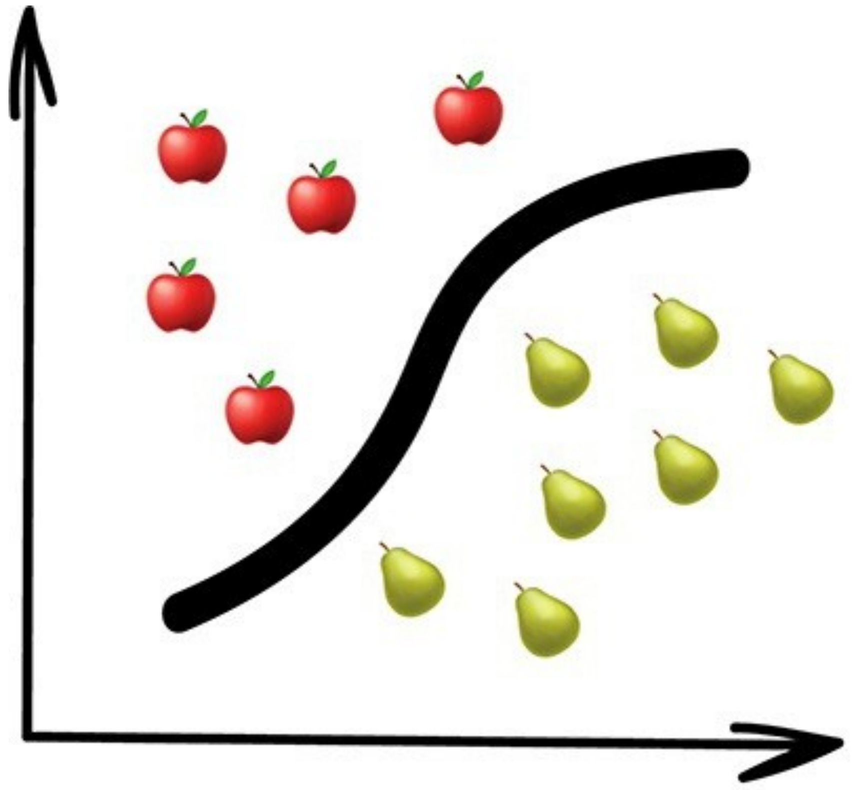
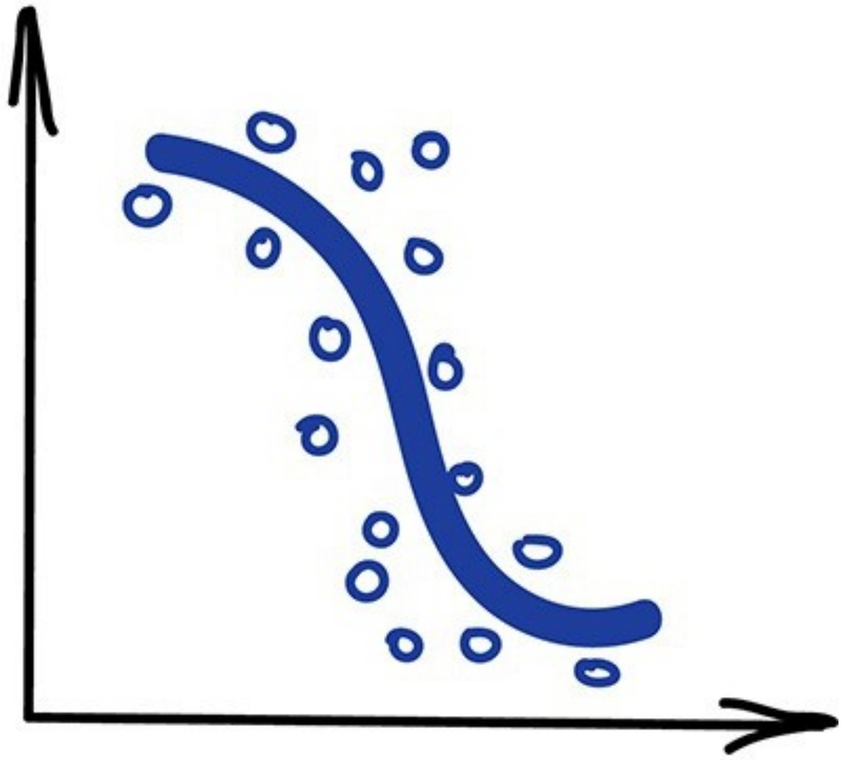
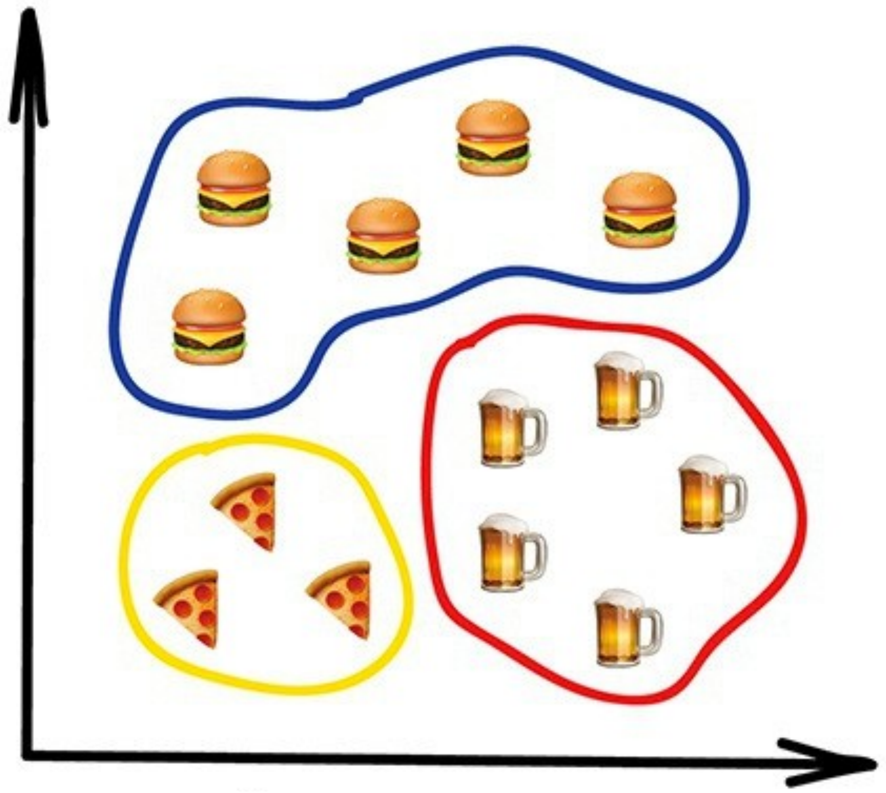
Занятие №2

