

УДК 519.21

О.В. Серая, Д.И. Брезинский

Национальный технический университет «Харьковский политехнический институт»

НЕЙРОСЕТЕВАЯ ОБРАБОТКА НЕЧЕТКИХ ДАННЫХ С АДАПТИВНОЙ АКТИВАЦИОННОЙ ФУНКЦИЕЙ

В статье рассматривается методика построения искусственной нейронной сети с адаптивной активационной функцией для случая, когда на вход поступает нечеткий набор данных.

нейронная сеть, адаптивная активационная функция, нечеткий набор данных

Введение

Постановка проблемы и анализ литературы. Хорошо известно, что искусственные нейронные сети (ИНС) являются надежным и эффективным инструментом решения задач установления зависимости в системе «факторы – отклик» [1 – 3]. Вместе с тем, проблемы остаются, если имеющаяся выборка наблюдений недостаточна для построения искомой зависимости с высоким уровнем адекватности. Как показано в [4, 5], дополнительные возможности улучшения качества модели возникают, если использовать адаптивную активационную функцию. Дело в следующем. При построении реальных ИНС выбирают некоторый определенный тип активационной функции из стандартного набора:

– пороговая униполярная

$$\varphi_1(s) = \begin{cases} 0, & s < a; \\ 1, & s \geq a; \end{cases}$$

– пороговая биполярная

$$\varphi_2(s) = \begin{cases} -1, & s < a; \\ 1, & s \geq a; \end{cases}$$

– логистическая

$$\varphi_3(s) = \frac{1}{1 + e^{-as}};$$

– шаговая (линейная с насыщением)

$$\varphi_4(s) = \begin{cases} -1, & s < -1; \\ s, & -1 \leq s \leq 1; \\ 1, & s > 1; \end{cases} \quad (1)$$

– синусоидальная с насыщением

$$\varphi_5(s) = \begin{cases} -1, & s < -a; \\ \sin s, & -a \leq s \leq a; \\ 1, & s > a; \end{cases}$$

– экспоненциальная

$$\varphi_6(s) = 1 - e^{-as}.$$

В ситуации, когда наилучший вариант выбора активационной функции при решении конкретной задачи отыскания зависимости функции отклика от факторов априори не известен, приходится осуществлять их перебор. При этом удачный выбор типа активационной функции существенно увеличивает адекватность модели, что особенно существенно в условиях малой выборки. Понятно, что в этой ситуации радикальный путь упрощения решения задачи поиска адекватной модели состоит в использовании некоторой универсальной функции, вариацией параметров которой можно было бы эффективно аппроксимировать любую активационную функцию из набора (1). В качестве такой функции в [4, 5] предложена следующая:

$$\varphi_0(x) = b_0 + b_1 \text{th}(c_0 + c_1 x + c_2 x^2) = b_0 + b_1 \text{th}(x). \quad (2)$$

Там же показано, что надлежащим выбором параметров функции (2) можно эффективно с высокой точностью аппроксимировать любую активационную функцию из набора (1). При этом, однако, открытым остался вопрос о том, насколько важно при описании $y(x)$ в (2) использовать именно квадратичный полином, а не линейный вида

$$y(x) = a_0 + a_1 x. \quad (3)$$

Далее, исследование эффективности использования (2) при адаптации активационной функции проведено в предположении, что каждому конкретному набору факторов однозначно соответствует конкретный отклик. Однако в реальности дело может обстоять не так. В практических задачах установления связи «факторы – отклик» достаточно часто возникает ситуация, когда одному и тому же на-

бору факторов соответствует не одно, а некоторое множество значений функции отклика. Если при этом выборка наблюдений мала и получить обоснованное теоретико-вероятностное описание результатов экспериментов в виде плотности вероятности значений функции отклика не представляется возможным, то наиболее естественным является нечеткое представление отклика. Проблемы использования адапционных активационных функций в этой ситуации ранее не обсуждались.

Целью статьи является следующее: во-первых, оценить необходимость использования квадратичного описания $y(x)$ в (2) и, во-вторых, предложить обоснованную методику построения ИНС для обработки нечетких данных с адаптивной активационной функцией.

