

УДК 004.93

Є. О. Гофман, С. О. Субботін, А. О. Олійник

Запорізький національний технічний університет
вул. Жуковського, 64, 69063 Запоріжжя, Україна

Метод синтезу нейро-нечітких мереж із використанням дерев розв'язків

Виконано аналіз методів побудови нейро-нечітких мереж. Запропоновано метод синтезу нейро-нечітких мереж, який на основі побудованого дерева розв'язків визначає структуру мережі та обчислює параметри нейроелементів. Проведено експерименти по розв'язанню практичних задач.

Ключові слова: *дерево розв'язків, нейро-нечітка мережа, функція належності, база правил.*

Вступ

При розробці систем обробки даних часто виникає необхідність побудови моделей об'єктів і процесів, що досліджуються [1]. Як розпізнаючі моделі ефективно можуть застосовуватися нейро-нечіткі (нейро-фаззи) мережі, що представляють собою систему нечіткого виведення у вигляді нейронної мережі, зручної для навчання, поповнення знань, аналізу та використання [2–4].

Проте синтез нейро-нечітких мереж (ННМ) на основі навчальної вибірки шляхом відображення її екземплярів як правил у структуру мережі приводить до побудови громіздкої та складної моделі з дуже низькими узагальнюючими властивостями. Велика складність синтезованої мережі, що викликана великою кількістю структурних елементів (нейронів) та їхніх параметрів, призводить до необхідності використання значного обсягу пам'яті, а також до збільшення витрат часу роботи мережі при послідовній реалізації обчислень на ЕОМ [2–6]. Крім того, необхідність виділення нечітких термів при синтезі ННМ вимагає або участі користувача, що знижує рівень автоматизації розроблюваної інформаційної технології, або вирішення задачі кластерного аналізу, коли терми визначаються як проєкції меж кластерів на осі ознак. ННМ, як правило, навчаються за допомогою градієнтних методів, які є методами локальної оптимізації і характеризуються високою ітеративністю, а також невизначеністю вибору початкової точки пошуку.

Тому виникає необхідність розробки нових методів побудови ННМ, що вільні від зазначених недоліків. У представленій роботі для синтезу ННМ пропонується використовувати дерева розв'язків (ДР), що представляють собою один з

методів аналізу даних, який задає спосіб представлення правил виду «Якщо — то» в ієрархічній послідовній структурі, де кожному об'єкту відповідає єдиний вузол з рішенням [4, 7–10]. Доцільність застосування ДР для синтезу ННМ обумовлюється наступними причинами:

— у процесі побудови ДР виявляється набір інформативних ознак (вирішується задача відбору інформативних ознак), що дозволяє будувати прості моделі. Крім того, ознаки ранжуються за ступенем їхньої важливості;

— ДР може бути перетворено на набір продукційних правил, число яких за рахунок узагальнення може бути меншим, ніж кількість екземплярів, що також знижує громіздкість і складність синтезованої моделі. Підвищити узагальнення додатково можливо за рахунок використання процедур усікання ДР;

— перевірки в ДР можуть бути використані для визначення числа термів за кожною змінною і параметрів інтервалів чіткого розбиття. Значення знайдених меж інтервалів (на основі вузлів ДР) при відображенні правил ДР у ННМ дозволяють здійснити не тільки структурну ідентифікацію, а й задати початкову точку пошуку для градієнтних методів, яка буде близькою до оптимальної;

— ДР може бути також використано для побудови ієрархічних мереж.

Метою даної роботи є розробка методу синтезу нейро-нечітких мереж на основі дерев розв'язків.

Постановка задачі

Нехай задана навчальна вибірка $\langle X, Y \rangle$, де $X = \{X_i\}$ — набір значень ознак, що характеризують даний об'єкт або процес; $Y = \{y_p\}$ — масив значень вихідного параметра в заданій вибірці; $X_i = \{x_{ip}\}$ — i -а ознака у вибірці, $i = 1, 2, \dots, L$; $x_{ip} \in [x_{\min i}; x_{\max i}]$ — значення i -ї ознаки для p -го екземпляра вибірки, $p = 1, 2, \dots, m$; $x_{\min i}$ та $x_{\max i}$ — мінімальне та максимальне значення i -ї ознаки, відповідно; $y_p \in D_Y$ — значення прогнозованого параметра для p -го екземпляра; $D_Y = \{y^{(t)}\}$ — множина значень вихідного параметра, $t = 1, 2, \dots, N_y$; L — загальне число ознак у вихідному наборі; m — число екземплярів вибірки; $N_y = |D_Y|$ — число значень вихідного параметра Y в множині D_Y .

