

О ПРЕОБРАЗОВАНИИ ДЕРЕВЬЕВ РЕШЕНИЙ В МНОГОСЛОЙНЫЕ НЕЙРОННЫЕ СЕТИ

Шибзухов З.М., Шеожев А.М.

НИИ прикладной математики и автоматизации КБНЦ РАН,
360000, Россия, КБР, г. Нальчик, ул. Шортанова, 189-а
sz@zmail.ru, alik@w3.to

ABSTRACT

Problem of the transformation of the decision tree structure into architecture of multilayer neural network is considered. The methods of the transformation of the decision tree into 3- or 4-layer neural network are suggested. We build these networks using threshold-, Σ - and Π -elements.

Одним из способов синтеза нейросетевых архитектур является преобразование дерева решений [1,2] в многослойную нейронную сеть [3,4].

Пусть y_1, \dots, y_m – набор целевых признаков, y_i ($1 \leq i \leq m$) принимает целочисленные значения в интервале $[0, r_i]$, x_1, \dots, x_n – набор наблюдаемых признаков, x_j ($1 \leq j \leq n$) принимает целочисленные значения в интервале $[0, q_j]$. В зависимости от конкретной постановки 0 может интерпретироваться, как “неизвестное значение” или “отсутствующее значение” (“пропуск”).

Дерево решений содержит вершины двух типов:

I. *решающие вершины*, которые описываются правилами вида:

$$\rho: y_{i(1)} = y_1 \wedge \dots \wedge y_{i(l)} = y_l \wedge \rho',$$

где ρ – метка текущего правила, $y_{i(1)}, \dots, y_{i(l)}$ – целевые переменные с индексами $1 \leq i(1) < \dots < i(l) \leq m$, соответственно, $y_1 \in [0, r_{i(1)}], \dots, y_l \in [0, r_{i(l)}]$ – их значения, ρ' – метка правила для нахождения остальных решений;

II. *селекторные вершины*, которые описываются правилами вида:

$$\rho: x = x_1 \rightarrow \rho_1 \vee \dots \vee x = x_k \rightarrow \rho_k \vee x \notin \{x_1, \dots, x_k\} \rightarrow \rho_{k+1},$$

где ρ – метка текущего правила, x переменная из набора входных переменных, x_1, \dots, x_k – ее значения, $\rho_1, \dots, \rho_k, \rho_{k+1}$ – метки других правил. Правило интерпретируется следующим образом: если значение переменной x равно x_l для некоторого $1 \leq l \leq k$, то поиск решения продолжается по правилу с меткой ρ_l , в противном случае решение ищется по правилу ρ_{k+1} .

Среди селекторных вершин выделяется одна, называемая корнем дерева. Поиск решения начинается в корне и завершается в одной из решающих вершин.

Строится многослойная нейронная сеть полиномиального типа, эквивалентная исходному дереву решений: на одинаковых наборах значений наблюдаемых признаков они возвращают одинаковые значения целевых признаков. Архитектура таких нейронных сетей будет включать два специальных слоя.

Первый – *слой δ -элементов*, который состоит из элементов вида:

$$\delta_{i,j} = \delta(x_i - j),$$

где x_i , ($1 \leq i \leq n$) – входной признак, $0 \leq j < q_i$, $\delta(x)$ – характеристическая функция единицы,

$$\delta(x) = \begin{cases} 1, & \text{если } x=0; \\ 0, & \text{иначе.} \end{cases}$$

Второй специальный слой, который условно назовем *else-слой*, состоит из Σ -элементов. Они строятся для каждой селекторной вершины, правило которой содержит else-часть

(т.е. правило $x \notin \{x_1, \dots, x_k\} \rightarrow \rho_{k+1}$.) Соответствующий такой вершине Σ -элемент имеет вид:

$$\varepsilon_\rho = 1 - \sum_{j \in J_\rho} \delta_{i,j},$$

где i – индекс переменной x_i , с которой ассоциирована селекторная вершина, J_ρ – множество значений этой переменной, с которыми ассоциированы поддеревья дерева решений. Архитектура многослойных нейронных сетей включает также *ассоциативный слой*, состоящий из П-элементов. Они строятся для каждой ветви, начинающейся в корне дерева решений и заканчивающейся в решающей вершине. Соответствующие П-элементы имеют вид:

$$u_\beta = \prod_{(i,j) \in S_\beta} \delta_{i,j} \prod_{\rho \in R_\beta} \varepsilon_\rho,$$

где β – ветвь дерева, S_β – множество пар индексов (i,j) , которые соответствуют дугам, выходящим из селекторных узлов ветви β , R_β – множество меток правил, которые описывают селекторные узлы с else-дугами, лежащими на ветви.