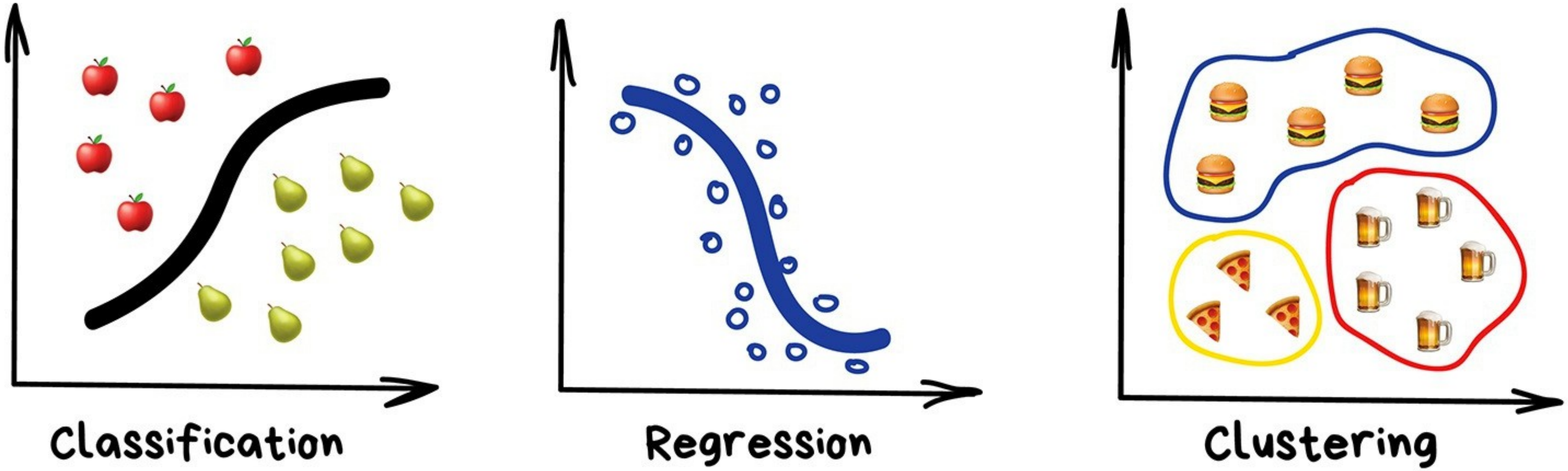
Классификация. Логистическая регрессия и SVM

