоритм работает в предположении, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора.



# Метод опорных векторов. SVM. Решение задачи.

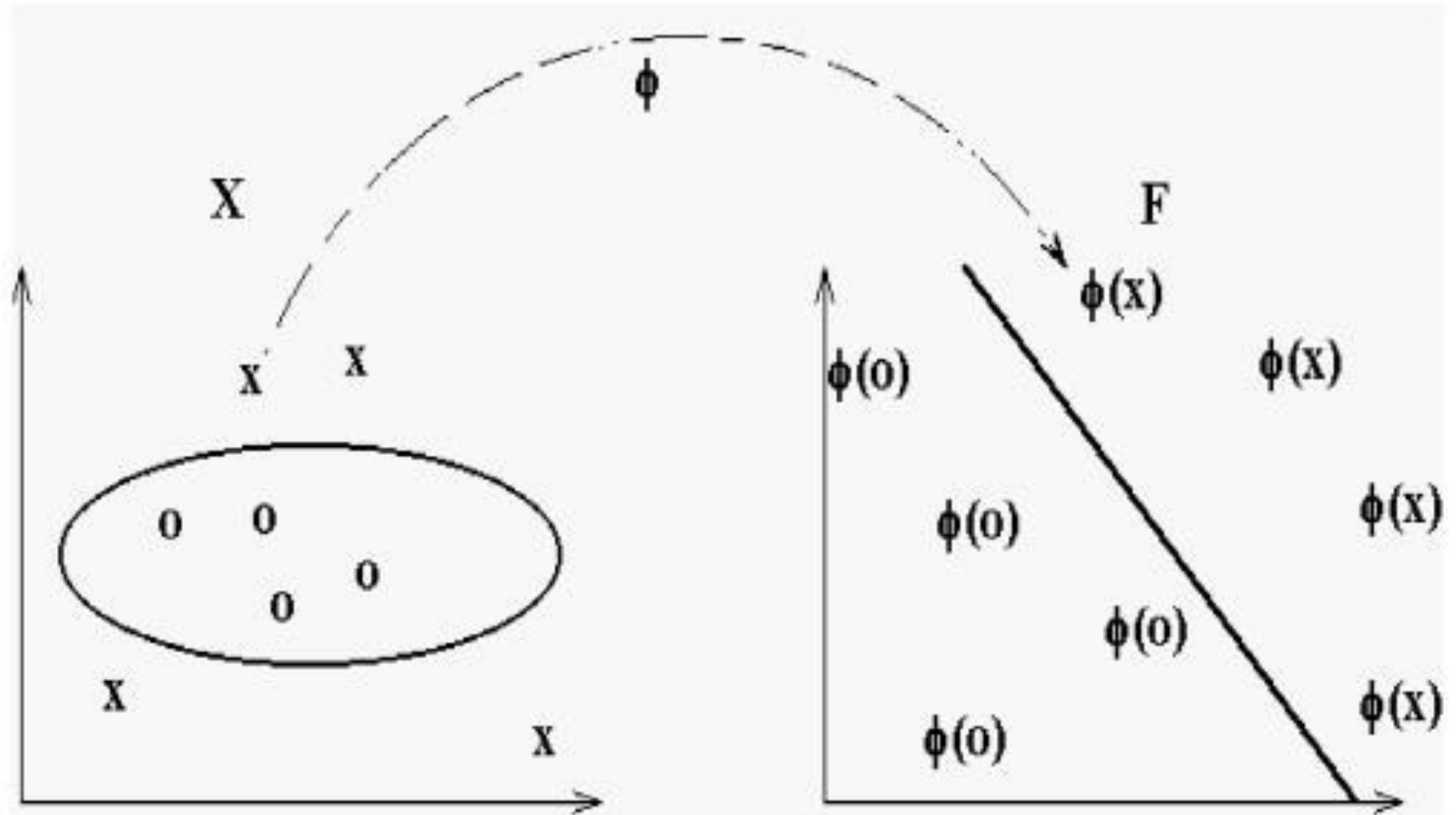
$$\begin{cases} \mathbf{L}(\mathbf{w}, \mathbf{b}; \lambda) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \lambda_i (c_i ((\mathbf{w} \cdot \mathbf{x}_i) - b) - 1) \rightarrow \min_{w,b} \max_{\lambda} \\ \lambda_i \geq 0, \quad 1 \leq i \leq n \end{cases}$$