Тоді задача синтезу розпізнаючої моделі на основі ННМ полягає в ідентифікації структури ННМ і значень її параметрів, для яких досягається мінімум заданого значення критерію оптимальності $\xi(\text{ННМ}, X, Y) \rightarrow \min$, де $\xi(\text{ННМ}, X, Y)$ — критерій, що визначає ефективність використання ННМ для апроксимації залежності між набором вхідних параметрів X та відповідному йому вектору значень вихідного параметра Y .

Як правило, як критерій оптимальності ННМ використовується середньоквадратична помилка:

$$\varepsilon = \sum_{p=1}^m (y_p - y(\text{ННМ}, X_p))^2,$$

де X_p — набір значень ознак для p -го екземпляру; $y(\text{ННМ}, X_p)$ — значення виходу ННМ, обчислене для набору значень X_p .

Метод синтезу нейро-нечітких мереж на основі дерев розв'язків

У розробленому методі побудови ННМ пропонується використовувати ДР, що синтезовано на основі заданої вибірки даних, для витягання продукційних правил вигляду «Якщо – то», обчислення параметрів функцій належності, завдання структури і налаштування параметрів нейроелементів ННМ. Синтез ННМ пропонується виконувати у вигляді такої послідовності етапів.

Етап 1. Побудова бази правил типу «Якщо — то». На даному етапі за навчальною вибіркою $\langle X, Y \rangle$ за допомогою методів синтезу ДР [4, 5, 7, 11, 12] виконується побудова дерева розв'язків $DT = \{d_k\}$, де $d_k = \langle cd_k, ld_k, rd_k \rangle$ — k -й вузол дерева DT , що представляє собою структуру, в якій cd_k — функція прийняття рішень (умова типу $x_i \in A_{ij}$) на основі значень вхідних змінних (у випадку, якщо вузол є внутрішнім) або значення вихідної змінної (для зовнішніх вузлів); ld_k та rd_k — посилання на лівого та правого нащадків k -го вузла відповідно, що представляють собою структури, аналогічні d_k ; $A_{ij} = [l_{ij}; r_{ij}]$ — інтервал значень ознаки x_i з умови cd_k , що обмежується значеннями l_{ij} та r_{ij} .

Для побудови набору правил R шляхом вилучення правил з ДР виконується прямий обхід дерева: від кореня дерева до кожного вузла. При цьому для кореня дерева, як першого вузла-батька, викликається рекурсивно процедура обходу лівого та правого піддерев. У процесі обходу дерева DT при кожному відвідуванні нового вузла d_k відбувається модифікація поточного антецедента $Antecedent_j$ правила R_j шляхом додавання умови cd_k , що знаходиться у вузлі d_k :

$$Antecedent_j = Antecedent_j \cap cd_k.$$

У випадку, коли в умові cd_k значення деякої ознаки X_i виходить за межі діапазону її можливих значень $[x_{\min i}; x_{\max i}]$, виконується коригування лівої l_{ij} та правої r_{ij} меж значень ознаки X_i в умові cd_k :

$$l_{ij} = x_{\min i}, \text{ якщо } l_{ij} < x_{\min i},$$

$$r_{ij} = x_{\max i}, \text{ якщо } r_{ij} > x_{\max i}.$$

Коли вузол d_k виявляється листом (посилання на лівого та правого нащадків є порожніми), виконується поповнення бази правил R новим правилом R_j . При цьому як антецедент приймається $Antecedent_j$, а як консеквент — значення cd_k листа дерева:

$$\text{Antecedent}_{j+1} = \text{Antecedent}_j / cd_k.$$

Обхід дерева виконується до тих пір, поки не будуть пройдені всі його листя (кожен вузол дерева відвідується до тих пір, поки не відвідані всі його нащадки).

