Результати обчислень концентрації компонента SiO_2 при T_0=I,I8; $\bar{\varkappa t}$ = 0,5; $\bar{D}$ = I показані на рисунку. Криві I,2,3 відповідають значенням κ^T= 0,4; κ^T= 0,3; κ^T = 0,I.

I. А л е к с е е в В.В., П р и ш и н А.М. Физическая газодинамика реагирующих сред. М., I979. 2. И с а е в С.И. Теория тепломассообмена: Учебник для вузов /Под ред. А.И.Леонтьева. М., I979.

Стаття надійшла до редколегії I2.02.9I

УДК 519.92.

С.Г.Дещенко, Б.А.Остудін

ЗАСТОСУВАННЯ МЕТОДІВ ТЕОРІЇ НЕЧІТКИХ МНОЖИН ПРИ РОЗВ"ЯЗУВАННІ НЕКОРЕКТНИХ ЗАДАЧ КЛАСИФІКАЦІЇ

Існують різні способи розв"язування задач класифікації та оцінки їх якості. Алгоритми класифікації, описані в літературі, вимагають визначення наперед кількості класів і, як правило, залежать від порядку розгляду початкових даних, що, в свою чергу, ускладнює аналіз результатів. Початкові дані при цьому описують за допомогою кількісних ознак.

Ми ж пропонуємо алгоритм класифікації для малого об"єму початкових даних, коли не можна використати класичну теорію ймовірностей. При цьому результати виконання алгоритму не залежать від порядку перегляду початкових даних, які містять ознаки, отримані за допомогою шкали порядку та шкали найменувань. Зауважимо також, що розглядається ситуація, коли кількість класів є невідомою.

I. Аналіз проблеми та необхідні відомості з теорії нечітких множин

Нехай M – множина початкових даних, яку необхідно розбити на класи подібних елементів. Припустимо, що $M=M_1\times M_2\times\ldots\times M_n$ і $card\, M = card\, M_i = m$, $1\leqslant i\leqslant n$ де m – потужність множини M та множини значень i -ї ознаки елементів M – M_i. Отже, $\forall x_p\in M$, $x_p=(x_{p_1}, x_{p_2}, \ldots, x_{p_n})$, де $x_{p_i}\in M_i$. Нехай також r – міра подібності між елементами M, яка постає як відображення $M\times M\to R^+$ таке, що $r(x,y)=r(y,x)$, $\forall x, y\in M$; $r(x,x)=1$, $\forall x\in M$.

Для визначення конкретної міри подібності розглянемо відображення $r_j^н : M \times M \to R^+$. При цьому будемо вважати, що M_j – множина значень j -ї ознаки, яка отримана за допомогою шкали найменувань. Нехай

$$r_j^н(x_p, x_q) = \frac{1}{m} \sum_{k=1}^{m} d^k(x_{p_j}, x_{q_j}), \qquad x_p, x_q \in M, \qquad /1/$$

де

$$d^k(x_{p_j}, x_{q_j}) = 1 \Leftarrow (x_{pj} = x_{kj} \wedge x_{qj} = x_{kj}) \vee (x_{pj} \neq x_{kj} \wedge x_{qj} \neq x_{kj}); \qquad /2/$$

$$d^k(x_{p_j}, x_{q_j}) = 0 \Leftarrow (x_{pj} = x_{kj} \wedge x_{qj} \neq x_{kj}) \vee (x_{pj} \neq x_{kj} \wedge x_{qj} = x_{kj}). \qquad /3/$$

Звідси $0 \leq r_j^н \leq 1$.

Якщо M_j – множина значень ℓ -ї ознаки, яка отримана за допомогою шкали порядку, то відображення $r_\ell^п : M_\ell \times M_\ell \to R^+$ – визначимо наступним чином:

$$r_\ell^п(x_p, x_q) = \frac{1}{m} \sum_{k=1}^{m} d^k(x_{p\ell}, x_{q\ell}), \qquad x_p, x_q \in M, \qquad /4/$$

