УДК 004:37

О.Меняйленко, канд. техн. наук
Луганський національний педагогічний університет імені Тараса Шевченка

СУБОПТИМАЛЬНІ АДАПТИВНІ АЛГОРИТМИ ДИФЕРЕНЦІЮВАННЯ УЧНІВ НА РІВНІ

У роботі розглядаються модифіковані алгоритми диференціювання учнів на рівні (класи) залежно від первинних параметрів, що утворюють образ учня. Наводяться результати їх чисельного дослідження.

Умовні позначення

$\Omega_j(\omega_1, \ldots, \omega_i, \ldots, \omega_n)$ – первинні ознаки учня j;
$O_j(\omega_1, \ldots, \omega_l)$ – образ учня j;
$P(\omega_l)$ – суміш розподілів первинних ознак учня, що утворюють його образ;
k – рівень учня;
m – кількість рівнів;
P_k – імовірність учня з рівнем k;
$p_k\left(\omega_l / \Omega_i, \sigma_k^2\right)$ – розподіл учнів усередині рівня k;
σ_k^2 – дисперсія всередині рівня k;
$\omega_l[n]$ – елементи суміші;
$M_i[n]$ – центр класу i;
$F_i(\omega_l[n], M_i[n-1])$ – функція втрат;
$f_{im}(\omega_l, M_i, M_m)$ – відокремлювальна функція;
Ω_i, Ω_m – множини ситуацій $\omega_l[n]$ класів i й m із центрами $M_i[n-1]$ та $M_m[n-1]$;
n – дискретні моменти часу 0, 1, …, n;
$\sigma_i^2[n]$, $\sigma_m^2[n]$ – дисперсії елементів $\omega_l[n]$ суміші в підмножинах Ω_i та Ω_m;
ρ – елементів суміші ω_l від “центрів” класів M;
$\rho_k(\omega_l, M_k)$ – функція втрат (штрафів) для k-го класу;
R – середній ризик (середній штраф) для всіх класів m;
$F_i(\cdot)$, $F_k(\cdot)$ – результуючі функції втрат для i й k класів у суміші;
$\sigma_i^2[n-1]$, $\sigma_k^2[n-1]$ – дисперсії i й k класів;
$F_i^*(\cdot)$, $F_k^*(\cdot)$ – оптимальні функції втрат для i й k класів;
$D_{ki}[n]$ – параметри нелінійної корекції відокремлювальної функції;
n – порядковий номер учня з параметрами первинної ознаки $\omega_i[n]$ $(1 \le n \le N_o)$;
N_o – загальна кількість учнів;
N^1, N^2, N^3 – кількість учнів, для яких $\omega_i[n] \in k_1$, $\omega_i[n] \in k_2$, $\omega_i[n] \in k_3$;
t_n^a – час навчання (адаптації);
$\overline{M}_1$, $\overline{M}_2$, $\overline{M}_3$ – сталі значення оцінок центрів класів у суміші;
$N_{er\%}$ – кількість помилок класифікації параметрів суміші розподілів.

Одним з напрямів створення та використання сучасних інформаційних технологій у педагогіці є індивідуалізація процесу навчання учнів на основі диференціювання їх на рівні (класи, підгрупи) (рис. 1) [1–5]. При побудові інформаційних технологій навчання треба формалізувати процес поділу учнів на рівні залежно від їхніх когнітивних психолого-фізіологічних і педагогічних особливостей. У роботі [6] уведено поняття первинних ознак $\Omega_j(\omega_1, \ldots, \omega_i, \ldots, \omega_n)$ учня j, його образу $O_j(\omega_1, \ldots, \omega_l)$, $l \le n$, та рівня k, а в [7] показано, що параметри, які утворюють образ учня, являють собою гетероскедастичні суміші розподілів неперетинних, слабко і сильно перетинних класів з їх кількістю від 1 до 3.

Тоді математичну модель учнів, що відображає їхні індивідуальні особливості, можна записати у вигляді такої суміші розподілів:

$$P(\omega_l)=\sum_{k=1}^{m}P_k p_k\left(\omega_l/\Omega_i,\sigma_k^2\right);\ \sum_{k=1}^{m}P_k=1\,;P_k\geq 0\,,\ (1)$$

де $P(\omega_l)$ – суміш розподілів первинних ознак учня, що утворюють його образ; P_k – імовірність учня з рівнем k; m – кількість рівнів; $p_k\left(\omega_l/\Omega_i,\sigma_k^2\right)$ – розподіл учнів усередині рівня k.

Суміші розподілів досліджуваних психологічних параметрів у (1) мають різну дисперсію σ_k^2 усередині кожного рівня k, тому на сьогодні не розроблено формалізованих алгоритмів класифікації (диференціювання) учнів на класи (рівні, підгрупи), що не дозволяє будувати ефективні автоматизовані навчальні системи, які враховують педагогічні вимоги до процесу навчання [5, 7].

