

Нейронные сети и их практическое применение.

Лекция 3. Проблема «Исключающего или». Многослойный персептрон

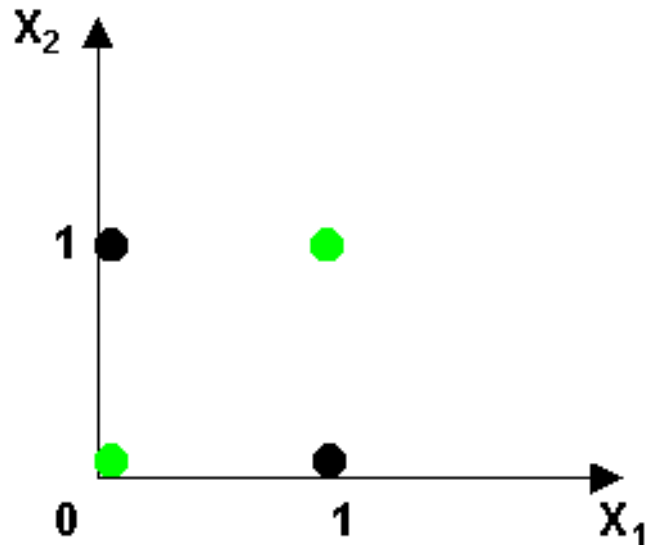
Дмитрий Буряк
к.ф.-м.н
dyb04@yandex.ru

Однослойная НС.

Проблема «исключающего ИЛИ»

Функция XOR:

