УДК 004.852

ПРИНЦИПЫ НЕЧЕТКОЙ ЛОГИКИ НА ПРИМЕРЕ НЕЧЕТКИХ НЕЙРОННЫХ СЕТЕЙ

[1]Мищенко В.А., [2]Коробкин А.А.

[1]Воронежский государственный педагогический университет, Воронеж
Воронеж, Россия (394000, г. Воронеж, ул. Ленина, 86)
[2]Воронежский государственный университет, Воронеж
Воронеж, Россия (394006, г.Воронеж, Университетская пл., 1)

В данной статье рассмотрены принципы построения систем, основанных на нечеткой логике, кроме того, определен принцип построения логического вывода. Также рассматривается структура организации нечетких нейронных сетей на примере сети Ванга – Менделя. Описывается схема организации такой сети, ее структура, в частности, определены слои нейронной сети и описаны принципы функционирования каждого слоя. Кроме того, рассмотрен процесс обучения нечеткой нейронной сети Ванга – Менделя, включающий в себя подстройку весовых коэффициентов сети и настройку параметров функции Гауса. А также рассмотрен процесс обучения сети в случае, когда нахождения решения процесса обучения невозможно, а поиск параметров осуществляется таким образом, что все условия выполняются в некоторой степени. Также в статье проведен сравнительный анализ различных типов архитектур интеллектуальных систем.

Ключевые слова: нечеткая логика, нечеткие нейронные сети.

PRINCIPLES OF INDISTINCT LOGIC BY THE EXAMPLE OF INDISTINCT NEURAL NETWORKS

[1]Mishchenko V.A., [2]Korobkin A.A.

[1]The Voronezh state pedagogical university, Voronezh
Voronezh, Russia (394000, Voronezh, street. Lenin, 86)
[2]The Voronezh state university, Voronezh
Voronezh, Russia (394006, Voronezh, the University square, 1)

In given article principles of construction of the systems based on indistinct logic besides the principle of construction of a logic conclusion is determined are considered. As the structure of the organization of indistinct neural networks is examined by the example of Vang-Mendel network. The circuit of the organization of such network, its structure is described, layers of a neural network in particular are determined and principles of functioning of each layer are described. Besides process of training of an indistinct neural network of Vang-Mendel, including fine tuning of weight factors of a network and adjustment of parameters of function Gaus is considered. And as process of training of a network in a case when presence of the decision of process of training it is not possible is considered, and search of parameters is carried out in such a manner that all conditions are carried out somewhat. As in article the comparative analysis of various types of architecture of intellectual systems is lead.

Key words: indistinct logic, indistinct neural networks.

Используемая в различных видах систем модель на основе нечеткой логики представляет собой базу знаний, построенную специалистами предметной области как множество нечетких правил вида:

$$\Pi_1 \text{ если } x \text{ есть } A_1, \text{ то } y \text{ есть } B_1,$$
$$\Pi_2 \text{ если } x \text{ есть } A_2, \text{ то } y \text{ есть } B_2,$$
$$\ldots$$
$$\Pi_n \text{ если } x \text{ есть } A_n, \text{ то } y \text{ есть } B_n,$$

где x и y- входная и выходная переменная соответственно, а A и B- функции принадлежности [2].

Нечеткий логический вывод формируется в несколько шагов:

- введение нечеткости: на этом этапе функции принадлежности применяются к фактическим значениям входных переменных;

- логический вывод: вычисляется значение истинности для предпосылок каждого правила и применяется к заключениям каждого правила. Это приводит к одному нечеткому подмножеству, которое будет назначено каждой переменной вывода для каждого правила;
- композиция: нечеткие подмножества, назначенные каждой переменной вывода, объединяют в одно множество для всех переменных вывода;
- приведение к четкости: используется в случаях, когда необходимо преобразовать нечеткий набор выводов в четкое число.

На этих принципах построено большое количество сетей, рассмотрим подробнее одну из них – сеть Ванга – Менделя. Структура такой сети представляет собой четырехслойную нейронную сеть, в которой первый слой выполняет фазификацию входных переменных, второй – агрегирование значений активации условия, третий – агрегирование M правил вывода (первый нейрон) и генерацию нормализующего сигнала (второй нейрон), тогда как состоящий из одного нейрона выходной слой осуществляет нормализацию, формируя выходной сигнал [1,3].