Формулювання і вирішення задачі диференціювання учнів на класи (рівні)

У даній роботі розглядається задача розробки субоптимальних алгоритмів класифікації (диференціювання) учнів на класи (рівні), у яких використовується адаптивний підхід до побудови самонавчальних систем класифікації. Як апріорна інформація береться кількість класів m, з яких утворено результуючу суміш розподілів первинних параметрів, що становлять образ учня O_j. Неповнота апріорної інформації про інші параметри суміші розподілів долається навчанням.

Якість поділу (класифікації, диференціювання) учнів на класи (рівні, підгрупи) визначається вибором функції втрат (відокремлювальної функції) [5, 8, 9]. Аналіз публікацій показує, що використовуються різні підходи до вибору конкретних функцій утрат: від рекомендацій вибирати найпростішу квадратичну функцію втрат, до тверджень, що вибір функції втрат є спірним і нерозв’язним [5–9]. У роботах [8, 9] дано розв’язання зазначеної задачі в загальному вигляді, показано, що залежно від рівня апріорної інформації про заваду вибирають оптимальну функцію втрат, а також отримано оптимальні функції втрат для типових щільностей розподілу завади. Так, для нормальної (гауссової) функції розподілу завад оптимальна функція втрат квадратична, а для експоненціальної (лапласової) – оптимальна функція втрат модульна.

В адаптивних алгоритмах класифікації можна виділити два основних типи функцій утрат: 1) функції втрат, що залежать тільки від відстані елементів суміші $\omega_l[n]$ від своїх “центрів” класів $M_i[n-1]$; 2) функції втрат, що залежать від параметрів інших класів суміші. Як показано в роботі [9], оптимальні функції втрат першого типу, як правило, є парними, але тоді є правильним таке твердження.

<u>Твердження 1</u>. Якщо функція втрат $F_i(\omega_l[n],M_i[n-1])$ є парною і залежить тільки від відстані між елементами суміші $\omega_l[n]$ і центрами класів $M_i[n-1]$, то відокремлювальна функція $f_{im}(\omega_l,M_i,M_m)$ не залежить від конкретного вигляду функції втрат $F_i(x[n],u_i[n-1])$, а розв’язувальне правило має такий вигляд:

$$\omega_l[n]\in\Omega_i \text{ при } \omega_l[n]\leq\frac{M_i[n-1]+M_m[n-1]}{2};$$

$$\omega_l[n]\in\Omega_m \text{ при } \omega_l[n]>\frac{M_i[n-1]+M_m[n-1]}{2}, \qquad (2)$$

де Ω_i,Ω_m – множина ситуацій $\omega_l[n]$ класів i й m із центрами $M_i[n-1]$ та $M_m[n-1]$; n – дискретні моменти часу 0, 1, …, n.

Доказ твердження 1. Оскільки для будь-якої парної функції утрат $F_i(\omega_l[n], M_i[n-1])$ виконується така рівність (за визначенням [9]):

$$F_i(\omega_l[n], M_i[n-1]) = F_i(-\omega_l[n], M_i[n-1]), \tag{3}$$

а, виходячи із залежності функції втрат тільки від відстані $(\omega_l[n] - M_i[n-1])$, також виконується рівність

$$F_i(\omega_l[n] - M_i[n-1]) = F_m(\omega_l[n] - M_i[n-1]). \tag{4}$$

Рівняння відокремлювальної лінії, відповідно до [8], знайдемо з виразу

$$f_{im}(\omega_l, M_i, M_m) = F_i(\omega_l[n] - M_i[n-1]) - F_m(\omega_l[n] - M_m[n-1]) = 0. \tag{5}$$

Тоді

$$F_i(\omega_l[n] - M_i[n-1]) = F_m(\omega_l[n] - M_m[n-1]), \tag{6}$$

але в цьому разі спостережувана ситуація $\omega_l[n]$ знаходитиметься між центрами класів $M_i[n-1]$ та $M_m[n-1]$, тобто

$$M_i[n-1] < \omega_l[n] < M_m[n-1], \tag{7}$$

а це дозволяє переписати рівність (6) з урахуванням умов (3) і (4) у такому вигляді:

$$F_i(\omega_l[n] - M_i[n-1]) = F_i(-(\omega_l[n] - M_m[n-1])). \tag{8}$$

Тоді можна записати також рівність аргументів, тобто

$$\omega_l[n] - M_i[n-1] = -(\omega_l[n] - M_m[n-1]), \tag{9}$$

звідки відокремлювальна лінія визначається рівністю

$$\omega_l[n] = \frac{M_i[n-1] + M_m[n-1]}{2}. \tag{10}$$

Але ця відокремлювальна лінія відповідає розв'язувальному правилу (2), отже, твердження 1 доведено.

Загальною вадою відокремлювальних функцій, що реалізують розв'язувальне правило (2), є зсув оцінок центрів $M_i[n-1]$ класів навіть у випадку неперетинних сумішей. Доведемо це в такому твердженні.