Содержание

Введение. Задачи машинного обучения

Введение. Задачи машинного обучения

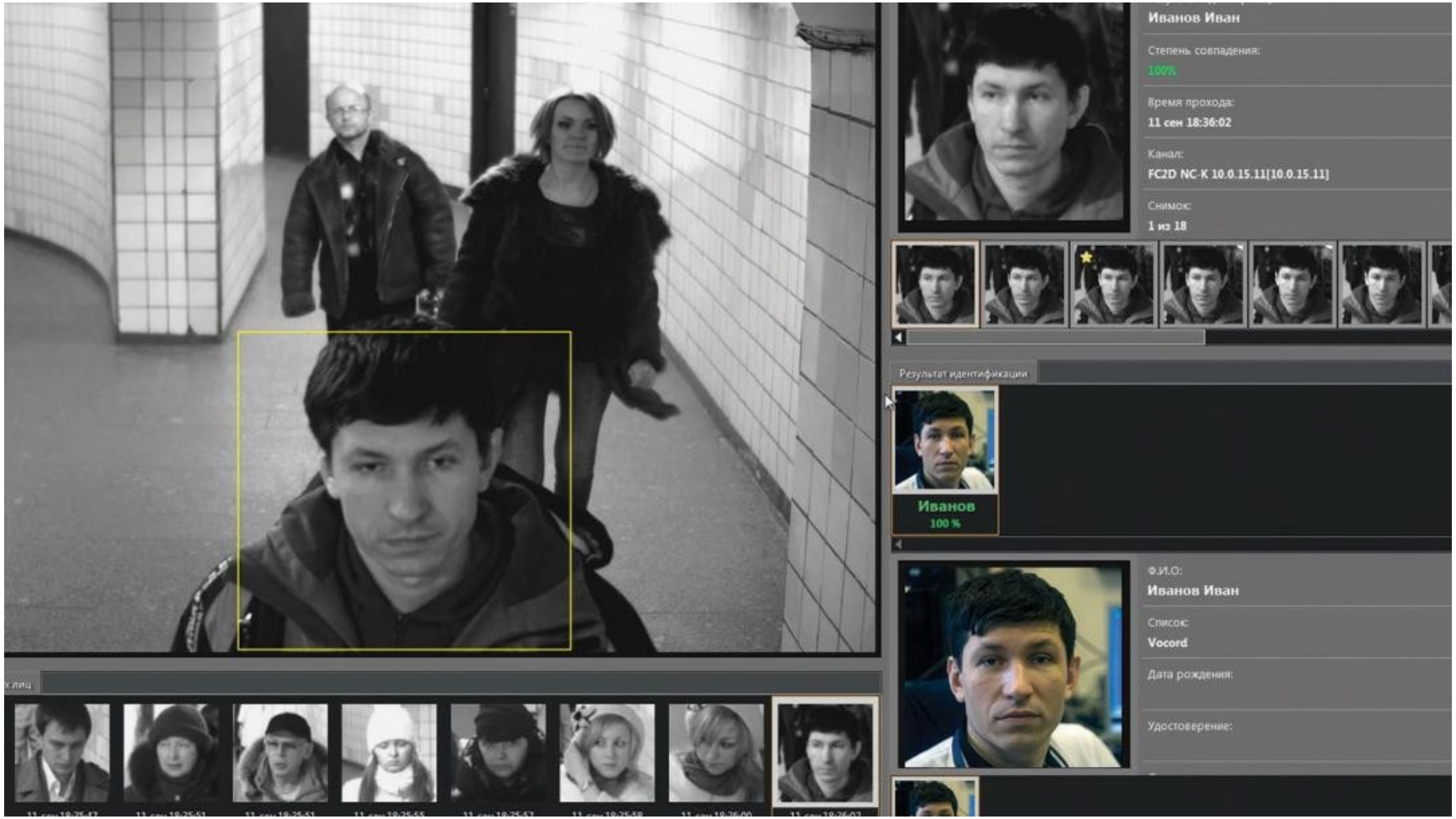


Введение. Классификация

Введение. Классификация

Введение. Классификация

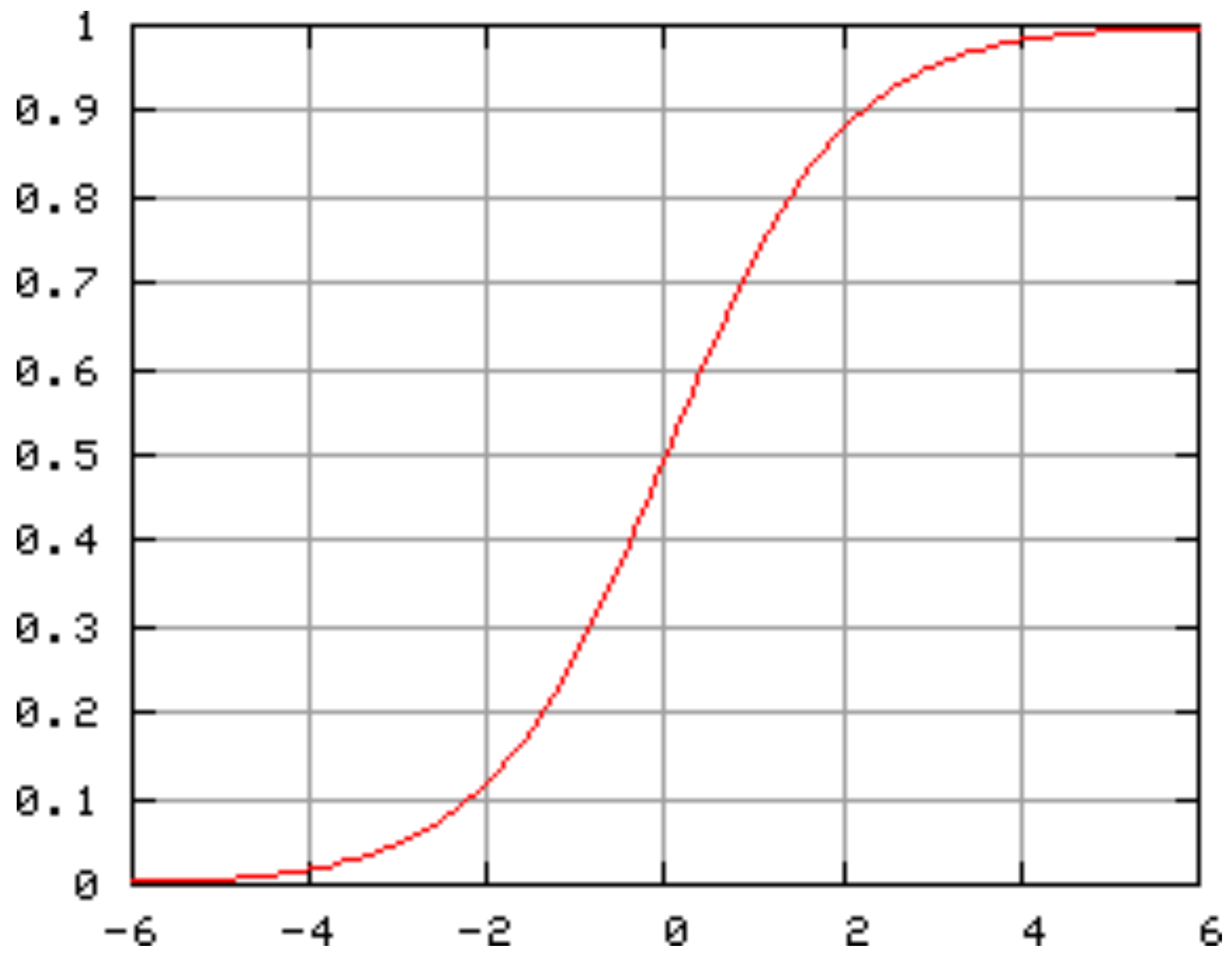
Введение. Классификация



Логистическая регрессия

применяется для прогнозирования вероятности возникновения некоторого события по значениям множества признаков

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



Логистическая регрессия

Знакомо? ;)

