Выполнила: Исмакова Даяна

49. ВИДЫ ФУНКЦИИ АКТИВАЦИИ НЕЙРОНОВ

**Функция активации** определяет выходное значение нейрона в зависимости от результата взвешенной суммы входов и порогового значения.

Обозначив функцию активации как ***f*** получаем выходной сигнал Y:

где *xi* и *wi* – входное значение и вес *i*-го входа, а b – смещение.

На рисунке 1 представлена модель нейрона.

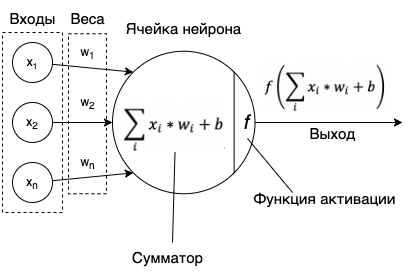


Рисунок 1. Модель нейрона

Основные функции активации:

* ступенчатая,
* линейная,
* сигмоидная,
* гиперболического тангенса,
* ReLu (Rectified linear unit).

**Ступенчатая функция**

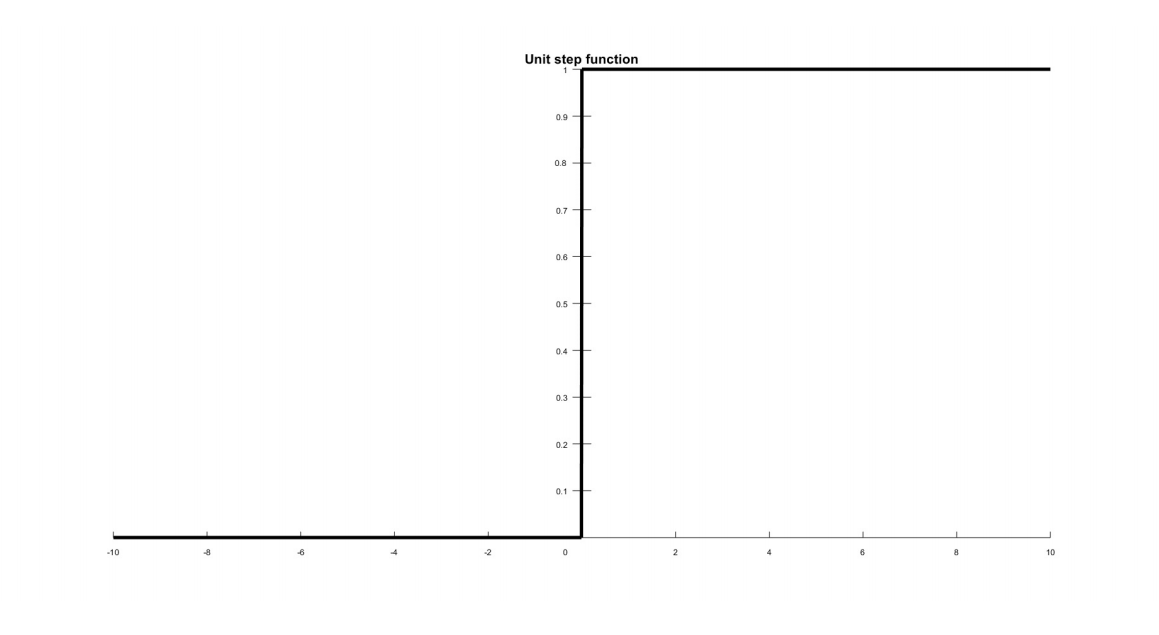
Если взвешенная сумма входов больше или меньше некоторого значения, то нейрон становится активированным:

,

где z – взвешенная сумма входов.

Используется для бинарной классификации, не подходит для классификаций более 2-х классов. Диапазон: [0,1]. На рисунке 2 представлена графическая интерпретация ступенчатой функции.

Рисунок 2. Ступенчатая функция



**Линейная функция**

Линейная функция представляет собой прямую:

Возможна классификация с большим количеством классов. Диапазон:

(-∞, +∞).Основные проблемы использования:

* невозможность использования метода обратного распространения ошибки,
* невозможность создать многослойную сеть, так как любое количество слоев может быть заменено одним слоем.

На рисунке 3 представлена графическая интерпретация линейной функции.