Твердження 2. Якщо відокремлювальна функція $f_{im}(\omega_l, M_i, M_m)$ приводить до розв'язувального правила (2), а дисперсії $\sigma_i^2[n-1]$ і $\sigma_m^2[n-1]$ елементів $\omega_l[n]$ суміші в підмножинах Ω_i та Ω_m різні (гетероскедастичні суміші) і класи є неперетинними за визначенням [6, 7], то розв'язувальне правило (2) даватиме незсунуті оцінки центрів класів $M_i[n-1]$ і $M_m[n-1]$ тільки у випадку:

$$3\sigma_i[n-1] \le \frac{|M_m[n-1] - M_i[n-1]|}{2} \ge 3\sigma_m[n-1]. \tag{11}$$

Доказ твердження 2. Очевидно, що спостережувані ситуації $\omega_l[n]$ в підмножинах Ω_i й Ω_m класів i та m не повинні переходити через відокремлювальну лінію, визначувану виразом (10). Ця умова й визначає нерівність (11). В іншому разі частина спостережуваних ситуацій $\omega_l[n]$ одного з класів i або m розв'язувальне правило (2) відноситиме до другого класу, що й приводитиме до зсунутих оцінок

$M_i[n-1]$ або $M_m[n-1]$. Покажемо, коли можливе порушення нерівності (11). Нехай класи *i* й *m* не перетинаються, а

$$3\sigma_i[n-1] > \frac{|M_m[n-1] - M_i[n-1]|}{2} \tag{12}$$

або

$$3\sigma_m[n-1] > \frac{|M_m[n-1] - M_i[n-1]|}{2}. \tag{13}$$

Тоді розв'язувальне правило (2) при виконанні нерівності (12) відноситиме частину значень $\omega_l[n] \in \Omega_i$, що виходять за відокремлювальну лінію (10), до підмножини Ω_m, а при виконанні нерівності (13) частину значень $\omega_l[n] \in \Omega_m$, що також виходять за відокремлювальну лінію (10), – до підмножини Ω_i. Таким чином, розв'язувальне правило (2) навіть у випадках неперетинних класів і при різних дисперсіях елементів $\omega_l[n]$ у класах (гетероскедастичних сумішах) даватиме зсунуті оцінки центрів класів $M_i[n-1]$ і $M_m[n-1]$ у суміші. Отже, твердження 2 доведено.

У роботі [8] для усунення зазначеної хиби пропонується вводити у функцію втрат "зсув" на величину квадрата протилежного класу (алгоритм Дорофеюка). У такій системі класифікації відокремлювальна лінія завжди зміщається до центру меншого класу (наприклад, при $m > i$ – до $M_i[n-1]$), що приводитиме до зсунутих оцінок центрів класів. Це викликано тим, що в суміші розподілів завжди має виконуватися нерівність (при $m > i$):

$$M_m[n-1] > M_i[n-1], \tag{14}$$

з якої випливає, що чим більше $M_m[n-1]$, тим більше відокремлювальна лінія зміщатиметься до центру $M_i[n-1]$. Отже, така корекція також не усуває зсуву оцінок центрів класів.

При розробці алгоритмів класифікації (диференціювання) учнів на класи (рівні, підгрупи) застосуємо адаптивний підхід до побудови самонавчальних систем класифікації, розвинутий у роботах [8, 9, 12], а також у роботах автора [5–7, 10, 11]. Як апріорна інформація в зазначених алгоритмах (підходах) використовується тільки кількість класів *m*, з яких утворено результуючу суміш (1), і спостережувані ситуації – значення первинних ознак учнів, що становлять образ учня O_j. Неповнота апріорної інформації P_k, $p_k(\omega_l/\Omega_j, \sigma_k^2)$ в (1) долається навчанням.

З математичної точки зору будемо припускати, що суміш (1) є скінченною розрізнюваною сумішшю. Таке обмеження пов'язане з тим, що за скінченною кількістю *n* спостережень не можна визначити нескінченну кількість компонентів у суміші розподілів. Також будемо припускати, що існують щільності $p_k\left(\omega_l/\Omega_i, \sigma_k^2\right)$, які є відомими функціями своїх аргументів $\omega_l, \Omega_i, \sigma_k^2$. Спостерігаючи реалізацію $\omega_l[n]$ в дискретні моменти часу $n = 0, 1, \ldots,$ суміші (1), необхідно визначити приналежність учня j до класу k.

Додержуючись [8, 9, 11], уведемо відстань ρ елементів суміші ω_l від "центрів" класів M, задавши опуклу функцію F_k різниці $\rho_k(\omega_l, M_k) = F_k(\omega_l - M_k)$. Функцію $\rho_k(\omega_l, M_k)$ можна розглядати як функцію втрат (штрафів) для *k*-го класу. Тоді середній ризик (середній штраф) для всіх класів m можна подати виразом:

$$R = \sum_{k=1}^{m} P_k \int_{\omega k} F_k(\omega_l - M_k) p_j(\omega_l / \Omega_j, \sigma_k^2) dx, \tag{15}$$