Постановка задачи. Пусть для каждого набора факторов $F_j = (F_{1j}, F_{2j}, \dots, F_{mj})$, $j = 1, 2, \dots, n$ известно соответствующее этому набору множество значений функции отклика $Y_j = \{y_{1j}, y_{2j}, \dots, y_{Rj}\}$. С использованием Y_j зададим описание функции принадлежности $\mu(y_j)$ нечеткого значения y_j функции отклика для входного набора факторов F_j , например, в виде

$$\mu(y_j) = \exp\left\{-\frac{(y_j - \bar{y}_j)^2}{2\sigma_j^2}\right\};$$

$$\bar{y}_j = \frac{1}{R} \sum_{r=1}^R y_{rj}; \quad \sigma_j^2 = \frac{1}{R} \sum_{r=1}^R (y_{rj} - \bar{y}_j)^2, \quad j = 1, 2, \dots, n.$$

Поставим задачу построения ИНС, максимально адекватно отображающей вход (набор факторов) в нечеткий выход с фиксацией соответствующего значения функции принадлежности.

Основные результаты

При решении первой задачи было осуществлено оценивание параметров адапционных активационных функций (2) и (3) при аппроксимации базового набора активационных функций (1). Параметры аппроксимирующих функций отыскивались методом Нелдера-Мида. Результаты приведены на рис 1 – 6, где обозначены:

- базовая функция;
- аппроксимация (2);
- — — — — аппроксимация (3).

Из анализа соответствующих кривых следует:

- 1) во всех случаях квадратический полином обеспечивает лучшую аппроксимацию, нежели линейный;
- 2) в ряде случаев линейная аппроксимация не слишком сильно уступает квадратической, что делает целесообразным её использование при малой выборке наблюдений.

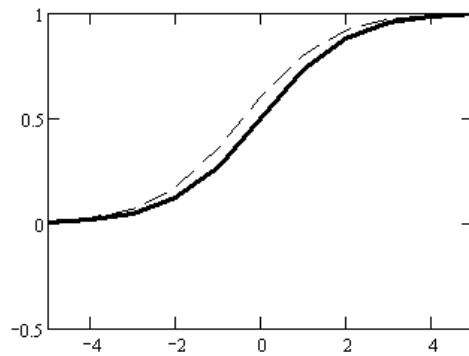


Рис. 1. Логистическая функция

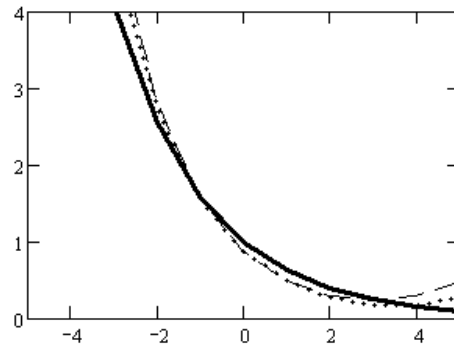


Рис. 2. Экспоненциальная функция

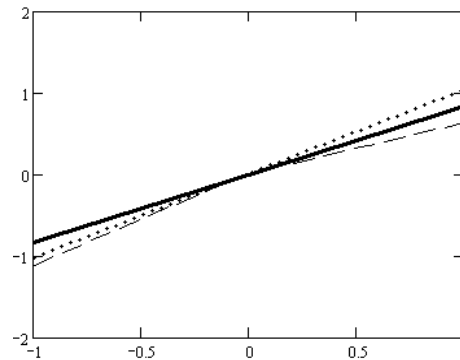


Рис. 3. Синусоидальная функция

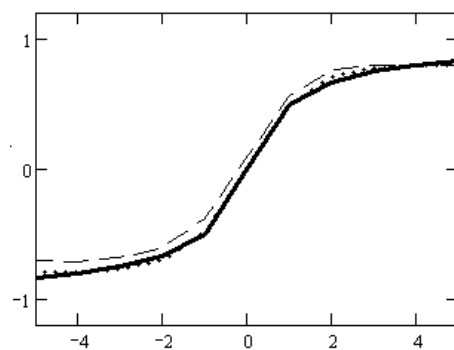


Рис. 4. Сигмоидальная функция

При решении второй задачи принципиальным является вопрос о выборе скалярного критерия качества обучения нейронной сети. Предлагается следующая схема формирования критерия. Для фиксированного набора параметров адаптации $(b_0, b_1, c_0, c_1, c_2)$ активационной функции и заданного набора весов $\{W_k\}$ каждому входному набору факторов F_j ставится

в соответствие значение функции отклика \hat{y}_j и отвечающее ему значение функции принадлежности $\mu(\hat{y}_j)$. Тогда для множества входных наборов F_j , $j=1,2,\dots,n$, может быть вычислено значение

$$\eta_1(b_0, b_1, c_0, c_1, c_2, \{W_k\}) = \sum_{j=1}^n \mu(\hat{y}_j), \quad (4)$$

которое является количественной мерой качества набора управляемых (настраиваемых) параметров ИНС. При этом с использованием любой процедуры обучения, например, методом обратного распространения ошибок, можно получить наиболее эффективную настройку сети. Понятно, что качество обучения существенно зависит от объема выборки, т.е. от числа входных наборов факторов F_j , $j = 1, 2, \dots, n$.