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

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

Логистическая регрессия. Максимизация правдоподобия

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

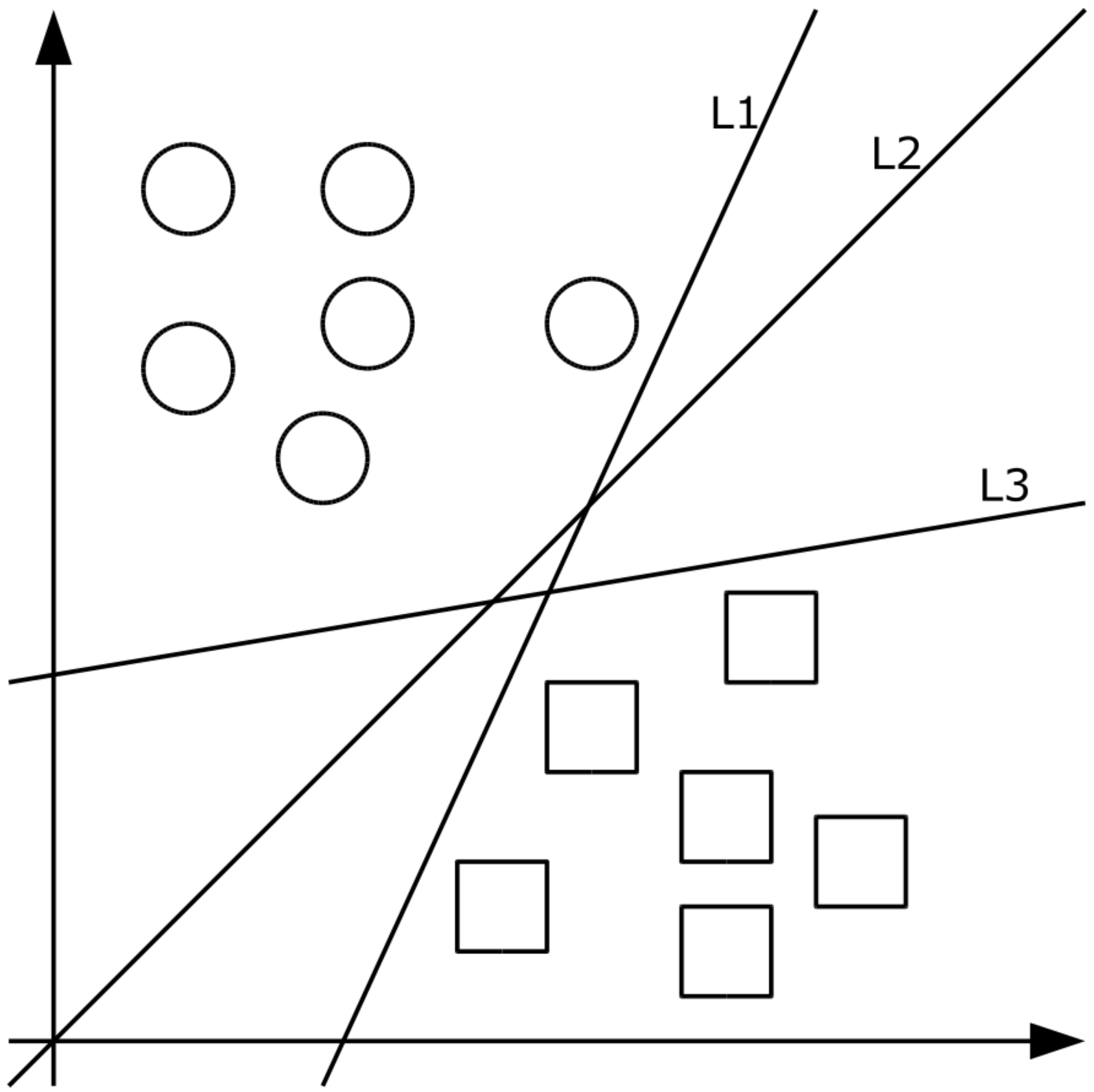
Чтобы модель могла обучаться, ей необходимо получать «штраф» за то, что она ошибается.