або, підставивши сумісну щільність розподілу ймовірності суміші (1), середній ризик (15) можна записати:

$$R = \sum_{k=1}^{m} \int_{\omega k} F_k(\omega_l - M_k) P(\omega_l) d\omega_l. \tag{16}$$

При цьому припускається, що класи в суміші розподілів не перетинаються. Умова мінімуму середнього ризику має такий вигляд [7–9, 11]:

$$\int_{\omega k} \nabla_{Mk} F_k(\omega_l - M_k) P(\omega_k) d\omega_l; \tag{17}$$

$$f(\omega_l, M_k, M_i) = F_k(\omega_l - M_k) - F_i(\omega_l - M_i) = 0. \tag{18}$$

Умова (17) визначає значення центрів класів M_k в суміші (1), а (18) – лінію (поверхню), що розділяє області Ω_i та Ω_k на класи k й i. Як випливає з (15)–(18), якість розподілу суміші (1) на класи визначатиметься видом функції втрат $F_k(\omega_l - M_k)$. Розглянемо функції втрат такого вигляду:

$$F_i(\omega_l[n], \sigma_k[n-1], M_i[n-1]) = F_i^*(\omega_l[n], M_i[n-1]) + \sigma_k^2[n-1]; \tag{19}$$

$$F_k(\omega_l[n], \sigma_i[n-1], M_k[n-1]) = F_k^*(\omega_l[n], M_k[n-1]) + \sigma_i^2[n-1], \tag{20}$$

де $F_i(\cdot)$, $F_k(\cdot)$ – результуючі функції втрат для *i* й *k* класів у суміші; $\sigma_i^2[n-1]$, $\sigma_k^2[n-1]$ – дисперсії *i* й *k* класів; $F_i^*(\cdot)$, $F_k^*(\cdot)$ – оптимальні функції втрат для *i* й *k* класів, що залежать тільки від відстані між елементами суміші $\omega_l[n]$ та центрами класів $M_i[n-1], M_k[n-1]$.

Тоді відокремлювальна функція матиме такий вигляд:

$$\begin{aligned} & f_{ik}(\omega_l, M_i, M_k, \sigma_i, \sigma_k) = \\ & = F_i^*(\omega_l[n], M_i[n-1]) - F_k^*(\omega_l[n], M_k[n-1]) + \sigma_k^2[n-1] - \sigma_i^2[n-1] = 0. \end{aligned} \tag{21}$$

Введемо таку нелінійну корекцію відокремлювальної функції (21):

$$\operatorname{sgn}(\sigma_k[n], M_k[n], M_i[n]) = \begin{cases} 1, \text{ якщо } 3\sigma_k[n] \geq 1/2(|M_k[n] - M_i[n]|); \\ 0, \text{ якщо } 3\sigma_k[n] < 1/2(|M_k[n] - M_i[n]|); \end{cases} \tag{22}$$

$$D_{ki}[n] = \sigma_k^2[n] \operatorname{sgn}(\sigma_k[n], M_k[n], M_i[n]). \tag{23}$$

Підставивши замість $\sigma_k^2[n]$ у виразі (21) значення $D_{ki}[n]$ (23) і випустивши проміжні перетворення, дістанемо таке розв'язувальне правило:

$$\omega_l[n] \in \Omega_i \text{ при } \omega_l[n] \leq \frac{M_k[n-1] + M_i[n-1] + D_{ik}[n-1] - D_{ki}[n-1]}{2};$$

$$\omega_l[n] \in \Omega_k \text{ при } \omega_l[n] > \frac{M_k[n-1] + M_i[n-1] + D_{ik}[n-1] - D_{ki}[n-1]}{2}. \tag{24}$$

Запропонована корекція відокремлювальної функції (24) дозволяє врахувати гетероскедастичні особливості розглянутих сумішей розподілів (1). Наприклад, при виконанні нерівності (11) $D_{ki}[n] = 0$ і $D_{ik}[n] = 0$ (див. (22)) отримуємо розв'язувальне правило (2), що відповідає неперетинним класам [6, 11]. При виконанні нерівностей (12) або (13) здійснюється корекція відокремлювальної функції на величину $D_{ki}[n]$ або $D_{ik}[n]$. Для випадків слабо і сильно перетинних класів у суміші [6, 11] одержуємо розв'язувальне правило (24). При цьому очевидно, що алгоритм [13, 14] є окремим випадком розв'язувального правила (24).

Оцінку параметрів $M_i[n], M_k[n]$, $\sigma_i^2[n], \sigma_k^2[n]$ доцільно отримати на основі методів стохастичної апроксимації [8, 9, 12], оскільки в цьому разі одночасно здійснюється оцінка зазначених параметрів за спостережуваними значеннями, а також обчислюється розв'язувальне правило (24).