Последний слой – *решающий*, состоит из m Σ -элементов. k -ый элемент, вычисляющий значение y_k имеет вид:

$$y_k = \sum_{\beta} y_{k,\beta} u_\beta,$$

где β пробегает ветви дерева, в окончании которых в решающем узле есть решение вида $y_k = y_{k,\beta}$.

В частном случае, когда во всех селекторных вершинах дерева решений отсутствует else-часть, архитектура НС упрощается за счет исключения else-слоя. В результате получается 3-слойная сеть:

I слой: $\delta_{i,j} = \delta(x_i - j)$,

II слой: $u_\beta = \prod_{(i,j) \in S_\beta} \delta_{i,j}$,

III слой: $y_k = \sum_{\beta} y_{k,\beta} u_\beta$,

Поскольку все $\delta_{i,j}$, $\varepsilon_\rho \in \{0,1\}$, то в тех случаях, когда аппаратная реализация П-элементов вызывает затруднения, их можно заменить на композиции пороговых элементов и Σ -элементов. Соответствующие композиции будут иметь вид:

$$u_\beta^* = \text{sign} \left[\left(\sum_{(i,j) \in S_\beta} \delta_{i,j} + \sum_{\rho \in R_\beta} \varepsilon_\rho \right) - |S_\beta| - |J_\beta| \right],$$

где $|S_\beta|$, $|J_\beta|$ – количество элементов множеств S_β и J_β соответственно. В том случае, если отсутствует else-часть, данные элементы будут иметь вид:

$$u_\beta^* = \text{sign} \left[\sum_{(i,j) \in S_\beta} \delta_{i,j} - |S_\beta| \right].$$

На практике число входов П-элемента всегда ограничено, в то время как глубина дерева (а следовательно и число входов элементов ассоциативного слоя) может превышать эту величину. Для решения этой проблемы необходимо ввести дополнительные слои, состоящие из П-элементов. Пусть D – глубина дерева, d – максимально возможное число входов П-элемента. Тогда ассоциативный слой заменяется на $L = \lceil \log_d (D-1) \rceil + 1$ слоев из П-элементов.

ЛИТЕРАТУРА

1. Тимофеев А.В., Шибзухов З.М. Методы синтеза и оптимизации баз знаний по базам данных на основе локально-оптимальных логико-вероятностных алгоритмов. – International journal "Information theories and applications". 1995. Vol. 3. № 2. Pp 12-19.
2. Тимофеев А.В., Шибзухов З.М. Методы синтеза и оптимизации баз знаний по базам данных на основе локально-оптимальных алгоритмов. – Proc. XXII International conference CAD-95 (Grimea, Gurzuff, may, 1995). Part I. Pp.28.
3. Тимофеев А.В. Методы синтеза диофантовых нейросетей минимальной сложности. Доклады АН, 1995, т.345, № 1, с. 32-35.
4. Тимофеев А.В., Шибзухов З.М. Методы синтеза и минимизации сложности диофантовых нейронных сетей над конечным полем. – Автоматика и телемеханика, 1997, № 4, с.204-212.
5. Шибзухов З.М., Шеожев А.М. О преобразовании деревьев решений в многослойные нейронные сети полиномиального типа// Тезисы конференции “Нейроинформатика-2000”. – М., 2000.
6. Шибзухов З.М., Шеожев А.М. Преобразование деревьев решений в многослойные нейронные сети полиномиального типа// Тезисы IV Всероссийского симпозиума “Математическое моделирование и компьютерные технологии”. – Кисловодск, 2000.