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

Метод опорных векторов. SVM

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

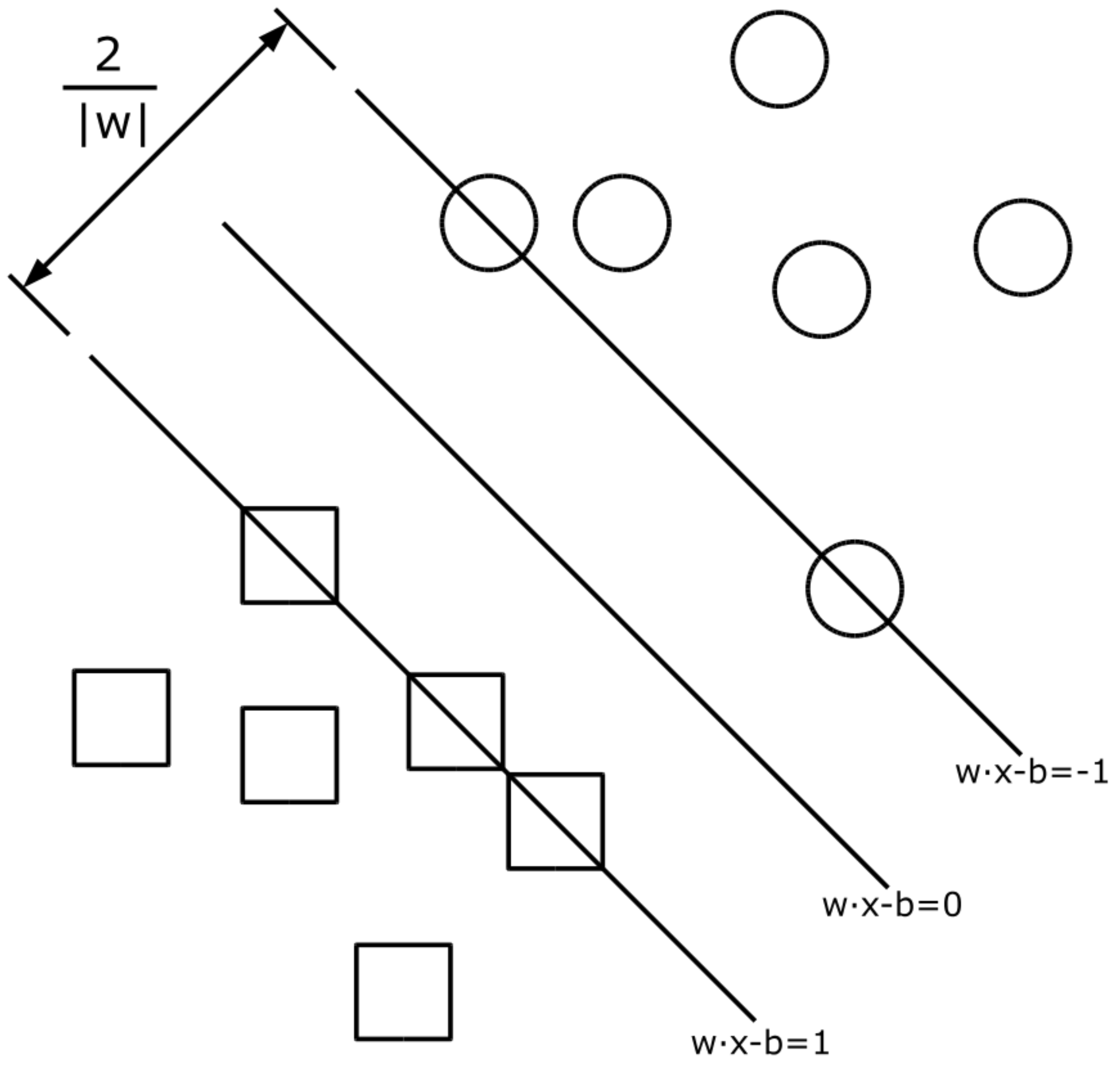
$$r = y\frac{\mathbf{w}^T\mathbf{x} + b}{\|\mathbf{w}\|}$$



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.