На основі викладеного вище підходу до задачі класифікації (диференціювання) учнів на класи (рівні, підгрупи) при кількості класів $m = 3$ адаптивні алгоритми класифікації гетероскедастичної суміші можна записати в такому вигляді [5, 11]:

$$n \in k_1;$$

$$M_1[n] = M_1[n-1] - \gamma_1[n](M_1[n-1] - \omega_i[n]);$$

$$\sigma_1^2[n] = \sigma_1^2[n-1] - \gamma_1[n]\left[\sigma_1^2[n-1] - (\omega_i[n] - M_1[n-1])^2\right];$$

$$M_2[n] = M_2[n-1];$$

$$\sigma_2^2[n] = \sigma_2^2[n-1];$$

$$M_3[n] = M_3[n-1];$$

$$\sigma_3^2[n] = \sigma_3^2[n-1];$$

при

$$\left((\omega_i[n] - M_1[n-1])^2 - D_{12}[n-1] + D_{21}[n-1] - (\omega_i[n] - M_2[n-1])^2 \right) \le 0;$$

$$\left((\omega_i[n] - M_1[n-1])^2 - D_{13}[n-1] + D_{31}[n-1] - (\omega_i[n] - M_3[n-1])^2 \right) \le 0;$$

або

$$n \in k_2;$$

$$M_1[n] = M_1[n-1];$$

$$\sigma_1^2[n] = \sigma_1^2[n-1];$$

$$M_2[n] = M_2[n-1] - \gamma_2[n](M_2[n-1] - \omega_i[n]);$$

$$\sigma_2^2[n] = \sigma_2^2[n-1] - \gamma_2[n]\left[\sigma_2^2[n-1] - (\omega_i[n] - M_2[n-1])^2\right];$$

$$M_3[n] = M_3[n-1];$$

$$\sigma_3^2[n] = \sigma_3^2[n-1];$$

при

$$\left((\omega_i[n] - M_2[n-1])^2 - D_{21}[n-1] + D_{12}[n-1] - (\omega_i[n] - M_1[n-1])^2 \right) \le 0;$$

$$\left((\omega_i[n] - M_2[n-1])^2 - D_{23}[n-1] + D_{32}[n-1] - (\omega_i[n] - M_3[n-1])^2 \right) \le 0;$$

або

$$n \in k_3;$$

$$M_1[n] = M_1[n-1];$$

$$\sigma_1^2[n] = \sigma_1^2[n-1];$$

$$M_2[n] = M_2[n-1];$$

$$\sigma_2^2[n] = \sigma_2^2[n-1];$$

$$M_3[n] = M_3[n-1] - \gamma_3[n](M_3[n-1] - \omega_i[n]);$$

$$\sigma_3^2[n] = \sigma_3^2[n-1] - \gamma_3[n]\left[\sigma_3^2[n-1] - (\omega_i[n] - M_3[n-1])^2\right]$$

в іншому разі;

$$\mathrm{sgn}(\sigma_1[n], M_1[n], M_2[n]) = \begin{cases} 1, \text{ якщо } 3\sigma_1[n] \geq 1/2(|M_1[n] - M_2[n]|); \\ 0, \text{ якщо } 3\sigma_1[n] < 1/2(|M_1[n] - M_2[n]|); \end{cases}$$

$$D_{12}[n] = \sigma_1^2[n]\,\mathrm{sgn}(\sigma_1[n], M_1[n], M_2[n]);$$

$$\mathrm{sgn}(\sigma_2[n], M_1[n], M_2[n]) = \begin{cases} 1, \text{ якщо } 3\sigma_2[n] \geq 1/2(|M_1[n] - M_2[n]|); \\ 0, \text{ якщо } 3\sigma_2[n] < 1/2(|M_1[n] - M_2[n]|); \end{cases}$$

$$D_{21}[n] = \sigma_2^2[n]\,\mathrm{sgn}(\sigma_2[n], M_1[n], M_2[n]);$$

$$\mathrm{sgn}(\sigma_2[n], M_3[n], M_2[n]) = \begin{cases} 1, \text{ якщо } 3\sigma_2[n] \geq 1/2(|M_3[n] - M_2[n]|); \\ 0, \text{ якщо } 3\sigma_2[n] < 1/2(|M_3[n] - M_2[n]|); \end{cases}$$

$$D_{23}[n] = \sigma_2^2[n]\,\mathrm{sgn}(\sigma_2[n], M_3[n], M_2[n]);$$

$$\mathrm{sgn}(\sigma_3[n], M_3[n], M_2[n]) = \begin{cases} 1, \text{ якщо } 3\sigma_3[n] \geq 1/2(|M_3[n] - M_2[n]|); \\ 0, \text{ якщо } 3\sigma_3[n] < 1/2(|M_3[n] - M_2[n]|); \end{cases}$$

$$D_{32}[n] = \sigma_3^2[n]\,\mathrm{sgn}(\sigma_3[n], M_3[n], M_2[n]);$$

$$\mathrm{sgn}(\sigma_1[n], M_3[n], M_1[n]) = \begin{cases} 1, \text{ якщо } 3\sigma_1[n] \geq 1/2(|M_3[n] - M_1[n]|); \\ 0, \text{ якщо } 3\sigma_1[n] < 1/2(|M_3[n] - M_1[n]|); \end{cases}$$