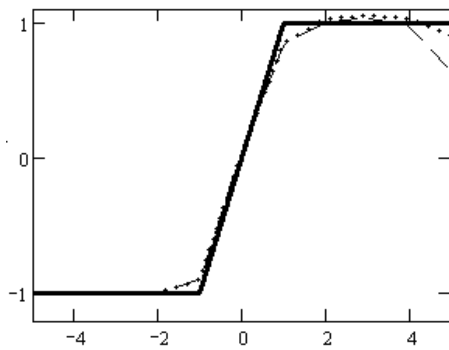


Рис. 5. Шаговая функция

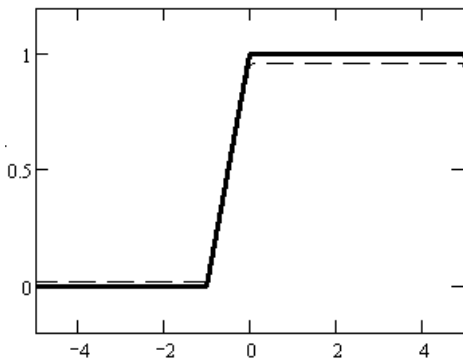


Рис. 6. Пороговая униполярная функция

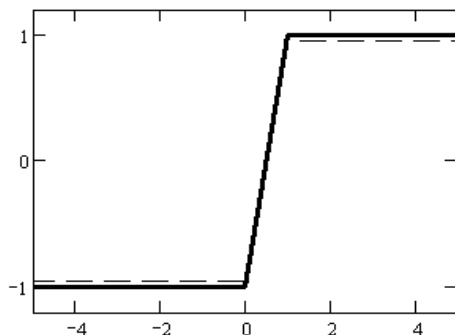


Рис. 7. Пороговая биполярная функция

В этой задаче во многих случаях может оказаться более предпочтительным другой, минимаксный критерий, формируемый в соответствии с соотношением

$$\eta_2(b_0, b_1, c_0, c_1, c_2) = \min_j \{\mu(\hat{y}_j)\}. \quad (5)$$

При этом настройка параметров сети направлена на максимизацию (5).

Использование этого критерия устраняет конструктивный недостаток критерия (4), состоящий в том, что максимальное значение суммы (4) может быть достигнуто в ситуации, когда качество предсказания отклика для отдельных или группы входных наборов будет неудовлетворительным.

Выводы

Таким образом, предложена методика обработки информации в искусственной нейронной сети с адаптивной активационной функцией для случая, когда результаты наблюдений представлены нечетко.

Введены естественные критерии качества обучения сети.

Направление дальнейших исследований может быть связано с улучшением качества аналитического описания функции принадлежности нечетких входных данных.

Список литературы

1. Руденко О.Г., Бодянский Е.В. Основы теории нейронных сетей. – Ч.: ТЕЛЕТЕХ, 2002. – 317 с.
2. Галушкин А.И. Теория нейронных сетей. – М.: ИПРЖР, 2000. – 392 с.
3. Каллан Р. Основные концепции нейронных сетей: Пер. с англ. – М.: Изд. Дом «Вильямс», 2001. – 288 с.
4. Раскин Л.Г., Серая О.В., Фищукова Н.В. Повышение эффективности нейронных сетей путем использования универсальной активационной функции // Вестник НТУ «ХПИ». – Х.: НТУ «ХПИ». – 2002. – № 13. – С. 39-43.
5. Раскин Л.Г., Серая О.В., Фищукова Н.В. Адаптивная нейронная сеть с перестраиваемой активационной функцией // Вестник НТУ «ХПИ». – Х.: НТУ «ХПИ». – 2003. – № 7. – С. 53-56.

Поступила в редколлегию 28.02.2007

Рецензент: д-р техн. наук, проф. Л.Г.Раскин, Национальный технический университет «ХПИ», Харьков.