Таким чином, у результаті виконання першого етапу пропонованого методу виконується формування бази правил R типу «Якщо — то», де кожне j -е правило формується на основі умов (перевірок) типу $x_i \in A_{ij}$, що відносяться до гілок ДР, які розташовані на шляху від кореня до j -го листа:

$$\text{Правило 1: Якщо } \bigcap_{i=1}^L x_i \in A_{i1}, \text{ то } y = y_1,$$

$$\text{Правило 2: Якщо } \bigcap_{i=1}^L x_i \in A_{i2}, \text{ то } y = y_2,$$

.....

$$\text{Правило } m: \text{Якщо } \bigcap_{i=1}^L x_i \in A_{im}, \text{ то } y = y_m,$$

де A_{ij} — інтервал (множина) значень ознаки X_i , при яких виконується j -е правило.

Етап 2. Синтез блоків першого шару ННМ, що визначають належність розпізнаваного екземпляру до термів ознак. Для цього обчислюються ліва l_{ij} і права r_{ij} межі діапазонів значень кожної ознаки X_i в правилі R_j .

Далі для кожної i -ї ознаки X_i визначається число інтервалів розбиття діапазону його значень N_i по всій множині правил R .

Потім задаються функції належності розпізнаваного екземпляру до кожного з термів. При цьому використовуються межі чітких інтервалів $\Delta_{ij} = [l_{ij}; r_{ij}]$, що були знайдені раніше.

Як функції належності доцільно використовувати такі функції, які дозволяють обмежувати інтервал значень ознак:

— трапецієподібну функцію:

$$\mu_{ik} = \begin{cases} 0, & \text{якщо } x_i < a_{ik}, \\ \frac{x_i - a_{ik}}{l_{ik} - a_{ik}}, & \text{якщо } a_{ik} \leq x_i < l_{ik}, \\ 1, & \text{якщо } l_{ik} \leq x_i \leq r_{ik}, \\ \frac{b_{ik} - x_i}{b_{ik} - r_{ik}}, & \text{якщо } r_{ik} < x_i \leq b_{ik}, \\ 0, & \text{якщо } x_i > b_{ik}, \end{cases}$$

де значення a_{ik} , b_{ik} , r_{i0} , l_{iN_i+1} визначаються за формулами:

$$a_{ik} = \frac{3r_{ik-1} + l_{ik}}{4}, b_{ik} = \frac{r_{ik} + 3l_{ik+1}}{4}, r_{i0} = l_{i1} - 2, l_{iN_i+1} = l_{iN_i} + 2;$$

— П-подібну функцію:

$$\mu_{ik} = \mu S_{ik} \cdot \mu Z_{ik},$$

де μS_{ik} та μZ_{ik} — S-подібна та Z-подібна функції належності відповідно:

$$\mu S_{ik} = \begin{cases} 0, & \text{якщо } x_i < l_{ik}, \\ \frac{x_i - l_{ik}}{r_{ik} - l_{ik}}, & \text{якщо } l_{ik} \leq x_i \leq r_{ik}, \\ 1, & \text{якщо } x_i > r_{ik}, \end{cases}$$

$$\mu Z_{ik} = \begin{cases} 1, & \text{якщо } x_i < l_{ik}, \\ \frac{r_{ik} - x_i}{r_{ik} - l_{ik}}, & \text{якщо } l_{ik} \leq x_i \leq r_{ik}, \\ 0, & \text{якщо } x_i > r_{ik}. \end{cases}$$

Етап 3. Побудова другого шару ННМ, що складається з нейронів, які реалізують рядки-кон'юнкції антецедентів правил нечіткої бази знань. Виходи кожного нейрона цього шару визначають ступені виконання умов правил для розпізнаваного екземпляру.

При обчисленні ступенів виконання антецедентів правил доцільно враховувати достовірність (вагу, надійність, впевненість) кожного правила. Тому при розрахунку виходів нейронів другого шару пропонується використовувати також коефіцієнт упевненості кожного правила, що дозволить посилити внесок більш надійних правил у розв'язок.