$$D_{13}[n] = \sigma_1^2[n]\,\mathrm{sgn}(\sigma_1[n], M_3[n], M_1[n]);$$

$$\mathrm{sgn}(\sigma_3[n], M_3[n], M_1[n]) = \begin{cases} 1, \text{ якщо } 3\sigma_3[n] \geq 1/2(|M_3[n] - M_1[n]|); \\ 0, \text{ якщо } 3\sigma_3[n] < 1/2(|M_3[n] - M_1[n]|); \end{cases}$$

$$D_{31}[n] = \sigma_3^2[n]\,\mathrm{sgn}(\sigma_3[n], M_3[n], M_1[n]);$$

$$\gamma_1[n] = \frac{1}{N^1};\ \gamma_2[n] = \frac{1}{N^2};\ \gamma_3[n] = \frac{1}{N^3}, \tag{25}$$

де $M_1[n]$, $M_2[n]$, $M_3[n]$ – оцінки центрів класів 1, 2 і 3 у суміші розподілів первинних параметрів $\omega_i[n]$, за якими здійснюється класифікація (диференціювання, поділ) учнів; $\sigma_1^2[n]$, $\sigma_2^2[n]$, $\sigma_3^2[n]$ – оцінки дисперсій центрів класів у суміші розподілів первинних параметрів $\omega_i[n]$, що утворюють образ учня; $D_{12}[n]$, $D_{21}[n]$, $D_{23}[n]$, $D_{32}[n]$, $D_{13}[n]$, $D_{31}[n]$ – параметри нелінійної корекції відокремлювальної функції; n – порядковий номер учня з параметрами первинної ознаки $\omega_i[n]$ $(1 \leq n \leq N_o)$; N_o – загальна кількість учнів; N^1, N^2, N^3 – кількість учнів, для яких $\omega_i[n] \in k_1$, $\omega_i[n] \in k_2$, $\omega_i[n] \in k_3$; $N_o = N^1 + N^2 + N^3$.

Початкові значення в алгоритмах (25) вибирають, виходячи з таких виразів:

$$M_1[0] = \omega_i^{\min} + \left(\frac{\omega_i^{\max} - \omega_i^{\min}}{2m} \right);$$

$$M_2[0] = M_1[0] + \left(\frac{\omega_i^{\max} - \omega_i^{\min}}{m} \right); \tag{26}$$

$$M_3[0] = M_2[0] + \left(\frac{\omega_i^{\max} - \omega_i^{\min}}{m} \right);$$

$$\sigma_1^2[0] = \sigma_2^2[0] = \sigma_3^2[0] = 0;\ D_{12}[0] = D_{21}[0] = D_{23}[0] = D_{32}[0] = D_{13}[0] = D_{31}[0] = 0,$$

де $\omega_i^{\min}$, $\omega_i^{\max}$ – найменше й найбільше значення первинної ознаки $\omega_i[n]$, за якою здійснюється поділ суміші на класи; $M_1[0]$, $M_2[0]$, $M_3[0]$ – початкові значення центрів класів у суміші; $\sigma_1^2[0]$, $\sigma_2^2[0]$, $\sigma_3^2[0]$ – початкові значення для дисперсій класів у суміші; $D_{12}[0]$, $D_{21}[0]$, $D_{23}[0]$, $D_{32}[0]$, $D_{13}[0]$, $D_{31}[0]$ – початкові значення параметрів нелінійної корекції відокремлювальної функції; m – кількість класів у суміші розподілів ($m = 3$).

Якісні й кількісні оцінки роботи алгоритмів (25) стосовно до задачі класифікації учнів на рівні визначатимемо чисельним моделюванням з використанням еталонних сумішей, що мають різну міру перетинання класів [6, 11] при кількості класів $m = 3$. Це пов’язано з тим, що на сьогодні немає аналітичних методів дослідження зазначеного типу алгоритмів.

Час навчання (адаптації) t_n^a визначатимемо за моментом перетинання всіма оцінками центрів класів $M_1[n]$, $M_2[n]$ і $M_3[n]$ 5 % ліній від своїх сталих значень:

$$M_1[n] \le (1 \pm 0{,}05)\overline{M}_1[n];$$

$$M_2[n] \le (1 \pm 0{,}05)\overline{M}_2[n]; \tag{27}$$

$$M_3[n] \le (1 \pm 0{,}05)\overline{M}_3[n],$$

де $\overline{M}_1$, $\overline{M}_2$, $\overline{M}_3$ – сталі значення оцінок центрів класів у суміші.

Оцінки точності класифікації параметрів еталонних сумішей алгоритмами (25) визначимо через надійні інтервали для центрів класів $M_1[n]$, $M_2[n]$, $M_3[n]$, а також через кількість помилок класифікації параметрів суміші розподілів $N_{er\%}$, виражену у процентах до загальної кількості спостережуваних значень N_o. Якщо в надійні інтервали потрапить істинне (теоретичне) значення оцінюваного параметра, то оцінку, що дається алгоритмами (25), вважатимемо незсунутою. Надійні інтервали обчислюватимемо на рівні значимості 95 %.

