

Лекция по дисциплине «Нейронные сети»

Лекция 1.5 Функции активации – 1 час

Цель: - Рассмотреть основные понятия функции активации;

План: - Степенные функции активации.

- Функция активации квадратный корень

Степенные функции активации представляются выражением:

$$f(s) = s^n,$$

где постоянная величина $n = 2, 3, 4, \dots$ - показатель степени.

У степенных функций активации областью определения является $D(f) = (-\infty, +\infty)$, а область значений зависит от четности показателя степени: если значение показателя четное, то область значений есть $E(f) = (0, +\infty)$, иначе будет $E(f) = (-\infty, +\infty)$.

Требованиям монотонности, непрерывности и дифференцируемости, предъявляемым к функциям активации — в полной мере отвечают лишь степенные функции с нечетными показателями степени.

П.2.7. Примеры:

1. В степенной функции, если $n=2$, то речь идет о квадратичной функции активации. График квадратичной функции активации показан на рисунке П.2.8.1.

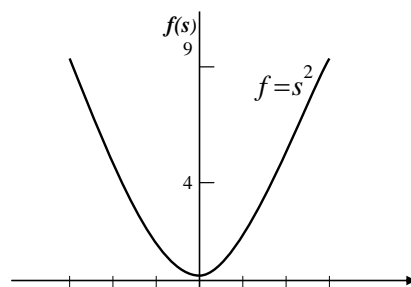


Рисунок П.2.8.1. График квадратичной функции активации.

2. В степенной функции, при $n=3$, имеем дело с кубической квадратичной функцией активации.

3. График кубической функции активации представлен на рисунке П.2.8.2.

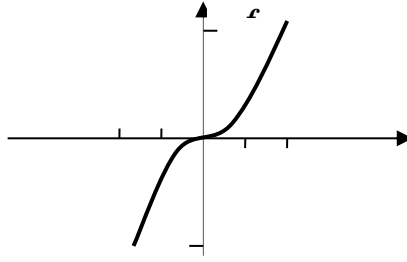


Рисунок П.2.8.2. График кубической функции активации.

Функция активации квадратный корень

В этом параграфе будет рассмотрена функция активации квадратный корень, предложены задания и сформулированы вопросы. При формировании учебного материала данного параграфа были использованы следующие источники [1-9].

Функция активации квадратный корень представляется следующим выражением:

$$f(s) = \sqrt{s}$$

Областью определения функции активации квадратный корень является $D(f)=(0,+\infty)$, а областью значений – $E(f) = (0, +\infty)$.

Функция активации квадратный корень используется в нейронной сети Кохонена. Квадратный корень преобразует активации сети Кохонена, т.е. квадраты расстояний, в выходные значения, представляющие сами расстояния. Сеть Кохонена рассчитана на обучение без учителя и применяется для разведывательного анализа данных (распознавания кластеров в данных, установления близости классов, решения задач классификации, обнаружения новых явлений).

Вопросы:

1. В каких нейронных сетях используется функция активации квадратный корень?
2. Что делает функция активации квадратный корень с нейронной сетью Кохонена?

3. Почему квадраты расстояний преобразуются в выходные значения нейронных сетей, представляющие сами расстояния?

Лекция 1.5 Функции активации – 1 час

Цель: - Рассмотреть основные понятия функции активации;

План: - Экспоненциальные функции активации.

- Функции активации, уменьшающие входные значения

Экспоненциальная функция активации имеет вид:

$$f(s) = e^{-s},$$

где $e = 2,71$ – постоянная величина.

Областью определения экспоненциальной функции активации является $D(f) = (-\infty, +\infty)$, а область значений – $E(f) = (0, +\infty)$.

График экспоненциальной функции активации представлен на рисунке П.2.10.

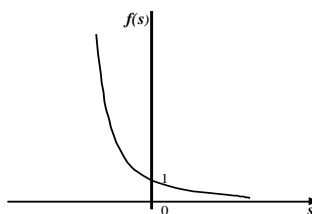


Рисунок П.2.10. График экспоненциальной функции активации.

Экспоненциальная функция активации используется в первом скрытом слое нейронов вероятностных нейронных сетей. Достоинством таких сетей является то, что их архитектура позволяет не только осуществить классификацию объектов, но и определить вероятность справедливости принимаемых решений.

Иногда используют экспоненциальные функции, выходы которых нормируются так, чтобы сумма всех активаций слоя равнялась 1. Их применяют в выходных слоях многослойных персептронов, специально сконструированных для задач классификации таким образом, чтобы выходы можно было интерпретировать как вероятности принадлежности к классу.

Она Экспоненциальная функция активации идеально подходит для радиальных элементов. Комбинация радиальной постсинаптической потенциальной функции (PSP-функции) и отрицательной экспоненциальной функции активации дает элемент, моделирующий гауссову функцию (колоколообразной формы) с центром в векторе весов. Стандартное

отклонение Гауссиана определяется по формуле $\sigma = \sqrt{1/d}$, где через d обозначено "отклонение" элемента, хранимое как его пороговое значение.

Напомним, что PSP-функция применяется к входным сигналам, его весам и порогам и выдает уровень активации этого элемента. Наиболее часто используются линейные (взвешенная сумма входов минус порог) и радиальные (промасштабированный квадрат расстояния от вектора весов до входного вектора) PSP-функции.

II.2.1. Функции активации, уменьшающие входные значения

В этом параграфе будут рассмотрены функции активации, уменьшающие входные значения, предложены задания и сформулированы вопросы. При подготовке учебного материала данного параграфа были использованы следующие источники [1-9].

Функция активации, уменьшающая входные значения (softmax), является экспоненциальной функцией и нормирует значение выхода так, чтобы сумма всех активаций нейронов равнялась единице.

В многослойной нейронной сети каждый i -тый нейрон softmax-группы будет иметь функцию активации, которая представляется следующим выражением:

$$y_i = \frac{e^{s_i}}{\sum_j e^{s_j}}, \quad (\text{II.2.11.1})$$

где $i \in N$ – индекс нейрона, $j \in N$ – номер слоя.

Областью определения функции активации softmax является $D(f) = (-\infty, +\infty)$, а областью значений – $E(f) = (0, +1)$.

Из выражения видно, что выход каждого нейрона зависит от сумматоров всех остальных нейронов softmax-группы, а сумма выходных значений всей группы равняется единице.

На каждом обучающем примере мы будем получать выход сети, который моделирует нужное нам вероятностное распределение, а для сравнения двух вероятностных распределений необходима корректная мера. В качестве такой меры будет использоваться перекрестная энтропия:

$$C = -\sum_j d_j \log y_j, \quad (\text{II.2.11.2})$$

где d_j – требуемые выходы для обучающего примера j , y_j – реальные выходы для обучающего примера j .

Функции активации softmax используются для решения задач классификации, когда выход нейрона нужно интерпретировать как вероятность принадлежности к определенному классу, т.е. когда необходимо нейронной сетью обеспечить способ моделирования вероятностного

распределения. Для этого используется многослойная нейронная сеть прямого распространения, такая что:

- сеть содержит некоторое количество скрытых слоев, все нейроны могут иметь свою собственную функцию активации;
- на последнем слое находится такое количество нейронов, которое соответствует количеству классов; все эти нейроны будут называться softmax-группой.

Вопросы:

1. Является ли функция активации softmax экспоненциальной функцией?
2. Что зависит от сумматоров всех остальных нейронов softmax-группы?
3. Чему равна сумма выходных значений всей группы?
4. Для какой задачи используется функция активации softmax?
5. Для чего используется многослойная нейронная сеть прямого распространения?