де

$$d^k(x_{p\ell}, x_{q\ell}) = 1 \Leftarrow (x_{p\ell} > x_{k\ell} \wedge x_{q\ell} > x_{k\ell}) \vee$$

$$\vee (x_{p\ell} < x_{k\ell} \wedge x_{q\ell} < x_{k\ell}) \vee (x_{p\ell} = x_{k\ell} \wedge x_{q\ell} = x_{k\ell};$$

$$d^k(x_{p\ell}, x_{q\ell}) = 0 \Leftarrow (x_{p\ell} > x_{k\ell} \wedge x_{q\ell} < x_{k\ell}) \vee \qquad /5/$$

$$\vee (x_{p\ell} < x_{k\ell} \wedge x_{q\ell} > x_{k\ell});$$

$$d^k(x_{p\ell}, x_{q\ell}) = 0.5 \Leftarrow (x_{q\ell} \wedge x_{q\ell} \neq x_{k\ell}) \vee \qquad /6/$$

$$\vee (x_{p\ell} \neq x_{k\ell} \wedge x_{q\ell} = x_{k\ell}). \qquad /7/$$

Нехай J_p, L_p – множини індексів ознак, які визначені для x_p у шкалі найменувань та шкалі порядку. Введемо позначення

$$J(p,q) = J_p \cup J_q, L(p,q) = L_p \cup L_q, \; N = card\, J(p,q) + card\, L(p,q). \qquad /8/$$

Враховуючи /I/ - /8/, розглянемо відображення $r: M \times M \to R^{+}$, де

$$r(x_p, x_q) = \frac{1}{N}\left\{ \sum_{j \in J(p,q)} r_j^{H}(x_p, x_q) + \sum_{l \in L(p,q)} r_l^{n}(x_p, x_q) \right\}. \qquad /9/$$

Т в е р д ж е н н я I. Відображення $r: M \times M \to R^{+}$ є мірою подібності для M, $0 \leqslant r(x_p, x_q) \leqslant 1$.

Відповідно до [I] міра подібності r на M є функцією належності деякого бінарного нечіткого відношення S на множині M, і визначає його. Тобто $r = \mu_S: M \times M \to [0, I]$.

Легко бачити, що із означення міри подібності r на M, яка є функцією належності μ_S для S, випливає

Т в е р д ж е н н я 2. Нечітке відношення S на M є нечітким відношенням подібності на M.

Нечітке відношення S має занадто загальні властивості й практично не враховує зв"язків між рівнем відношення подібності для різних пар множини M. Умови, які враховують цей зв"язок, називаються умовами транзитивності. Вони забезпечують можливість розбиття множини M на класи подібності, що не перетинаються. Тому розглянемо перетворення початкового нетранзитивного відношення у транзитивне. Відповідно до [3] введемо наступні означення.

О з н а ч е н н я I. Композицією нечітких відношень R і S на M називається нечітке відношення $R \circ S$ із функцією належності

$$\mu_{R \circ S}(x, z) = \max\{\min(\mu_R(x, y) \mu_S(y, z))\}, \quad \forall x, y, z \in M, \qquad /I0/$$

де μ_R і μ_S - функції належності нечітких відношень R та S.

О з н а ч е н н я 2. Нечітке відношення S є транзитивним на M, якщо $S \circ S \leqslant S$, або

$$\mu_S(x, z) \geqslant \min\{\mu_S(x, y), \mu_S(y, z)\}, \qquad \forall x, y, z \in M. \qquad /II/$$

О з н а ч е н н я 3. Транзитивним замиканням нечіткого відношення S на M називається нечітке відношення $\hat{S}$, яке визначається наступним чином: $\hat{S} = S^1 \cup S^2 \cup \ldots \cup S^k \cup \ldots$ причому відношення S^k, в свою чергу, визначається рекурсивно

$$S^1 = S, \; S^k = S^{k-1} \circ S, \quad k = 2, 3, \ldots .$$

У праці [I] показано, що транзитивне замикання $\hat{S}$ довільного нечіткого відношення S транзитивне і є узагальненням звичайного відношення еквівалентності. Якщо множина M містить m елементів і нечітке відношення S рефлексивне, то

$$\hat{S} = S^{m-1} . \qquad /12/$$

Означення 4. α - рівнем нечіткого відношення S називається звичайне відношення S, яке визначається для всіх $\alpha > 0$ наступним чином:

$$S_\alpha = \left\{ (x_p, x_q) \in M \times M : \mu_S (x_p, x_q) \geq \alpha \right\} . \qquad /13/$$

Оскільки множина M скінченна, то нечітке відношення S на цій множині можна представити за допомогою матриці розміром $m \times m$, на перетині рядка p і стовбця q якої розміщується елемент $r(x_p, x_q)$.