В этой сети первый и третий слой являются параметрическими: первый слой содержит *M*N*2* параметров функции Гаусса, а третий – M параметров w_i.

Выходной сигнал сети Ванга – Менделя рассчитывается по формуле:

$$y(x) = \frac{\sum_{i=1}^{M} w_i \prod_{j=1}^{N} \mu_{ij}(x_j)}{\sum_{i=1}^{M} \prod_{j=1}^{N} \mu_{ij}(x_j)}, \qquad (1)$$

где w_i – весовой коэффициент, $\mu_{ij}()$ – функция Гаусса с параметрами математического ожидания, которое определяет центр c_{ij} и параметрами разброса, которые определяются средним квадратическим отклонением d_{ij},

$$\mu_{ij}(x_j) = \frac{1}{1 + \frac{(x_j - c_{ij})^2}{d_{ij}^2}}$$ – функция Гаусса.

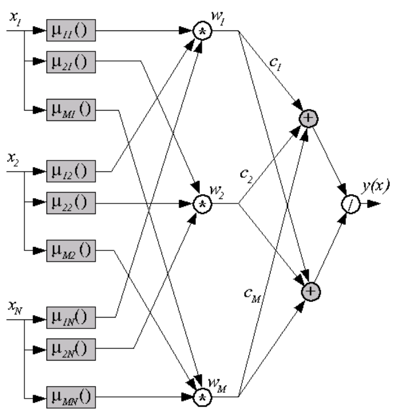

Рис. 1. Структура сети Ванга – Менделя

Задача сети состоит в построении такого отображения пар данных (*x*, *d*), чтобы ожидаемое значение, соответствующее входному вектору *x*, формировалось выходной функцией *y(x)*.

Обучение нечетких сетей, также как и классических сетей, может проводиться по алгоритму с учителем, основанному на минимизации целевой функции, задаваемой с использованием евклидовой нормы как

$E = \frac{1}{2}\sum_{i=1}^{p}(y(x^{(l)}) - d^{(l)})^2$, где *p* – количество обучающих пар (*x*, *d*).

Для обучения нечеткой нейронной сети применяют алгоритм, включающий последовательное чередование следующих шагов:

- для фиксированных значений параметров c_{ij} и d_{ij} первого слоя вычисляются значения параметров w_i третьего слоя сети;
- при зафиксированных значениях параметров w_i третьего слоя уточняются параметры c_{ij} и d_{ij} первого слоя сети.

Таким образом, на первом этапе для *K* обучающих выборок $< X^k, d^k >$, k=1, 2, … *K*, получаем систему *K* линейных уравнений $PV * W = D$, где *W* – вектор, составленный из линейных коэффициентов w_i, *D* – вектор эталонных ответов сети, $PV_i^k = \frac{\prod_{j=1}^{N} \mu_{ij}(x_j^k)}{\sum_{i=1}^{M}\prod_{j=1}^{N} \mu_{ij}(x_j^k)}$.

Количество строк *K* матрицы *PV* значительно больше количества ее столбцов. Решение этой системы линейных алгебраических уравнений может быть получено за один шаг следующим образом: $W = PV^+ * D$, где PV^+ – псевдообратная матрица для матрицы *PV*.

На втором этапе фиксируются значения коэффициентов полиномов третьего слоя и осуществляется уточнение (обычно многократное) коэффициентов функции Гаусса для первого слоя сети стандартным методом градиента: $c_{ij}^{k+1} = c_{ij}^k - \nu_c \frac{\partial E^k}{c_{ij}^k}$, $d_{ij}^{k+1} = d_{ij}^k - \nu_d \frac{\partial E^k}{d_{ij}^k}$, где k – номер очередного цикла обучения, ν_c – скорость обучения для коэффициентов c_{ij}, ν_d – скорость обучения для коэффициентов d_{ij}, $E = \frac{1}{2}\sum_{l=1}^{L}(y_l - y_l^e)^2$ – ошибка сети, где L – общее число обучающих выборок, y_l – выход сети Ванга-Менделя для данной выборки, y_l^e – эталонное значение выхода сети Ванга – Менделя [1,2].