Методика та результати чисельного дослідження алгоритмів класифікації учнів на класи (рівні)

Методика чисельного дослідження алгоритмів класифікації учнів на класи (рівні) (25) включає такі етапи: 1) генерування гетероскедастичних еталонних сумішей розподілів для неперетинних, слабко й сильно перетинних класів у суміші; 2) чисельне моделювання алгоритмів (25) для різних мір перетинання класів у суміші [6, 11]; 3) побудова графіків залежностей $M_1[n]$, $M_2[n]$, $M_3[n]$, $\sigma_1^2[n]$, $\sigma_2^2[n]$, $\sigma_3^2[n]$, ідентифікація класу спостережуваних значень $k[n]$, кількості помилок ідентифікації класів у суміші $N_{er\%}[n]$; 4) визначення часу навчання (адаптації) t_n^a, а також точності оцінки досліджуваних параметрів через надійні інтервали; 5) виявлення можливості класифікації учнів на класи (рівні) алгоритмами (25).

Для генерації еталонних гетероскедастичних сумішей використаємо мову MML, вбудовану в програму *Statistica 5.0*. Відповідно до алгоритмів (25) було розроблено прикладну програму "Ідентифікація суміші розподілів" на мові *Object pascal*. На рис. 2 показано типові перехідні процеси в алгоритмах (25) класифікації (поділу) суміші на класи (рівні), а в табл. 1 подано надійні інтервали для оцінок $M_1[n]$, $M_2[n]$ і $M_3[n]$, помилки класифікації $N_{er\%}$, а також час навчання (адаптації) t_n^a.

З наведених результатів видно, що оцінки центрів класів $M_1[n]$, $M_2[n]$, $M_3[n]$, $\sigma_1^2[n]$, $\sigma_2^2[n]$ та $\sigma_3^2[n]$ наближаються до своїх теоретичних значень. Оцінки $M_1[n]$, $M_2[n]$ та $M_3[n]$ для неперетинних і слабко перетинних класів потрапляють у надійні інтервали для значень еталонних гетероскедастичних сумішей M_{e1}, M_{e2} і M_{e3} (див. табл. 1), тобто отримані оцінки центрів класів є незсунутими. При збільшенні міри перетинання класів збільшується й величина зсуву. Так, для випадку сильно перетинних класів спостерігається відхил центрів класів $M_1[n]$ і $M_2[n]$ від припустимих значень надійних інтервалів, що не перевищує 5,4 % і 6,35 % від меж відповідних надійних інтервалів. Для виродженого випадку сильно перетинних класів з однаковими математичними сподіваннями й різними дисперсіями алгоритми (25) дають зсунуті оцінки для параметрів суміші (див. табл. 1).

Таблиця 1 - Основні результати ідентифікації параметрів гетероскедастичних еталонних сумішей

Результати ідентифікації параметрів еталонних сумішей	Міра перетинання класів у суміші			
	Неперетинні класи	Слабко перетинні класи	Сильно перетинні класи	Вироджений випадок сильно перетинних класів
M_1	18,020	17,780	19,20	26,030
Надійні інтервали для M_{e1}	17,79855–18,24510	17,75690–19,20544	17,41522–18,20332	31,63381–32,29824
M_2	31,860	31,910	35,490	26,130
Надійні інтервали для M_{e2}	31,48733–32,23157	31,91813–32,91000	30,42259–33,37667	31,41400–33,20089
M_3	54,010	54,940	53,660	34,020
Надійні інтервали для M_{e3}	53,23138–54,78253	53,05055–55,61780	52,40759–54,37062	29,94846–32,46344
Час навчання t_n^a	7	55	10	86
Помилки класифікації $N_{er\%}$	0	4,3	14,3	76,3

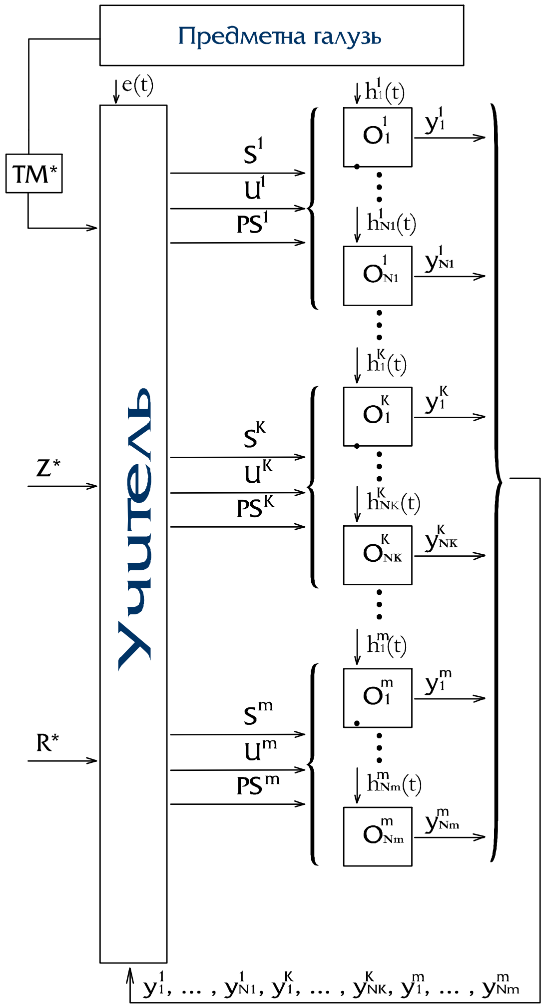

Рисунок 1 - Структурна схема процесу диференційованого навчання в традиційній системі вчитель–учень.