X_1	0	0	1	1
X_2	0	1	0	1
XOR	0	1	1	0

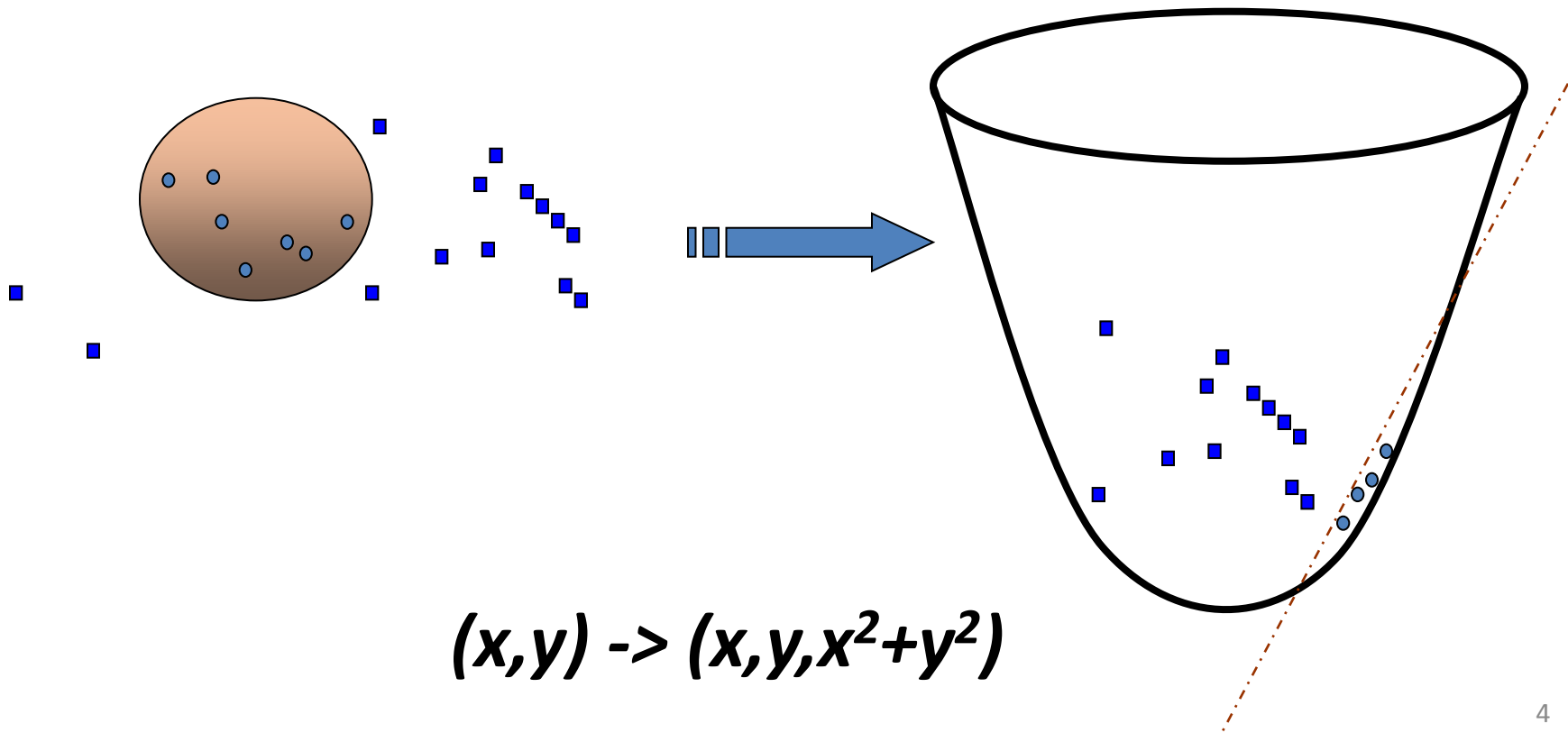


Невозможно построить однослойную НС, реализующую функцию XOR.

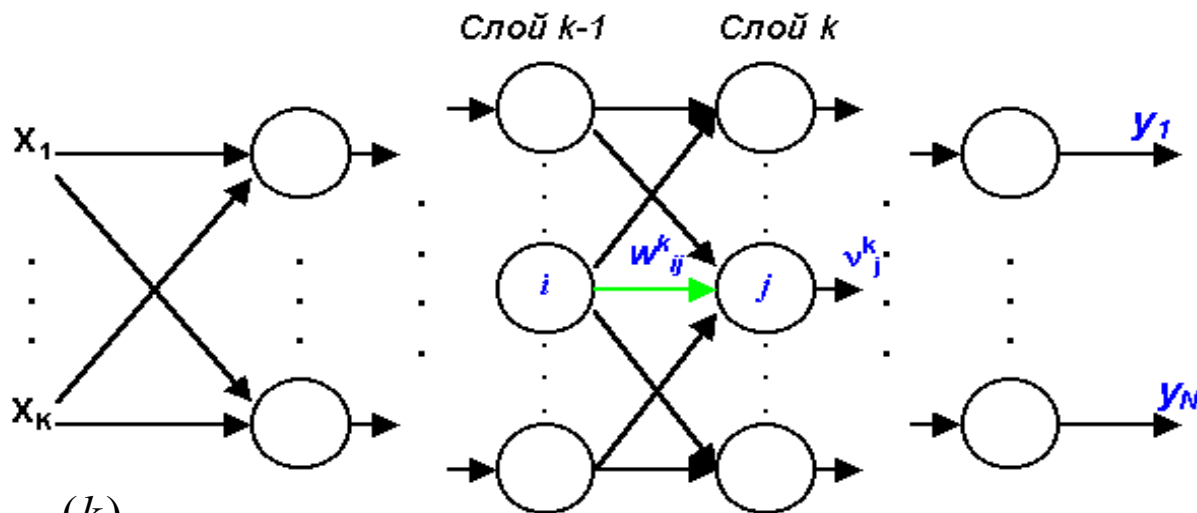
Применение линейных классификаторов

1. Предобработка входных образов для получения линейно разделимых классов.
2. SVM.
3. Специальные сети (SNoW - Sparse Network of Winnows).
4. AdaBoost

Пример. Линеаризация



Многослойный персептрон



$w_{ij}^{(k)}$ - вес связи от нейрона i слоя $k-1$ к нейрону j слоя k ;

$W^{(k)} = (w_{ij}^{(k)})$ - матрица весов для слоя k ;

$u_i^{(k)}$ - л/к нейрона i слоя k ;

$v_j^{(k)}$ - выход нейрона j скрытого слоя k ;

y_i - выход нейрона i выходного слоя.