Для обчислення коефіцієнта упевненості γ_j j -го правила пропонується по всій вибірці $\langle X, Y \rangle$ знайти число екземплярів n_j , для яких спрацювало j -е правило, а також визначити $n_{\text{ош } j}$ — число помилкових рішень при використанні даного правила.

Коефіцієнт упевненості j -го правила γ_j будемо розраховувати за формулою

$$\gamma_j = \alpha_j P_{\text{прав } j},$$

де $\alpha_j = \frac{n_j}{m}$ — ймовірність спрацювання j -го правила; $P_{\text{прав } j} = 1 - P_{\text{ош } j} = 1 - \frac{n_{\text{ош } j}}{n_j}$

— ймовірність правильних рішень при використанні j -го правила.

Тоді ступінь відповідності виконання антецедента j -го правила для розпізнаваного екземпляру може бути розрахована наступним чином:

$$\mu_{Aj} = \alpha_j \min \left\{ \max_{i,k} \left\{ w_{ik}^{(2j)}; \mu_{ik} \right\} \right\},$$

де μ_{Aj} — ступінь виконання антецедента j -го правила для розпізнаваного екземпляру; $w_{ik}^{(2j)}$ — ваги нейронів другого шару, що визначають наявність зв'язку від нейронів першого шару до другого і відображають наявність k -го терму i -ї ознаки в антецеденті j -го правила. Ваги $w_{ik}^{(2j)}$ визначаються за синтезованим ДР наступним чином: $w_{ik}^{(2j)} = 0$, якщо на шляху від кореня ДР до j -го листа є умова, за якої i -а ознака потрапляє в k -й інтервал діапазону її розбиття (k -й терм i -ї ознаки присутній в описі умов j -го правила), в іншому випадку: $w_{ik}^{(2j)} = 1$.

Етап 4. Визначення параметрів нейронів третього шару ННМ, які обчислюють ступені належності вхідного вектора до відповідних термів вихідної змінної, використовуючи формулу

$$\mu_{y^{(t)}} = \min \left\{ \max_j \left\{ w_j^{(3t)}; \mu_{Aj} \right\} \right\},$$

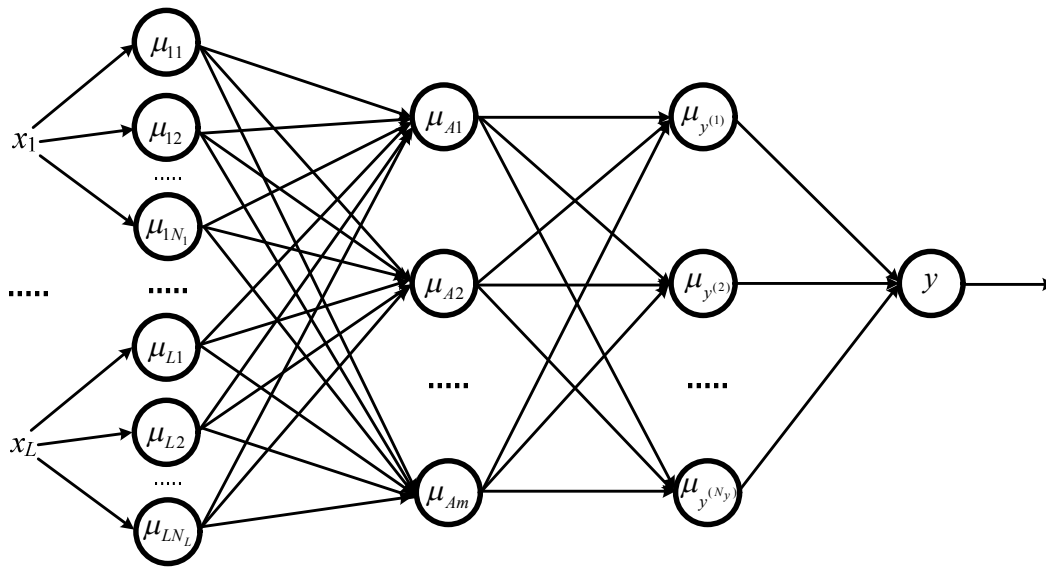
де $\mu_{y^{(t)}}$ — ступінь належності вхідного вектора до t -го терму вихідної змінної y ; $w_j^{(3t)}$ — ваги нейронів третього шару, що визначають наявність зв'язку від нейронів другого шару до третього та відображають відповідність t -го терму вхідного параметра до j -го правила нечіткої бази знань. Ваги $w_j^{(3t)}$ визначаються за синтезованим ДР: $w_j^{(3t)} = 0$, якщо значення j -го листа ДР дорівнює t -му значенню вхідного параметра $y^{(t)}$ з множини його значень D_y , в іншому випадку: $w_j^{(3t)} = 1$.