Визначимо композицію нечітких відношень через їх матриці. Нехай нечітким відношенням R і S відповідають матриці A і B, а $R \circ S - C$. Тоді матриці відношень $R \circ S$, R і S відповідно до введеної операції композиції /10/ пов"язані наступним чином:

$$c_{ij} = max \left\{ min \, (a_{ik} \, . \, b_{kj}) \right\} , \qquad /14/$$

де c_{ij}, a_{ik} і b_{kj} відповідно елементи матриць C, A і B.

Припустимо, що нечіткому відношенню S відповідає матриця D. Тоді, враховуючи /12/, матриця нечіткого транзитивного відношення $\hat{S}$, згідно з /14/, дорівнює D^{m-1}.

Нехай

$$N_1 = \{ (1,2), (1,3, \ldots, (1,n) \} , \quad N_2 = \{ (2,3), (2,4), \ldots, (2,n) \} ; \ldots , \quad N_{n-1} = \{ (n-1,n) \} , \quad N = N_1 \cup N_2 \cup \cdots \cup N_{n-1} .$$

Згідно з алгоритмом знаходження максимального зваженого паросполучення, викладеного у [4], розіб"ємо множину початкових даних M на такі пари елементів (x_p, x_q), щоб

$$\sum_{(p,q) \in N} \mu_S (x_p, x_q) \longrightarrow max .$$

Тим самим ми отримали деяке бінарне відношення

$$B - \left\{ (x_p, x_q) \in M \times M : \sum_{(p,q) \in N} \mu_S (x_p, x_q) \longrightarrow max \right\} .$$

Побудуємо множини $A_1' = pr_1 B$ та $A_2' = pr_2 B$, які є відповідно першою та другою проекціями відношення B. Із означення A_1' та A_2' маємо, що $M = A_1' \cup A_2'$. У випадку, коли кількість елементів M непарна, елемент x_6, який лишається, утворює пару (x_6, x_6).

Як відомо з /I/, α – рівень нечіткого транзитивного відношення $\hat{S} - \hat{S}_\alpha$ є алгоритмом класифікації. Причому до одного класу входять елементи, подібність між якими дорівнює або є більшою ніж α. Тому, застосовуючи $\hat{S}_\alpha$ до A_1' та A_2, отримаємо деяке їх розбиття на класи

$$\hat{S}_\alpha(A_1') = \{Q_{\alpha_1}^{A_1'}, Q_{\alpha_2}^{A_1'}, \ldots, Q_{\alpha_r}^{A_1'}\}, \quad \hat{S}_\alpha(A_2') = \{Q_{\alpha_1}^{A_2'}, Q_{\alpha_2}^{A_2'}, \ldots, Q_{\alpha_s}^{A_2'}\},$$

причому

$$A_1' = \bigcup_{k=1}^{r} Q_{\alpha_k}^{A_1'}, \quad Q_{\alpha_l}^{A_1'} \cap Q_{\alpha_m}^{A_1'} = \varnothing, \quad \forall l \neq m,$$

$$\mu_{\hat{S}_\alpha}(x_i', x_i'') > \max_y \{\mu_{\hat{S}_\alpha}(x_i', y), \mu_{\hat{S}_\alpha}(x_i'', y)\}, \forall y \in A_1' \setminus Q_{\alpha_1}^{A_1'},$$

$$\forall x_i', x_i'' \in Q_{\alpha_i}^{A_1'}, \quad i = 1, 2, \ldots, r,$$

та

$$A_2' = \bigcup_{k=1}^{s} Q_{\alpha_k}^{A_2'}, \quad Q_{\alpha_m}^{A_2'} \cap Q_{\alpha_2}^{A_2'} = \varnothing, \quad \forall l \neq m,$$

$$\mu_{\hat{S}_\alpha}(x_t', x_t'') > \max_y \{\mu_{\hat{S}_\alpha}(x_t', y), \mu_{\hat{S}_\alpha}(x_t', y)\},$$

$$\forall y \in A_2' \setminus Q_{\alpha_t}^{A_2'}, \quad \forall x_t', x_t'' \in Q_{\alpha_t}^{A_2'}, \quad t = 1 - s.$$

Надаючи α ($\alpha \in [0,1]$) різних значень, отримаємо множину класифікацій A_1'. Аналогічно отримуємо A_2'.