Необходимость нелинейности функции активации

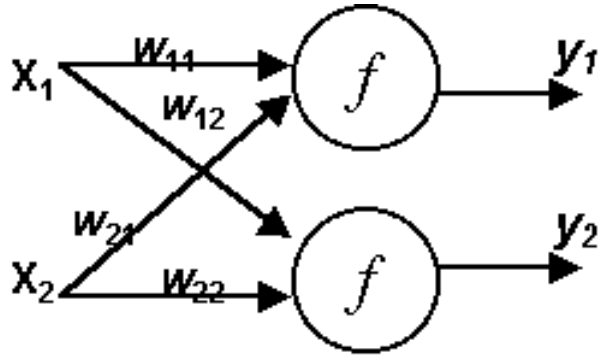
Предположим, что функция активации линейна

$$u^{(k-1)} = xW^{(k-1)}$$

$$y^{(k-1)} = f(u^{(k-1)}) = f(xW^{(k-1)}) = \alpha xW^{(k-1)}$$

$$u^{(k)} = y^{(k-1)}W^{(k)} = \alpha xW^{(k-1)}W^{(k)} = \alpha xW$$

Решение проблемы «исключающего ИЛИ»



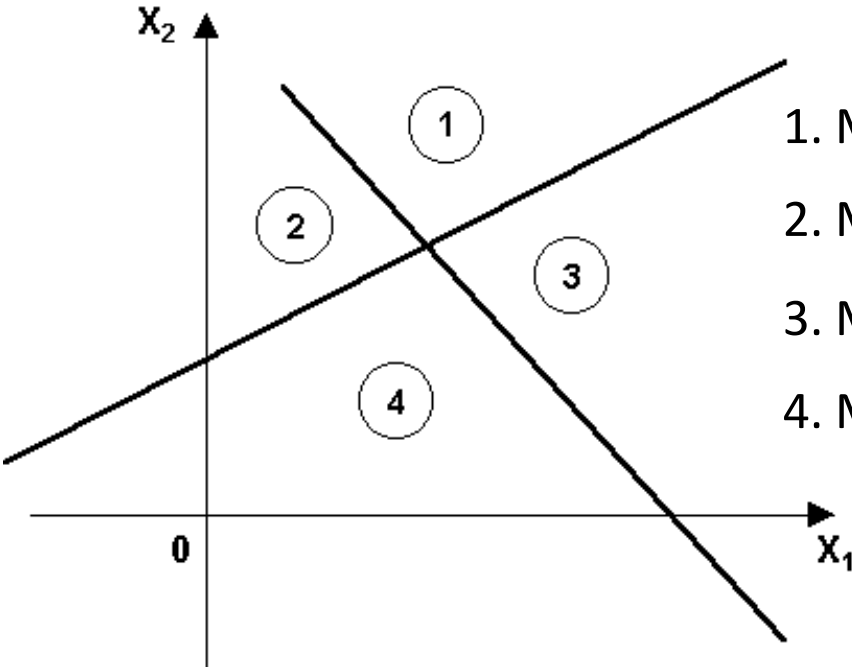
$$u_1 = w_{11}x_1 + w_{21}x_2 + w_{01}$$

$$u_2 = w_{12}x_1 + w_{22}x_2 + w_{02}$$

$$y_1 = f(u_1) = f(w_{11}x_1 + w_{21}x_2 + w_{01})$$

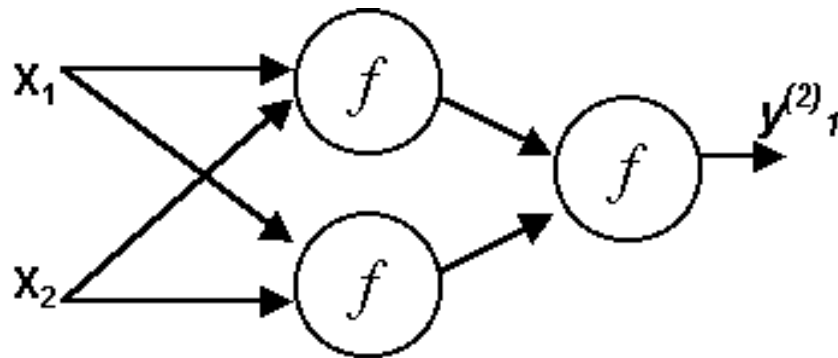
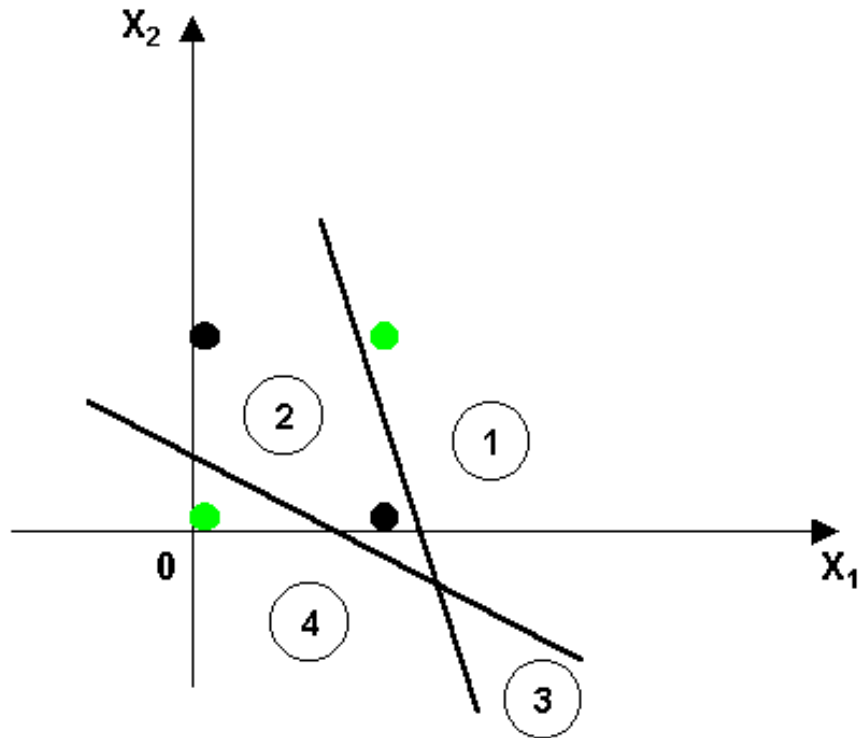
$$y_2 = f(u_2) = f(w_{12}x_1 + w_{22}x_2 + w_{02})$$

Решение проблемы «исключающего ИЛИ»

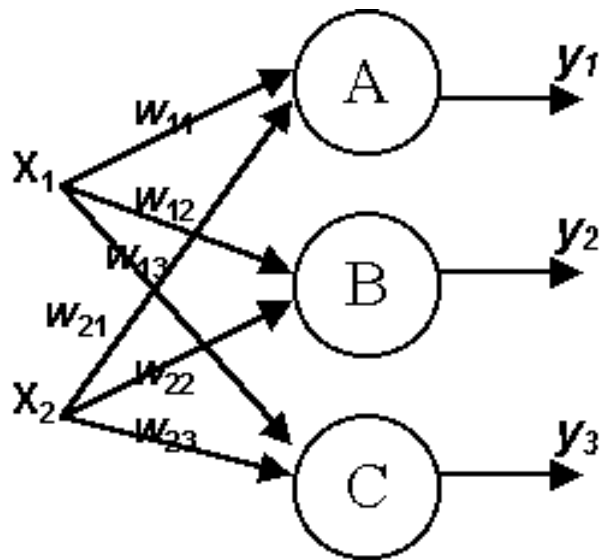


1. Множество точек, для которых $u_1 \geq 0 \quad u_2 \geq 0$
2. Множество точек, для которых $u_1 \geq 0 \quad u_2 < 0$
3. Множество точек, для которых $u_1 < 0 \quad u_2 \geq 0$
4. Множество точек, для которых $u_1 < 0 \quad u_2 < 0$

Решение проблемы «исключающего ИЛИ»



Влияние архитектуры НС на ее функциональность



$$u_1 = w_{11}x_1 + w_{21}x_2 + w_{01}$$

$$u_2 = w_{12}x_1 + w_{22}x_2 + w_{02}$$

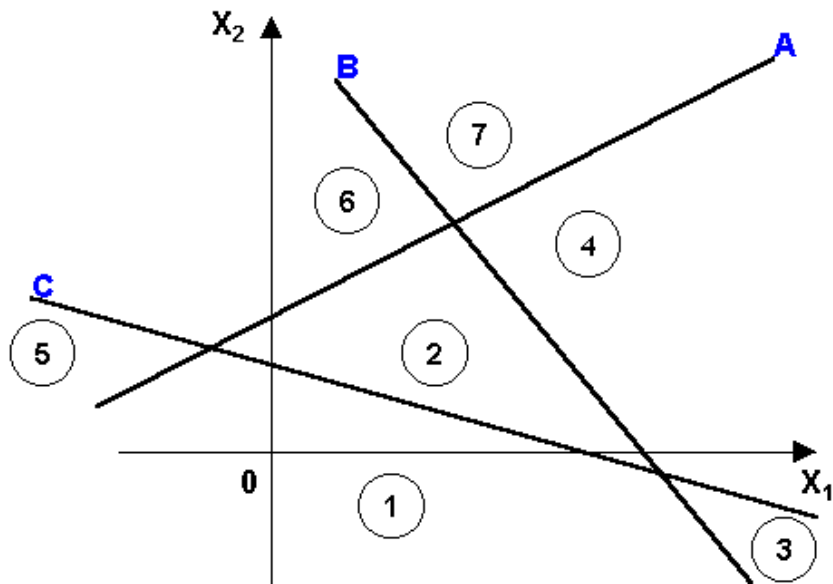
$$u_3 = w_{13}x_1 + w_{23}x_2 + w_{03}$$

$$y_1 = f(u_1) = f(w_{11}x_1 + w_{21}x_2 + w_{01})$$

$$y_2 = f(u_2) = f(w_{12}x_1 + w_{22}x_2 + w_{02})$$

$$y_3 = f(u_3) = f(w_{13}x_1 + w_{23}x_2 + w_{03})$$

Влияние архитектуры НС на ее функциональность



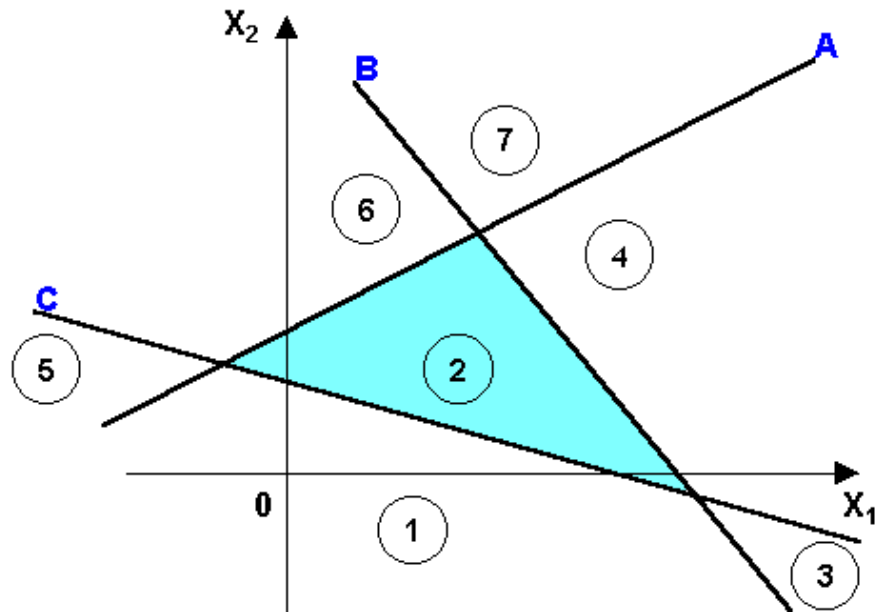
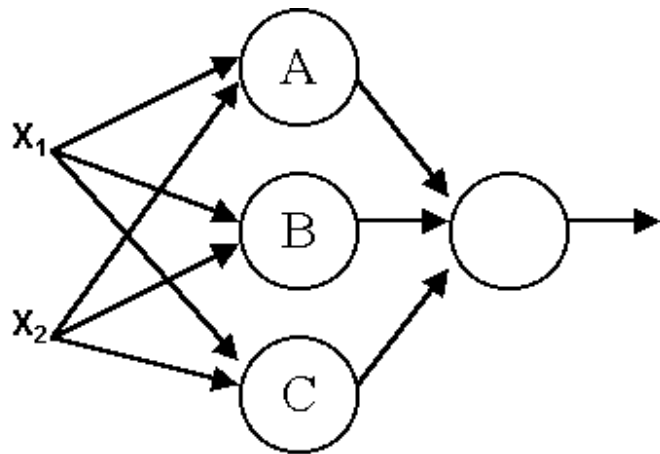
	A	B	C
1	0	0	0
2	0	0	1
3	0	1	0
4	0	1	1
5	1	0	0
6	1	0	1
7	1	1	1

$$A: w_{11}x_1 + w_{21}x_2 + w_{01} = 0$$

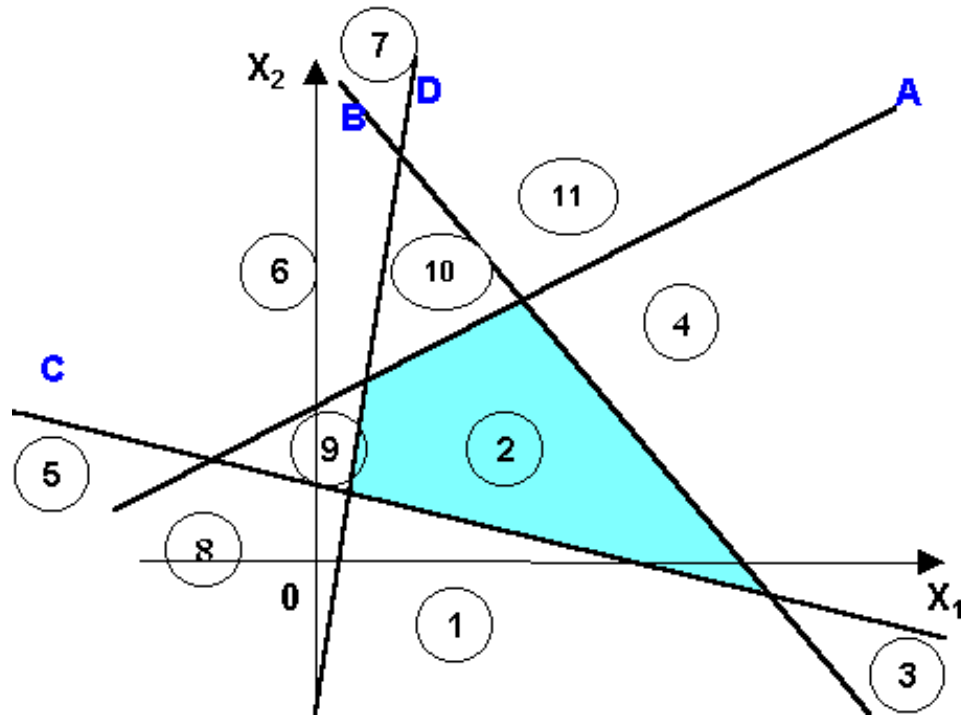
$$B: w_{12}x_1 + w_{22}x_2 + w_{02} = 0$$

$$C: w_{13}x_1 + w_{23}x_2 + w_{03} = 0$$

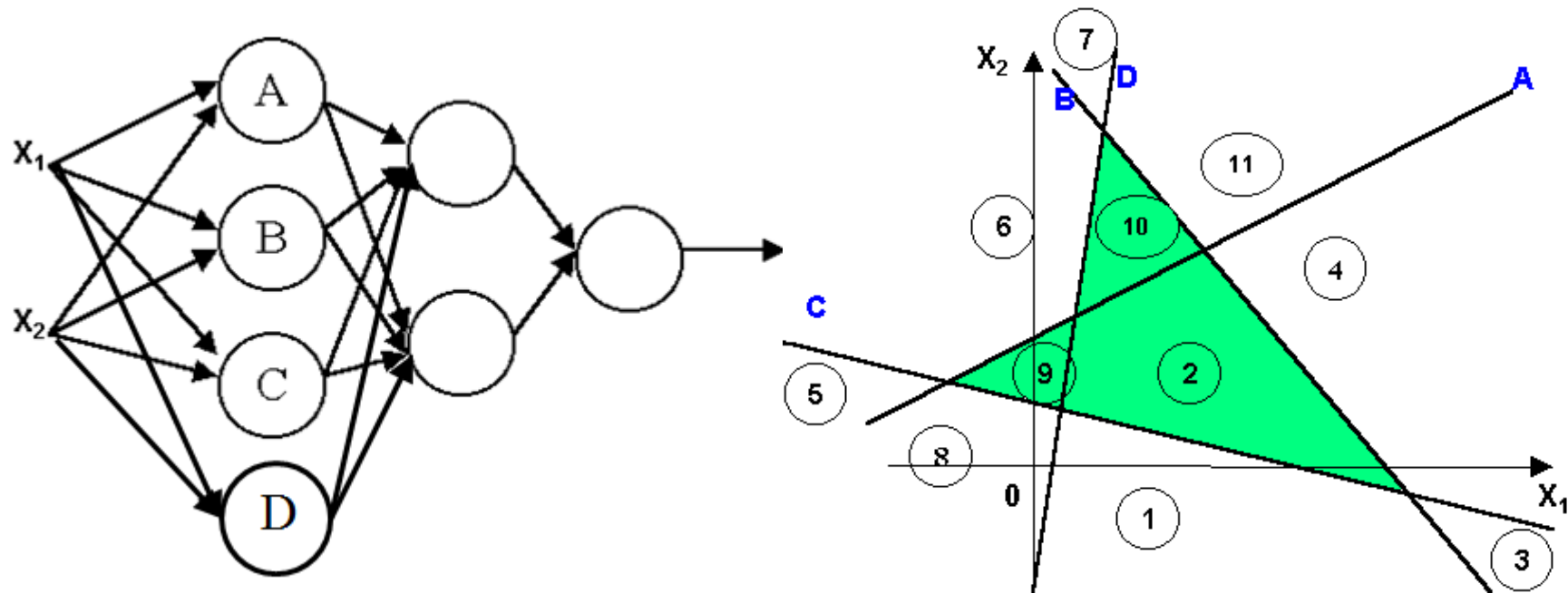
Влияние архитектуры НС на ее функциональность



Влияние архитектуры НС на ее функциональность



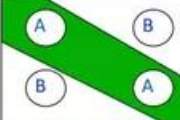
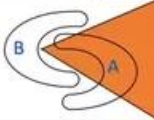
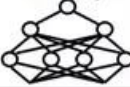
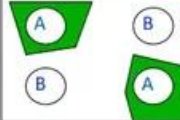
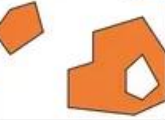
Влияние архитектуры НС на ее функциональность



Влияние архитектуры НС на ее функциональность

Выводы:

1. Добавление нейронов в слой позволяет изменять форму выделяемой области (при этом тип области не меняется), увеличивая точность разделения образов.
2. Добавление слоев позволяет изменять тип выделяемой области, расширяя классы классифицируемых образов.