На рис. 1 прийнято такі умовні позначення: *ТМ** – теорія й метод навчання, використовувані вчителем; *Z** – цілі й завдання навчання; *R** – обмеження на процес навчання й керування; $y_1^1, \ldots, y_{N1}^1, y_1^k, \ldots, y_{Nk}^k, y_1^m, \ldots, y_{Nm}^m$ – стан учнів (вихідна координата об'єкта навчання й керування); $e(t)$ – завади (шуми) для вчителя; S^1, U^1, PS^1, …, S^k, U^k, PS^k, …, S^m, U^m, PS^m – відповідно диференційовані навчальні, керуючі та педагогічні (стимулюючі) впливи вчителя для класів (рівнів) учнів $O_1^1, \ldots, O_{N1}^1, O_1^k, \ldots, O_{Nk}^k, O_1^m, \ldots, O_{Nm}^m$, диференційованих за певним параметром (параметрами), що визначає їх когнітивні характеристики; m – кількість класів (рівнів), на які диференційовано учнів, $1 \le k \le m$; $h_1^1(t), \ldots, h_{N1}^1(t)$, $h_1^k(t), \ldots, h_{Nk}^k(t)$, $h_1^m(t), \ldots, h_{Nm}^m(t)$ – завади (шуми), що впливають на учнів, диференційованих на m класів (рівнів).

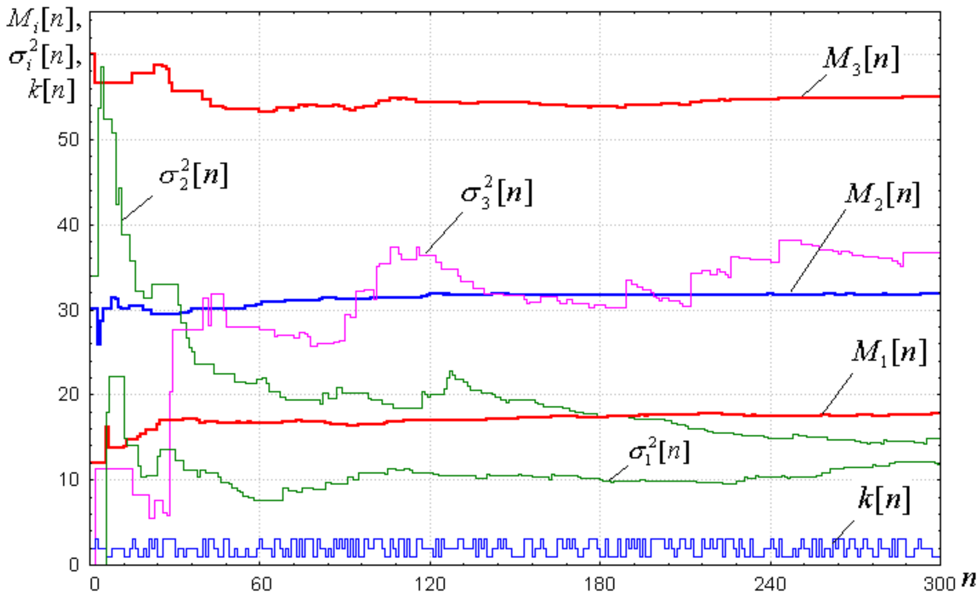

Рисунок 2 - Результати ідентифікації параметрів еталонних сумішей слабко перетинних класів ($M_{e1} = 18$; $\sigma_{e1} = 4$; $P_{k1}^* = 0{,}4$; $M_{e2} = 32$; $\sigma_{e2} = 2{,}0$; $P_{k2}^* = 0{,}3$; $M_{e3} = 54$; $\sigma_{e3} = 7{,}0$; $P_{k3}^* = 0{,}3$).

Час навчання (адаптації) t_n^a залежно від міри перетинання класів в еталонних сумішах змінюється від 7 до 86 кроків (див. табл. 1), помилки класифікації змінюються від 0% (для неперетинних класів) до 14,3% (для сильно перетинних класів). У виродженому випадку сильно перетинних класів з однаковими математичними сподіваннями й різними дисперсіями помилки класифікації становлять 76,3%.

Таким чином, виходячи з результатів чисельного дослідження алгоритмів (25), можна зробити висновок про ефективність і доцільність їх використання для диференціювання учнів на рівні (класи).

Основні висновки

1. