Лекция по дисциплине «Нейронные сети»

Лекция 1.9 Модели нейронов

Цель: - Рассмотреть основные понятия моделей нейронов;

План: - Кубические нейроны.

- Радиальные нейроны

Кубические нейроны относятся к нестандартной модели искусственного нейрона. В кубической модели нейронов вектор входных двоичных сигналов *X* рассматривается как адрес ячейки памяти, содержимое которой равно 0 или 1. Для размерности *n* вектора *X* существует 2^n возможных адресов.

Можно рассматривать ячейки памяти, как вершины $\boldsymbol{n}$-мерного *гиперкуба*. Ячейки памяти получают значения независимо друг от друга. Полезно рассматривать ячейки памяти как содержащие поляризованные двоичные значения ± 1. Тогда работа кубического модуля описывается следующим образом.

Двоичный вход *x* используется как адрес памяти, поляризованная двоичная величина считывается и конвертируется в неполяризованную форму функцией *h*. Обозначим значения по адресу *x* через v_x, так что $y = h(v_x)$. Такие модули мы будем называть кубическими, чтобы подчеркнуть геометрическое представление множества адресов значений активации как множество вершин гиперкуба.

Рассмотрим двухвходовый кубический модуль. Существует 4 значения активации $\{v_{00}, v_{01}, v_{10}, v_{11}\}$. Выражение для активации будет иметь следующий вид:

$$s = v_{00}(1 - x_1)(1 - x_2) + v_{01}(1 - x_1)x_2 + v_{10}x_1(1 - x_2) + v_{11}(1 + x_1)(1 + x_2)$$

$x = (x_1, x_2)$ - входной вектор. Такая запись вызвана тем, что только одно из произведений в сумме должно быть ненулевым. Для поляризованных входов x_1 и x_2 активация будет

$$s = [v_{00}(1 - x_1)(1 - x_2) + v_{01}(1 - x_1)x_2 + v_{10}x_1(1 - x_2) + v_{11}(1 + x_1)(1 + x_2)] / 4$$

В случае модуля с n входом получим

$$s=[\sum_{i=0}^{m} v_{ij} \prod_{j=1}^{n} (1-x_j)]/2^n$$

Кубические нейроны обучаются путем изменения содержимого ячейки их памяти. Обозначим через ‘+’ операцию инкремента –установки содержимого ячейки в +1, а через ‘-’ операцию декремента – установки в -1.

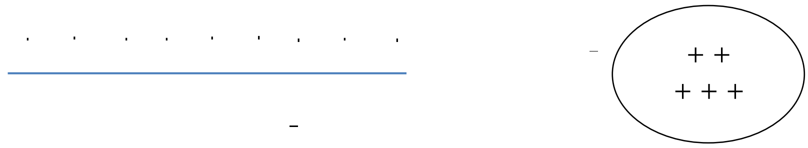
Пусть в начальном состоянии все ячейки кубического нейрона установлены в ноль. Обозначим ячейки, адресуемые обучающей выборкой, как центральные ячейки или центры. Ячейки, близкие к центрам в смысле расстояния Хемминга, будем настраивать на те же или близкие к ним значения, что и сами центры, т.е. должна происходить кластеризация значений ячейки вокруг центра. Это условие должно выполняться для сети из кубических нейронов.

Алгоритм обучения строит так называемое разбиение Вороного, при котором значение в ячейке определяется значением в ближайшем центре, а ячейки, равноудаленные от центров, остаются установленными в ноль. Кубические нейроны допускают большую функциональность, чем полулинейные, и поэтому, возможно, позволяют решать те же задачи при меньшем количестве модулей.

Радиальные нейроны

В этом параграфе рассмотрены радиальные нейроны, обсуждены возможности данных нейронов~~а~~, предложены задания и сформулированы вопросы. При подготовке учебного материала данного параграфа были использованы источники [1-9].

Радиальные нейроны относятся к нестандартной модели искусственного нейрона и представляет собой гиперсферу, которая осуществляет шаровое разделение пространства вокруг центральной точки. Именно с этой точки зрения он является естественным дополнением сигмоидального нейрона, поскольку в случае круговой симметрии данных позволяет заметно уменьшить количество нейронов, необходимых для разделения различных классов. На рисунке II.3.12.1 показаны способы разделения пространства данных сигмоидальным и радиальным нейронами.



Способы разделения пространства данных.
А) синмоидальным нейроном; В) радиальным нейроном.

Радиальные нейроны реализуют функцию, радиально изменяющуюся вокруг выбранного центра и принимающую ненулевые значения только в окрестности этого центра. Подобные функции, определяемые в виде $f(X,C)=\|X-C\|$, называются *радиальными базисными функциями*, где $X \in R^n$, $\boldsymbol{C}$ – центр.

Примерами радиальных базисных функций могут служить:

1. Мультиквадратичная функция: $f(X,C)=\left(\|X-C\|^2+a^2\right)^{\frac{1}{2}}$, где $a>0$.

2. Сплайн тонких пластин: $f(X,C)=\left(\|X-C\|^2\ln\left(\|X-C\|\right)\right)^{\frac{1}{2}}$.

3. Обратная мультиквадратичная функция: $f(X,C)=\dfrac{1}{\left(\|X-C\|^2+a^2\right)^{\frac{1}{2}}}, a>0$

4. Функция Гаусса: $f(X,C)=\exp\left(-\dfrac{\|X-C\|^2}{2\sigma^2}\right)$, где $\sigma>0$ – параметр, определяющий "ширину" функции.

Структура радиального нейрона показана на рисунке II.3.12.2.

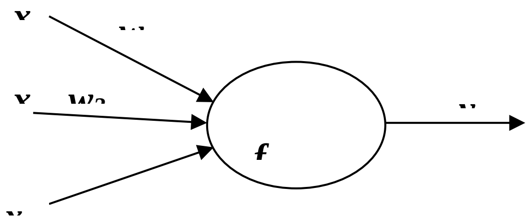


Рисунок II.3.12.2. Структура радиального нейрона.

На рисунке II.3.12.3 приведен график одномерной радиальной функции в скалярном варианте для различных значений $\boldsymbol{s}_i$.

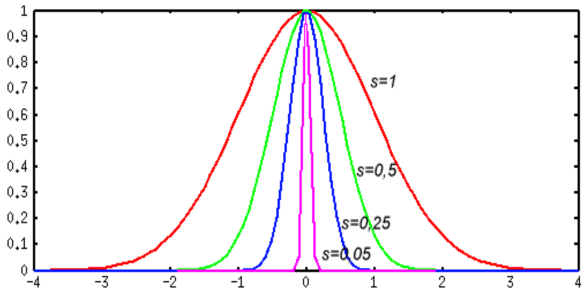


Рисунок II.3.12.3. График одномерной радиальной функции.

На рисунке II.3.12.4 приведен график двумерной радиальной функции в скалярном варианте для различных значений $\boldsymbol{s}_i$.

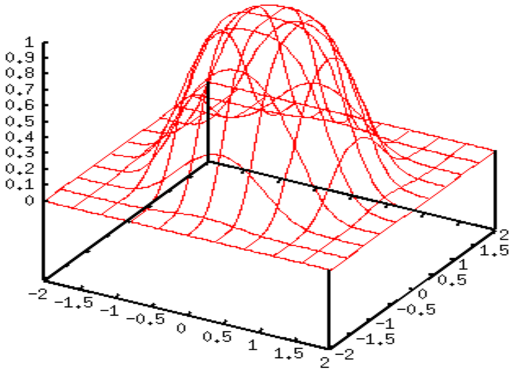


Рисунок II.3.12.4. График двумерной радиальной функции.

Принципиальное отличие радиального нейрона от сигмоидального нейрона и персептрона в том, что сигмоидальный нейрон разбивает многомерное пространство входных сигналов гиперплоскостью, а радиальный – гиперсферой.

Вопросы.

1. Какую функцию реализуют радиальные нейроны?
2. Как изменяются радиальные функции?
3. Какие значения принимают радиальные функции в окрестности центра?