Structure	Types of Decision Regions	Exclusive-OR Problem	Classes with Meshed regions	Most General Region Shapes
Single-Layer 	Half Plane Bounded By Hyperplane			
Two-Layer 	Convex Open Or Closed Regions			
Three-Layer 	Arbitrary (Complexity Limited by No. of Nodes)			

Визуализация решения задачи классификации:

<https://programforyou.ru/projects/feed-forward-network-visualizer>

Теорема об универсальной аппроксимации (1)

Многослойный персептрон - аппроксиматор функции $F: R^{m_0} \rightarrow R$

Каково минимальное количество скрытых слоев для аппроксимации любой функции?

Пусть $\varphi(\cdot)$ — ограниченная, не постоянная монотонно возрастающая непрерывная функция. Пусть I_{m_0} — m_0 -мерный единичный гиперкуб $[0, 1]^{m_0}$. Пусть пространство непрерывных на I_{m_0} функций обозначается символом $C(I_{m_0})$. Тогда для любой функции $f \in C(I_{m_0})$ и $\varepsilon > 0$ существует такое целое число m_1 и множество действительных констант α_i, b_i и w_{ij} , где $i = 1, \dots, m_1, j = 1, \dots, m_0$, что

$$F(x_1, \dots, x_{m_0}) = \sum_{i=1}^{m_1} \alpha_i \varphi \left(\sum_{j=1}^{m_0} w_{ij} x_j + b_i \right) \quad (4.86)$$

является реализацией аппроксимации функции $f(\cdot)$, т.е.

$$|F(x_1, \dots, x_{m_0}) - f(x_1, \dots, x_{m_0})| < \varepsilon$$

для всех $x_1, x_2, \dots, x_{m_0}$, принадлежащих входному пространству.

Теорема об универсальной аппроксимации (2)

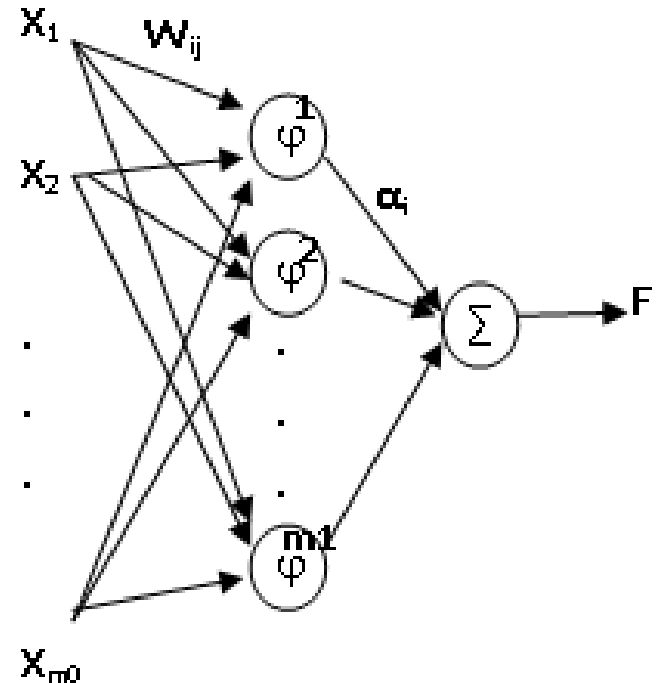
- Архитектура универсального аппроксиматора функции $F: R^{m_0} \rightarrow R$

Входной слой: m_0

Внутренний слой: m_1

Выходной слой: 1 линейный нейрон

- Теорема существования
- Один внутренний слой – не всегда оптимальное решение



Вопросы

1. Почему при помощи однослойной НС нельзя смоделировать функцию «Исключающее или»?
2. Какое минимальное число слоев должно быть в НС, чтобы иметь возможность аппроксимировать с помощью нее любую непрерывную функцию с произвольной точностью?