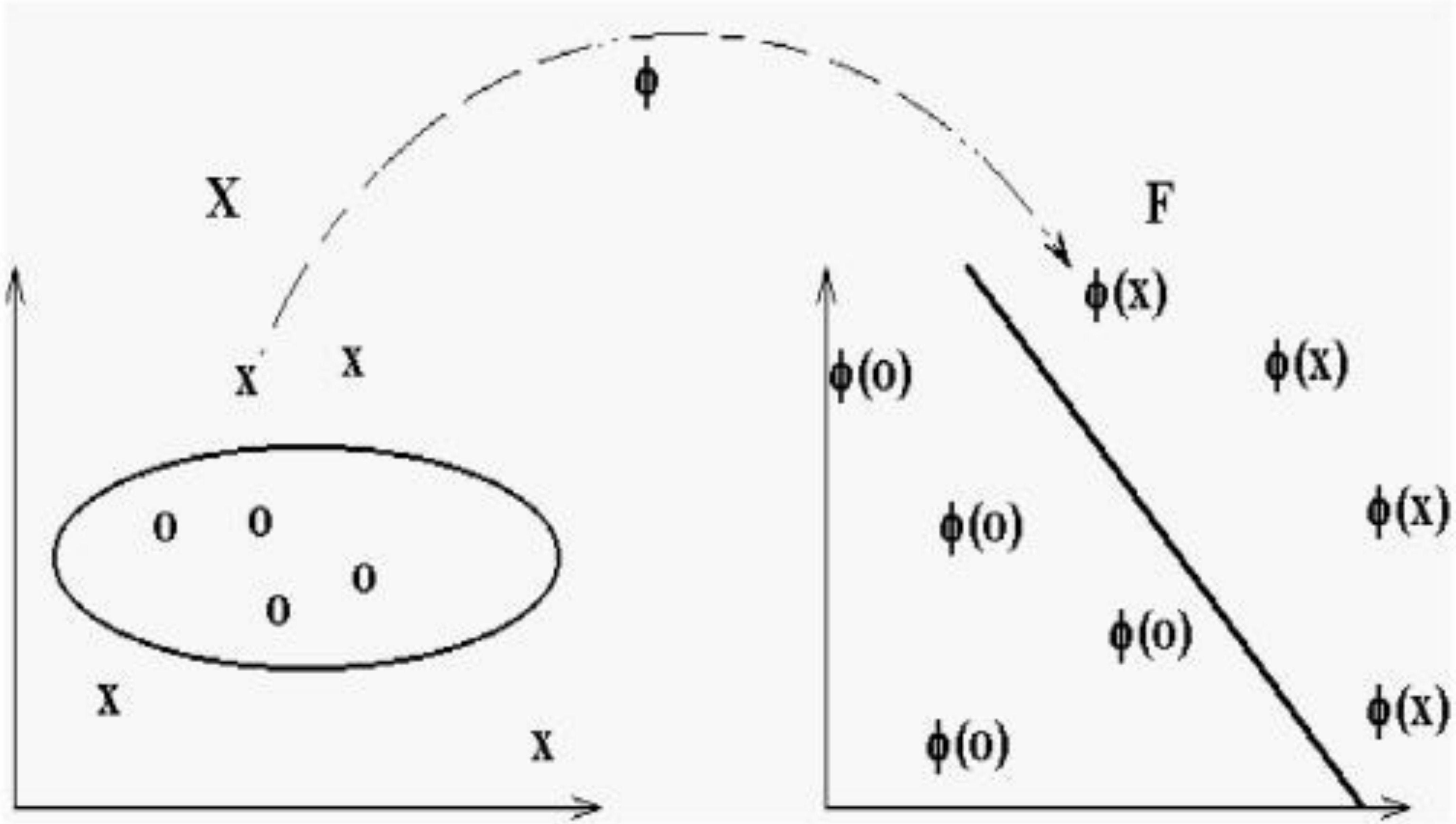


Метод опорных векторов. SVM. Решение задачи.

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

Метод опорных векторов. Полиномиальное ядро

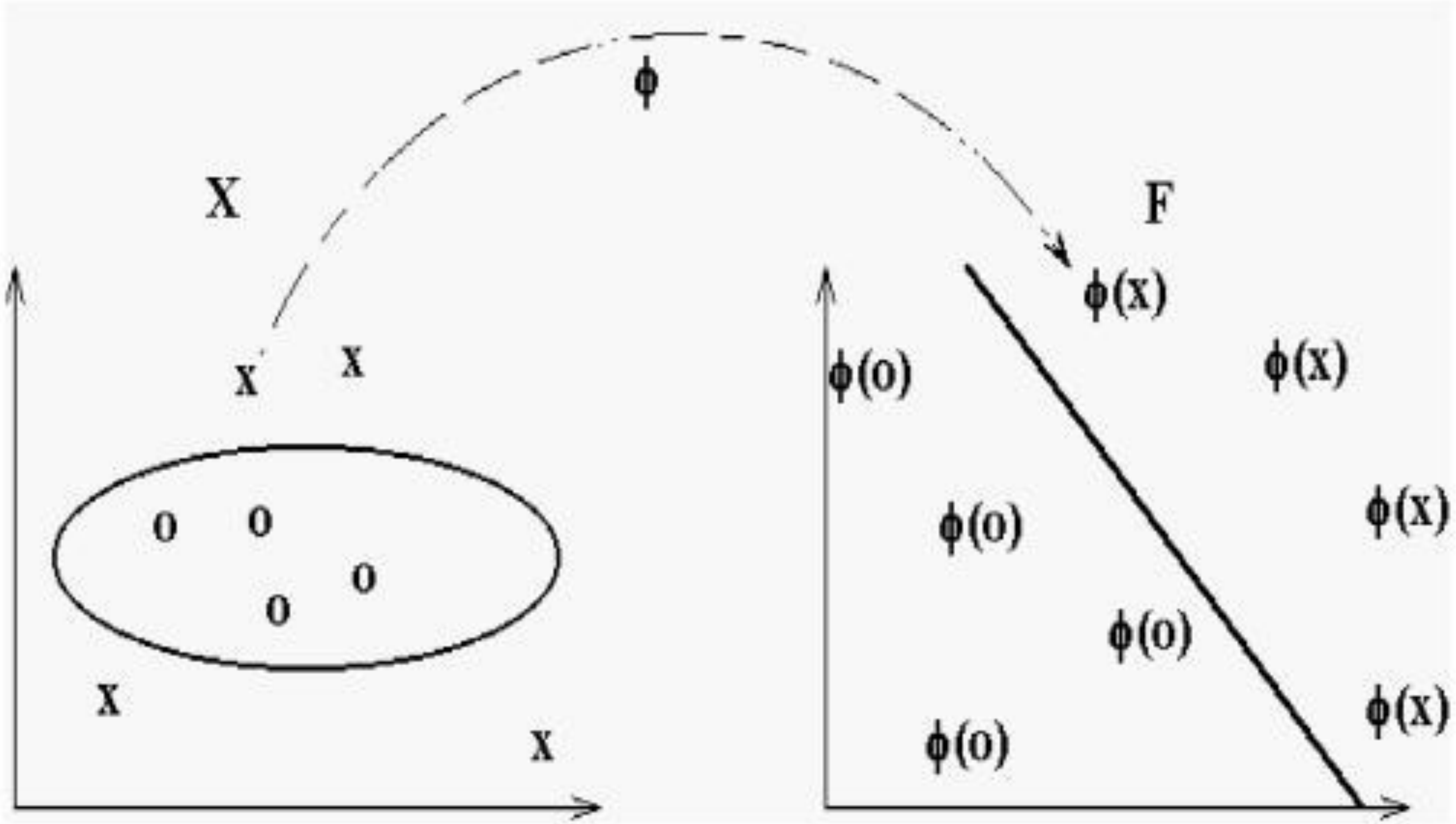
$$k(\mathbf{x}, \mathbf{x}') = (\mathbf{x} \cdot \mathbf{x}')^d$$