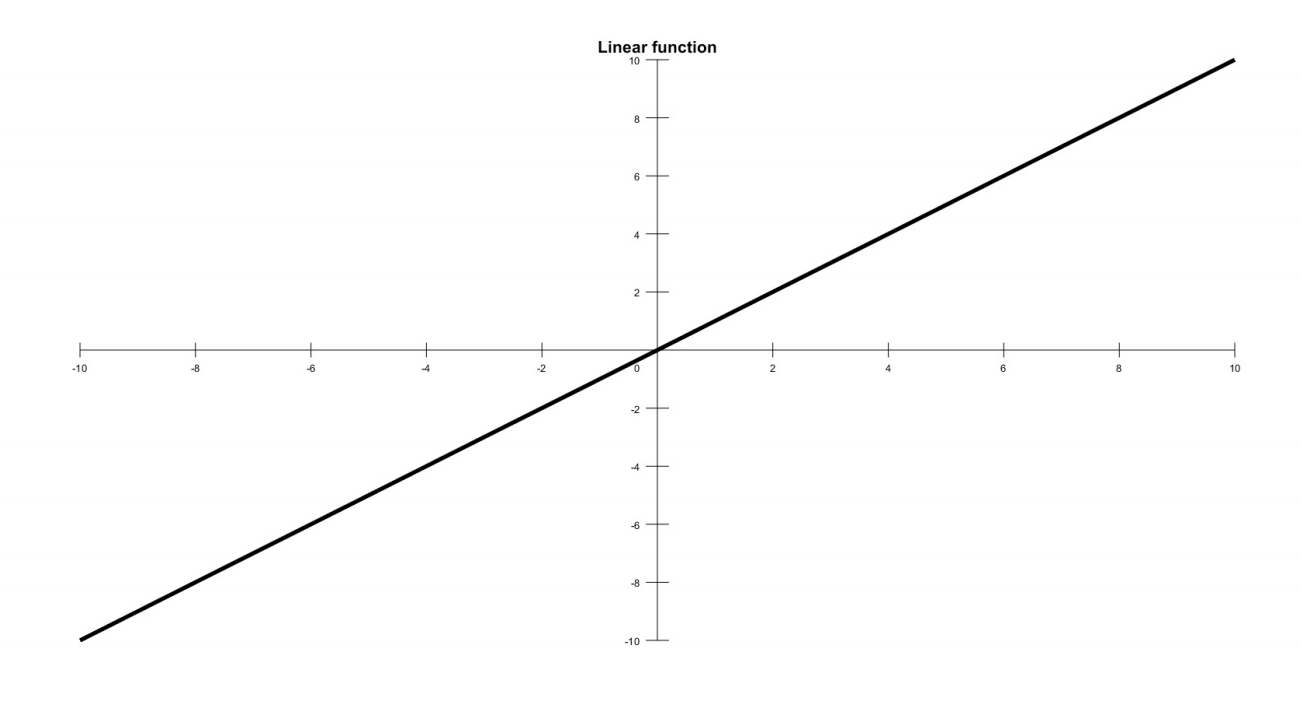


Рисунок 3. Линейная функция

**Сигмоидная функция**

Сигмоидная или логистическая функция является гладкой монотонно возрастающей нелинейной функцией:

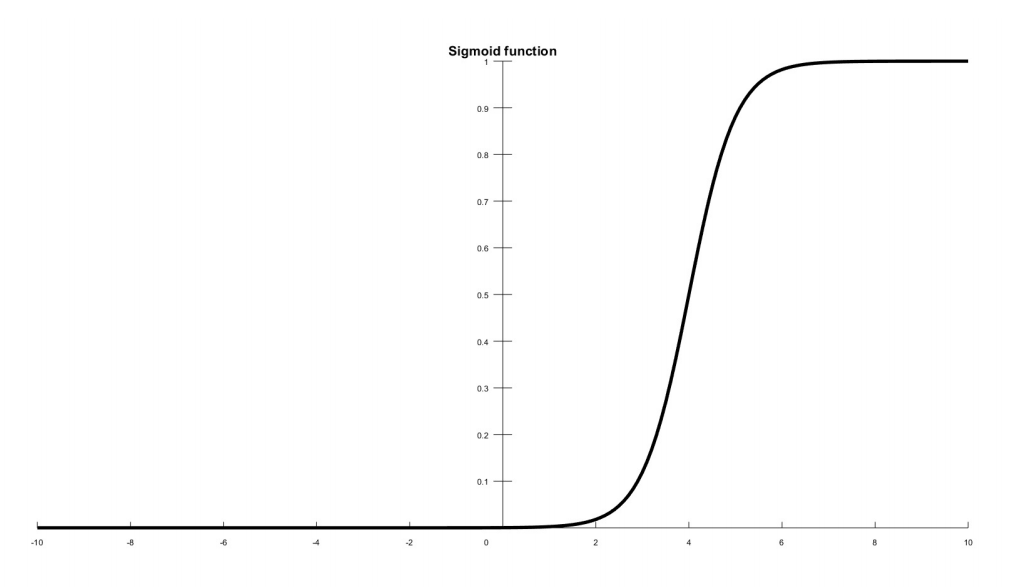


Рисунок 4. Сигмоидная функция

Возможна для классификации с большим количеством классов, а также она подходит для нейронных сетей с множеством слоев. Диапазон: (0,1).

Недостаток: так как значения Y на концах сигмоиды слабо реагируют на изменения в X, то градиент в таких местах принимает очень малые значения, что приводит к проблеме исчезающего градиента.

На рисунке 4 представлена графическая интерпретация сигмоидной функции.