Етап 5. Дефазифікація. Приведення результату в чітку форму здійснюється за допомогою нейрона четвертого шару ННМ, вихід якого може бути обчислений за допомогою центроїдного методу [4] за формулою:

$$y = \frac{\sum_{j=1}^m y_j \mu_{y^{(t)}}}{\sum_{j=1}^m \mu_{y^{(t)}}}.$$

Схема ННМ, яка побудована за допомогою запропонованого методу, наведена на рисунку.

Запропонований метод побудови ННМ на основі ДР був програмно реалізований у комп'ютерній програмі «Автоматизована система синтезу дерев рішень на основі інтелектуальних обчислень для неруйнівного діагностування та класифікації» на мові програмування C#.



Структура ННМ, яка побудована за допомогою запропонованого методу

За допомогою розробленого методу та програмного забезпечення вирішувалася практична задача діагностування кузовів автотранспортних засобів [11, 12]. Експериментальні дослідження проводилися на основі вибірки даних, яка формувалася за результатами вимірювань для кузовів автомобілів, виготовлених на підприємстві ЗАТ «ЗАЗ». Для синтезу ННМ була сформована вибірка зі 172 екземплярів, кожен з яких описувався 48 ознаками. Вихідний параметр приймав два значення: «0» — кузов необхідно допрацьовувати, «1» — кузов вважається придатним до експлуатації.

У процесі експериментального дослідження проводилося порівняння таких методів синтезу ННМ:

1) метод, при використанні якого вибірка даних відображається безпосередньо в правила (кожен екземпляр вибірки перетворюється в одне правило бази правил), таким чином формується структура мережі. Налаштування параметрів нейроелементів здійснюється за допомогою методу зворотного поширення помилки [4];

2) метод, який передбачає кластеризацію вибірки. При використанні такого підходу на основі знайдених центрів кластерів формуються правила, які відображаються в структуру ННМ. Параметри блоків ННМ настроюються за допомогою методу зворотного поширення помилки [4];

3) запропонований метод синтезу ННМ на основі дерев розв'язків.

Результати експериментів показали, що порівнювані методи характеризувалися наступними похибками: 3,1 %, 1,7 % і 1,4 % відповідно.

Важливо відзначити, що ННМ, синтезована на основі запропонованого методу, характеризувалася меншим числом нейроелементів, що забезпечується за рахунок використання дерев розв'язків, узагальнюючих у вигляді бази правил інформацію, що міститься у вибірці вихідних даних. Час побудови ННМ і час обчислення вихідного параметра за синтезованою мережею в другому та третьому методах є істотно меншим, ніж у першому, що пояснюється меншою кількістю нейроелементів.

Також за допомогою запропонованого методу було розв'язано ряд практичних задач [1, 6], основні характеристики яких та результати їхнього вирішення наведено в таблиці.

Практичні задачі та результати проведення експериментів

Назва задачі	Кількість вхідних ознак	Кількість екземплярів	Похибка побудованої ННМ за допомогою методу, %		
			1	2	3
Діагностування лопаток турбіни високого тиску газотурбінного авіадвигуна	515	32	3,1	6,25	3,1
Моделювання залежності частоти власних коливань від геометричних параметрів пера лопаток	33	100	3,0	4,0	2,0
Моделювання взаємозв'язків параметрів дослідів авіадвигунів у процесі проливики соплових апаратів	6	491	3,1	2,6	1,2

Таким чином, запропонований метод синтезу ННМ на основі дерев розв'язків дозволяє будувати моделі з відносно невеликим числом нейроелементів, що мають високий рівень узагальнення. На відміну від градієнтних методів налаштування параметрів мережі, запропонований метод не є високоітеративним, не відноситься до методів локальної оптимізації і не пов'язаний з проблемою невизначеності вибору початкової точки пошуку.