# Метод опорных векторов. Полиномиальное ядро

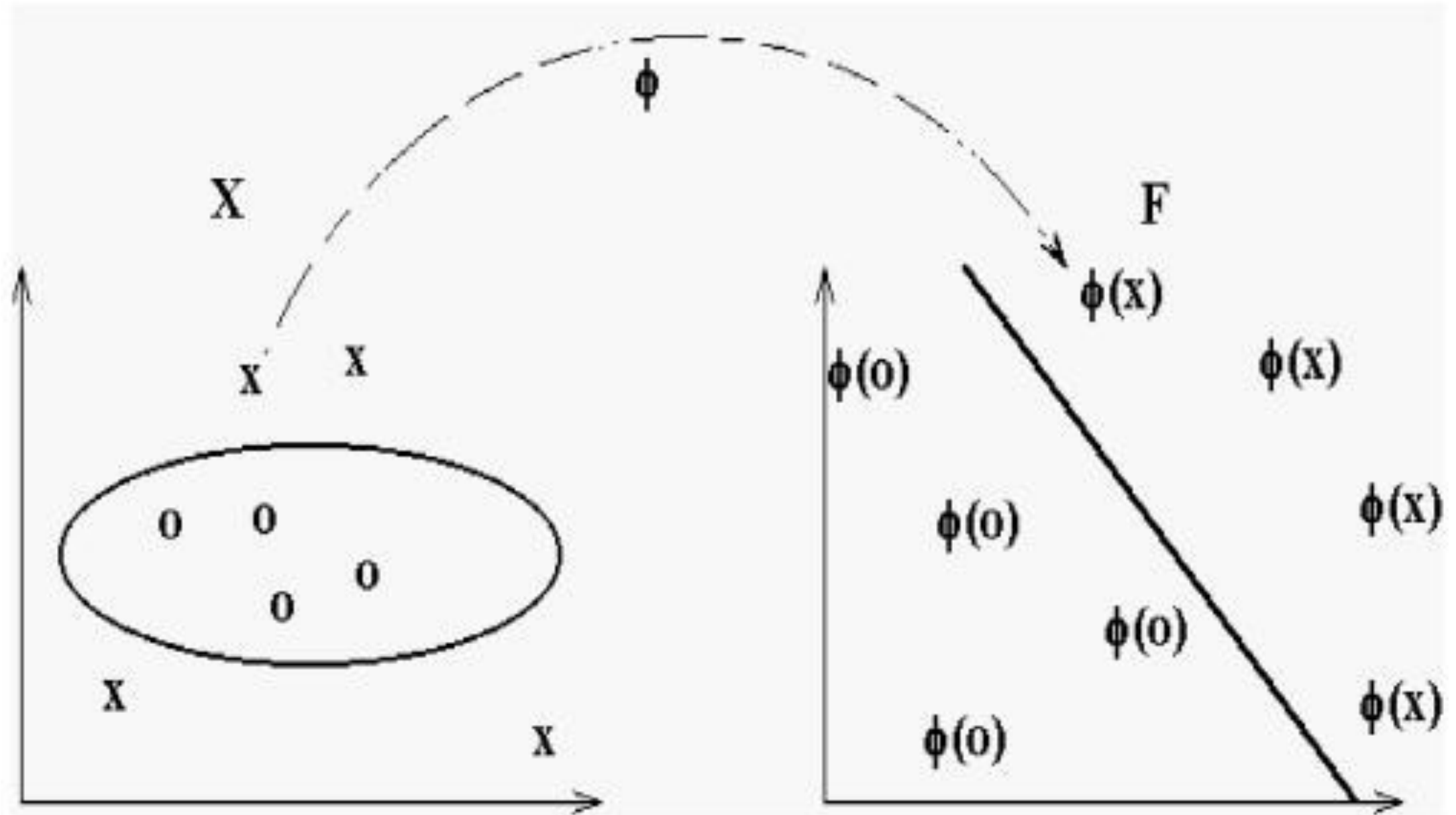
$$k(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}')^d$$