**Функция гиперболического тангенса**

Гиперболический тангенс очень похож на сигмоиду, поэтому имеет схожие характеристики. Формула:

Градиент больше, чем у сигмоидной. Диапазон: (-1, 1).

На рисунке 5 представлена графическая интерпретация функции гиперболического тангенса.

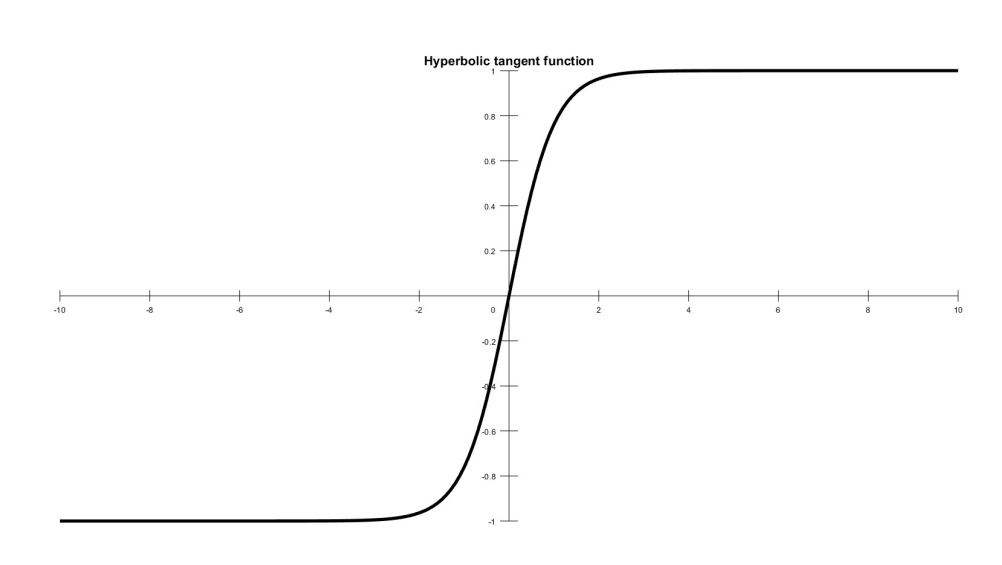


Рисунок 5. Функция гиперболического тангенса

**Функция ReLU**

Функция ReLU возвращает значение z, если *z* положительно, и 0 в противном случае:

Не имеет недостатков, которые есть в предыдущих функциях. Быстрый подсчет производных, разрежённость активаций. Диапазон: [0, inf).

Недостатки: проблема умирающего ReLU, так как часть производной функции равна нулю, то и градиент для нее будет нулевым, а то это значит, что веса не будут изменяться во время спуска и нейронная сеть перестанет обучаться.

На рисунке 6 представлена графическая интерпретация функции ReLU.

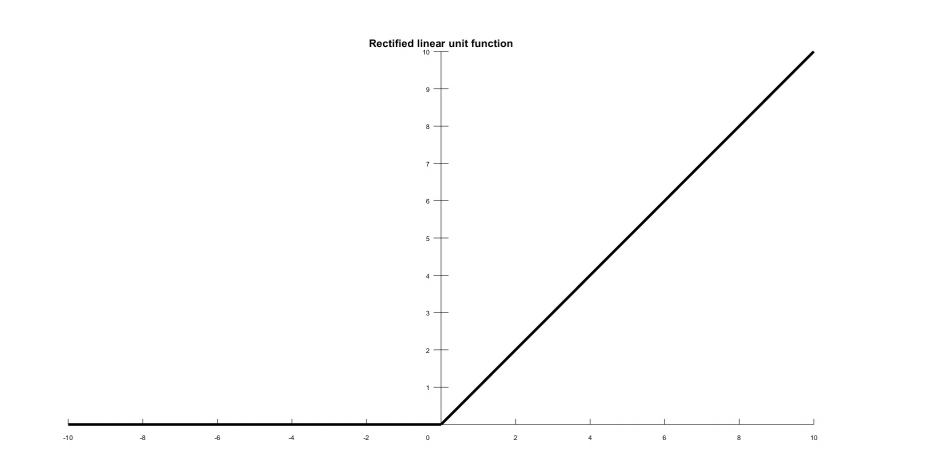


Рисунок 6. Функция ReLU

ИСТОЧНИКИ

1. Wikipedia, Функция активации. URL: <https://ru.wikipedia.org/wiki/Функция_активации> (Дата обращения: 4.12.2020).
2. Милютин И., Функции активации нейросети: сигмоида, линейная, ступенчатая, ReLu, tahn. URL:<https://neurohive.io/ru/osnovy-data-science/activation-functions/> (Дата обращения: 4.12.2020).
3. Университет ИТМО, Практики реализации нейронных сетей. URL: <https://neerc.ifmo.ru/wiki/index.php?title=Практики_реализации_нейронных_сетей#:~:text=Функция%20активации%20(англ.,ого%20входа%2C%20а%20—%20смещение> (Дата обращения: 4.12.2020).