Нехай $T(A_1') = \{\hat{S}_\alpha(A_1')\}_{\alpha \in [0,1]}$ та $T(A_2') = \{\hat{S}_\alpha(A_2')\}_{\alpha \in [0,1]}$. Розглянемо відображення $P: T(A_1') \times T(A_2') \to R^+$, де

$$P(\hat{S}_\alpha(A_1'), \hat{S}_\alpha(A_2')) = 4[\Delta Q / (r + K)]^2, \qquad /15/$$

$\hat{S}_\alpha(A_1') \in T(A_1')$, $\hat{S}_\alpha(A_2') \in T(A_2')$, r, K є кількість класів, на які алгоритм класифікації $\hat{S}_\alpha$ розбиває множини A_1' та A_2'.

ΔQ - кількість подібних класів у класифікаціях $\hat{S}_\alpha(A'_1)$ та $\hat{S}_\alpha(A'_2)$.

Означення 5. Клас $Q_{\alpha i}^{A'_1}$ називається подібним до класу $Q_{\alpha j}^{A'_2}$, якщо між їх елементами можна встановити таке бієктивне відображення, що для будь-якого $x_i^{r_1} \in Q_{\alpha i}^{A'_1}$ існує $x_j^{q_2} \in Q_{\alpha j}^{A'_2}$ такий, що $(x_i^{r_1}, x_j^{q_2}) \in B$ і навпаки, для будь-якого $x_j^{q_2} \in Q_{\alpha j}^{A'_2}$ існує $x_i^{r_1} \in Q_{\alpha i}^{A'_1}$ такий, що $(x_i^{r_1}, x_j^{q_2}) \in B$.

Твердження 3. Відображення $P: T(A'_1) \times T(A'_2) \to R^+$ є мірою подібності між елементами множин $T(A'_1)$ та $T(A'_2)$ і $0 \leq P \leq 1$.

Із означення P випливає, що $P(\hat{S}_\alpha(A'_1), \hat{S}_\alpha(A'_2)) = P(\hat{S}_\alpha(A'_2), \hat{S}_\alpha(A'_1))$. Враховуючи, що $\Delta Q \leq r$ і $\Delta Q \leq K$, отримаємо $0 \leq P \leq 1$.

2. Розв"язування задачі класифікації

Нехай M - множина початкових даних, які описані за допомогою шкали порядку та шкали найменувань. Розглянемо задачу класифікації, що полягає у знаходженні оптимального алгоритму класифікації $\hat{S}_{опт}$, який задовольняє такі умови. По-перше, розбиває множину M на класи $Q_1, Q_2, \dots, Q_t$, що не перетинаються / t - наперед невідоме, а

$$M = \bigcup_{i=1}^{t} Q_i \;, \quad Q_i \cap Q_j = \varnothing \;, \qquad \forall i \neq j),$$

причому міра подібності між елементами одного класу має бути більшою порівняно з мірою їх подібності до елементів інших класів. По-друге, алгоритм класифікації надає мірі подібності /I5/ максимуму.

Розв"язування задачі проведемо за методом групового визначення аргументів /МГВА/([2]), який передбачає створення множини алгоритмів класифікації та організацію їх перегляду, внаслідок чого отримується оптимальний алгоритм. Критерієм оптимальності виступає друга умова задачі, яка вимагає, щоби на подібних множинах оптимальний алгоритм здійснював близькі класифікації.