Висновки

У роботі вирішено актуальну задачу автоматизації побудови діагностичних і розпізнаючих моделей на основі нейро-нечітких мереж.

Наукова новизна роботи полягає в тому, що запропоновано новий метод синтезу нейро-нечітких мереж на основі дерев розв'язків, який для побудованого за навчальною вибіркою ДР автоматично формує нечітке розбиття ознакового простору і виділяє правила, на основі яких формує структуру мережі і визначає значення її параметрів, що дозволяє синтезувати моделі з невеликою кількістю нейроелементів, які характеризуються високим рівнем узагальнення, і не вимагає рішення задач оптимізації для налаштування значень параметрів моделі.

Практична цінність отриманих результатів полягає в тому, що на основі запропонованого методу розроблено програмне забезпечення, яке дозволяє синтезувати нейро-нечіткі мережі, а також вирішено практичну задачу діагностування автотранспортних засобів.

1. *Прогрессивные технологии моделирования, оптимизации и интеллектуальной автоматизации этапов жизненного цикла авиационных двигателей: монография* / [А.В. Богуслаев, Ал.А. Олейник, Ан.А. Олейник и др.]; под ред. Д.В. Павленко, С.А. Субботина. — Запорожье: ОАО «Мотор Сич», 2009. — 468 с.

2. *Гибридные* нейро-фаззи модели и мультиагентные технологии в сложных системах: монография / [В.А. Филатов, Е.В. Бодянский, В.Е. Кучеренко и др.]; под общ. ред. Е.В. Бодянского. — Дніпропетровськ: Системні технології, 2008. — 403 с.
3. *Nauck D.* Foundations of Neuro-Fuzzy Systems / D. Nauck, F. Klawonn, R. Kruse. — Chichester: John Wiley & Sons, 1997. — 305 p.
4. *Субботін С.О.* Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень: навч. посібник / С.О. Субботін. — Запоріжжя: ЗНТУ, 2008. — 341 с.
5. *Encyclopedia of Artificial Intelligence* / Eds.: J.R. Dopico, J.D. De la Calle, A.P. Sierra. — New York: Information Science Reference, 2009. — Vol. 1–3. — 1677 p.
6. *Субботін С.О.* Неітеративні, еволюційні та мультиагентні методи синтезу нечіткологічних і нейромережних моделей: монографія / С.О. Субботін, А.О. Олійник, О.О. Олійник; під заг. ред. С.О. Субботіна. — Запоріжжя: ЗНТУ, 2009. — 375 с.
7. *Rokach L.* Data Mining with Decision Trees. Theory and Applications / L. Rokach, O. Maimon. — London: World Scientific Publishing Co, 2008. — 264 p.
8. *Люгер Дж.Ф.* Искусственный интеллект: стратегии и методы решения сложных проблем / Дж.Ф. Люгер. — М.: Вильямс, 2005. — 864 с.
9. *Quinlan J.R.* Induction of Decision Trees / J.R. Quinlan // *Machine Learning*. — 1986. — N 1. — P. 81–106.
10. *A Comparison of Decision Tree Ensemble Creation Techniques* / R. Banfield, L.O. Hall, K.W. Bowyer, W.P. Kegelmeyer // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. — 2007. — Vol. 29 (1). — P. 173–180.
11. *Гофман Е.А.* Использование деревьев решений для диагностирования автотранспортных средств / Е.А. Гофман, А.А. Олейник, С.А. Субботин // Информационные управляющие системы и компьютерный мониторинг: II Международная научно-техническая конференция ИУС и КМ-2011, 11–13 апреля 2011 г.: матер. конф. — Донецк, 2011. — Т. 1. — С. 159–163.
12. *Гофман Е.А.* Эволюционный метод синтеза деревьев решений / Е.А. Гофман, А.А. Олейник, С.А. Субботин // *Штучний інтелект*. — 2011. — № 2. — С. 6–14.

Надійшла до редакції 09.02.2012