# Метод опорных векторов. Полиномиальное ядро

$$k(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}')^d$$

$$k(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}' + 1)^d$$

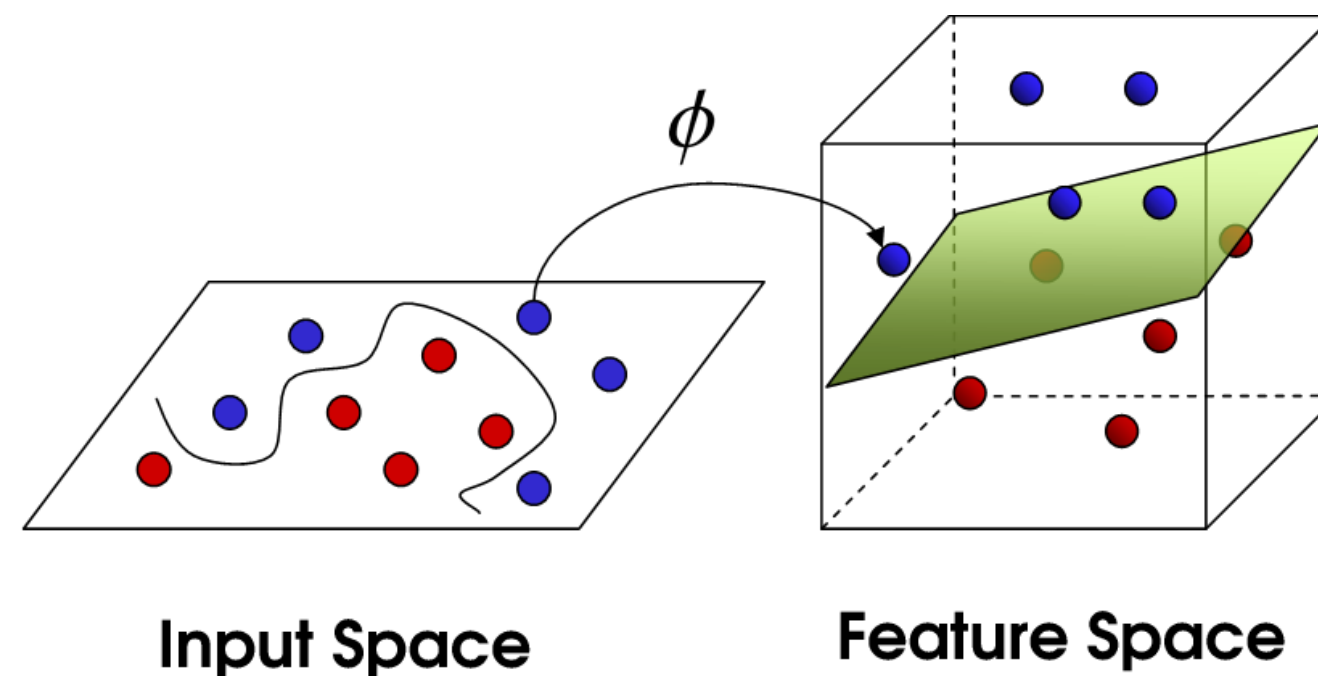




# Метод опорных векторов. Другие разновидности ядер

Радиальная базисная функция  $k(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2)$ , для  $\gamma > 0$

Радиальная базисная функция Гаусса  $k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right)$



**ПРАКТИКА**



---

Спасибо за  
внимание!

---