Згідно п. I нашої праці побудуємо на M нечітке транзитивне відношення подібності $\hat{S}$, його функцію належності μ_S та його α рівні. Нехай нечіткому віношенню $\hat{S}$ відповідає матриця

F з елементами f_{ij}, причому $0 \leq f_{ij} \leq 1, 1 \leq i, j \leq m$. Впорядкуємо різні елементи F за зростанням

$$q_1 < q_2 < \dots < q_r , \qquad /I6/$$

де

$$q_1 = \min_{ij} f_{ij} , \qquad q_2 = \max_{ij} f_{ij} = 1 .$$

Надаючи послідовно α всі значення із /I6/, отримаємо множину алгоритмів класифікації $\{\hat{S}_\alpha\}$ множини M. До $\{\hat{S}_\alpha\}$ не будемо включати $\hat{S}_{q_1}$ та $\hat{S}_{q_r}$, оскільки вони проводять тривіальну класифікацію, яка не дає жодної нової інформації відносно структури множини M.

Як показано в п. I, кожний $\hat{S}_\alpha$ розбиває M на класи, які задовольняють першу умову, причому міра подібності між елементами кожного класу дорівнює або є більшою від α. Для розв"язку задачі залишилось виділити з множини $\{\hat{S}_\alpha\}$ такий оптимальний алгоритм, який надає максимуму мірі подібності /I5/.

З [2] випливає, що критерії вигляду /I5/ виконуються, як правило, на декількох класифікаціях, серед яких є й оптимальна. В такому випадку розв"язок задачі є неоднозначним, і ми маємо некоректну задачу. Позначимо через E множину тих значень із послідовності /I6/, для яких класифікації $\hat{S}_\alpha$ надають максимуму мірі подібності /I5/. Для виділення єдиного і стійкого розв"язку скористаємося згідно з [5], регуляризацією. Для цього додамо до міри подібності /I5/ деякий стабілітрон, що надає можливість створити згладжуючий критерій.

Для побудови стабілітрона проведемо обмін між A_1^1 та A_2^1 елементами декількох пар. В результаті отримаємо нові підмножини A_1^2 та A_2^2. Застосовуючи аналогічну процедуру з елементами інших пар, отримаємо нові підмножини A_1^3 та A_2^3 ... A_1^ℓ та A_2^ℓ $(M = A_1^\ell \cup A_2^\ell$. Введемо відображення

$$P = 1/(\ell - 1) \sum_{k=2}^{\ell} P(\hat{S}_\alpha(A_1^k), \hat{S}_\alpha(A_2^k)), \qquad /I7/$$

яке є мірою подібності та яке будемо називати стабілітроном. Тоді згладжуючий критерій матиме вигляд

$$R = P(\hat{S}_\alpha(A_1^1), \hat{S}_\alpha(A_2^1)) + P \to max . \qquad /I8/$$

У [2] вказано, що критерії вигляду /I8/ досягають свого максимуму тільки на одній класифікації. В /I7/ практично треба взяти лише декілька доданків.

Використовуючи таку методику, внаслідок перебору множини алгоритмів $\{\hat{S}_\alpha\}_{\alpha \in E}$ можемо знайти такий алгоритм $\hat{S}_{\alpha\, opt}$, який задовольняє /I8/. Знайдена класифікація й буде розв"язком нашої задачі.

Запропонований підхід зі застосуванням методів теорії нечітких множин при використанні МГВА для розв"язку задачі класифікації дає змогу зменшити перебір з 2^m ($m = card\, M$) до величини, що не перебільшує $m(m-3)/2$ і суттєво зменшити обсяг обчислень.

Даний метод класифікації рекомендується застосовувати для розбиття множини початкових даних M, що описані за допомогою шкали найменувань і шкали порядку на подібні класи з невизначеною наперед їх кількістю у випадку, коли $n > m$.

I. Аверкин А.Н., Батыршин И.З., Блишун А.Ф. и др. Нечеткие множества в моделях управления и искусственного интеллекта. М., I986. 2. Ивахненко А.Г. Метод последовательного опробования /перебора/ кластеризаций-кандидатов по критериям дифференциального типа //Распознавание, классификация, прогноз. I989. Вып.2. 3. Кофман А. Введение в теорию нечетких множеств. М., I982. 4. Препарата Ф., Шеймос М. Вычислительная геометрия: Введение. М., I989. 5. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М., I986.

Стаття надійшла до редколегії 30.04.91