Производные $\frac{\partial E}{\partial c_{ij}}$ и $\frac{\partial E}{\partial d_{ij}}$ вычисляются по формулам:

$$\frac{\partial E}{\partial c_{ij}} = \sum_{l=1}^{L}(y_l - y_l^e) * \frac{\partial y_l}{\partial c_{ij}}, \quad \frac{\partial E}{\partial d_{ij}} = \sum_{l=1}^{L}(y_l - y_l^e) * \frac{\partial y_l}{\partial d_{ij}}.$$

Производные $\frac{\partial y}{\partial c_{ij}}$ и $\frac{\partial y}{\partial d_{ij}}$ можно найти по формулам:

$$\frac{\partial y}{\partial c_{ij}} = \frac{2*(x_i - c_{ij})*\prod_i \mu_{ij}(x_i)*\sum_j((w_j - w_j)*\prod_i \mu_{ij}(x_i))}{d_{ij}^2(\sum_j\prod_i \mu_{ij}(x_i))^2}, \quad \frac{\partial y}{\partial d_{ij}} = \frac{2*(x_i - d_{ij})^2*\prod_i \mu_{ij}(x_i)*\sum_j((w_j - w_j)*\prod_i \mu_{ij}(x_i))}{d_{ij}^3(\sum_j\prod_i \mu_{ij}(x_i))^2},$$

где $\mu_{ij}(x_j)$ – функция Гаусса

Поскольку в череде этапов этап уточнения параметров функции Гаусса имеет много меньшую скорость сходимости, то в ходе обучения реализацию этапа 1, как правило, сопровождает реализация нескольких этапов 2.

Часто требуется найти «решение» системы, которая решений (в обычном смысле) не имеет. Выходом из ситуации является нахождение таких значений неизвестных параметров, что все условия системы выполняются «в некоторой степени».

Матрица A^+ называется псевдообратной матрицей для матрицы *A*, если $AA^+A = A$. Отсюда сразу вытекает, что если матрица *A* имеет размер *m* х *n*, то псевдообратная матрица A^+ имеет размер *n* х *m* .

Опишем и другой, часто встречающийся в литературе подход к определению этого понятия. Сначала введём понятие псевдорешения системы уравнений. Пусть нам дана система уравнений

$$Ax = b, \qquad (2)$$

где A - матрица размера m х n, b - вектор из m элементов.

Любое решение этой системы является также и решением системы

$$A^T Ax = A^T b, \qquad (3)$$

Псевдорешением системы (2) называется решение системы (3) с минимальной нормой среди всех столбцов, имеющих минимальную невязку (норма вектора равна квадратному корню из суммы квадратов компонент вектора, а невязкой решения системы (2) называется норма вектора Ax-b).

Псевдообратной матрицей для матрицы A размера m х n называется матрица A^+, столбцы которой – псевдорешения систем вида $Ax = e_i$,

где e_i – i-ый столбец единичной матрицы порядка m.

К универсальным способам нахождения псевдообратной матрицы относятся рекуррентные алгоритмы Гревиля и Фадеева. В данной работе приведем алгоритм Гревиля для псевдообращения матриц.

Пусть дана матрица $A \in R^{mxn}$ и a_k – ее k-й столбец, k = 1, . . ., n.

Пусть A_k – матрица, составленная из k первых столбцов матрицы A:

$$A_k = [a_1 \quad a_2 \quad \dots \quad a_k]$$

При k = 1: $A_1 = a_1$, а при k = 2, . . . , n: $A_k = [A_{k-1} \quad a_k]$; $A_n = A$.

Матрица $A^+ \in R^{nxm}$ может быть вычислена с помощью рекуррентного алгоритма:

1. Инициализация.

$$A_1^+ = \begin{cases} 0, \ \text{если} \quad a_1 \\ \dfrac{a_1^T}{\|a_1\|^2}, \ \text{иначе} \end{cases}$$

2. Цикл по k=2, …, n.

$$A_k^+ = \begin{bmatrix} A_{k-1}^+(I - a_k f_k) \\ f_k \end{bmatrix}$$, где I – единичная матрица порядка m,

$$f_k = \begin{cases} \dfrac{c_k^T}{\|c_k\|^2}, \ c_k = (I - A_{k-1}A_{k-1}^+)a_k, \ c_k \neq 0, \\ \dfrac{a_k^T (A_{k-1}^+)^T A_{k-1}^+}{1 + \|A_{k-1}^+ a_k\|^2}, \ c_k = 0 \end{cases}$$

Полученная на последнем шаге матрица A_n^+ и есть псевдообратная матрица, которая является искомым решением.

Принцип нечеткой логики достаточно давно используется для решения задач, в которых исходные данные являются слабо формализованными или же ненадежными. Основными преимуществами сетей с такой структурой являются:

- удобство представления информации: описание постановки задачи и условий производится на языке близком к естественному;
- универсальность: согласно теореме нечеткой аппроксимации, любая математическая модель может быть аппроксимирована системой, построенной на нечеткой логике;
- эффективность: ряд теорем, подобных теоремам о полноте для искусственных нейронных сетей, показывают высокую эффективность работы таких сетей.

Однако, такой организации нейронных сетей присущ и ряд недостатков:

- исходный набор нечетких правил формируется человеком, что не всегда является объективным, а иногда неполным или даже противоречивым;
- вид и параметры данных, связывающих вход и выход, также определяются субъективно и не всегда отражают действительность.

Каждый тип архитектуры интеллектуальных систем обладает своими особенностями в части обучения сети, обработки данных и вычисления конечного результата, что позволяет использовать одни типы архитектур для решения задач, к которым не применимы другие. Так, например, использование искусственных нейронных сетей в задачах по распознаванию образов имеет широкое применение, однако, объяснить принцип работы сетей достаточно сложно. Сети могут самостоятельно получать данные и обрабатывать их, однако, процесс обучения сетей достаточно долог, кроме того, анализ полученной в конечном итоге сети достаточно сложен. При этом, ввод в нейронную сеть какой-либо заранее достоверной информации не возможен [5].

Рассматривая системы, построенные на нечеткой логике, можно утверждать обратное – данные, получаемые на выходе таких систем, легки в понимании, однако, такие системы не могут самостоятельно получать информацию, которую можно использовать в дальнейшем при формировании выходных данных.

Как мы видим, искусственные нейронные сети и системы с нечеткой логикой схожи между собой, однако, каждая из них имеет свои достоинства и недостатки. Данный вывод был взят за основу при создании нечетких нейронных сетей. Такие сети строят решение на основе аппарата нечеткой логики, однако функции принадлежности настраиваются с помощью алгоритмов обучения искусственных нейронных сетей [3,5]. Кроме того, такие сети не только могут обучаться, но и способны учитывать априорную информацию. По своей структуре нечеткие нейронные сети схожи с многослойными сетями, например, с сетью, обучающейся по алгоритму обратного распространения, но скрытые слои в нечетких сетях соответствуют этапам работы нечеткой системы: первый слой производит введение нечеткости, исходя из заданных признаков входов; второй слой определяет множество нечетких правил; третий слой выполняет функцию приведения к четкости. В каждом из указанных слоев имеется набор параметров, настройка которых производится так же, как и настройка обычной нейронной сети.

Список литературы

1. Аксенов С.В., Новосельцев В.Б. Организация и использование нейронных сетей (методы и технологии) / Под общ. ред. В.Б. Новосельцева. – Томск: Изд-во НТЛ, 2006. – 128 с.

2. Батыршин И.З. Нечеткие гибридные системы. Теория и практика / Под ред. Н.Г. Ярушкиной. – М.ФИЗМАТЛИТ, 2007. – 208 с.

3. Осовский С. Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.

4. Хайкин С. Нейронные сети: полный курс: Пер. с англ. 2-е изд. – М.: Издательский дом "Вильяме", 2006. – 1104 с.: ил. – Парад, тит. англ.

5. Яхъева Г.Э. Нечеткие множества и нейронные сети: Учебное пособие / Г.Э. Яхъева. – М.: Интернет-Университет Информационных технологий; БИНОМ. Лаборатория знаний, 2006. – 316 с.

Рецензенты:

Шашкин А.И., д.ф.-м.н., зав. кафедрой математического и прикладного анализа ФГБОУ ВПО «Воронежский государственный университет», г. Воронеж.

Кургалин С.Д., д.ф.-м.н., зав. кафедрой цифровых технологий ФГБОУ ВПО «Воронежский государственный университет», г. Воронеж.