Метод опорных векторов. Полиномиальное ядро

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

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

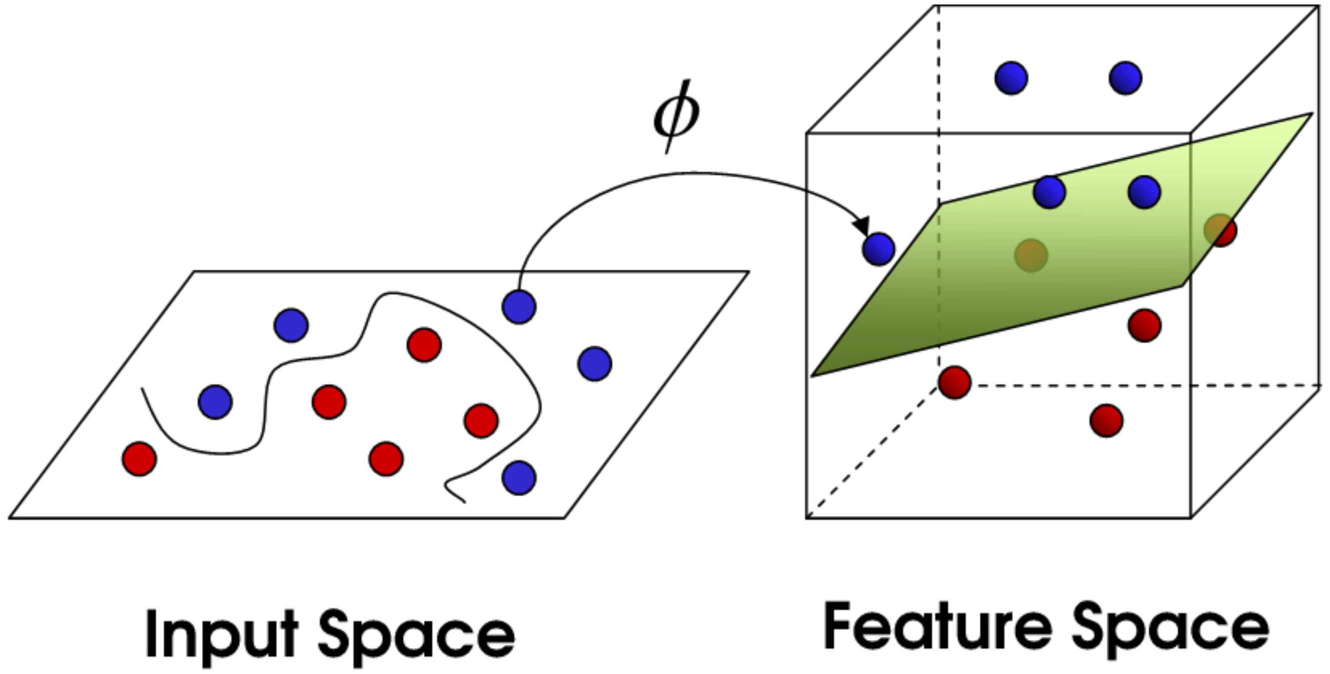


Метод опорных векторов. Другие разновидности ядер

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!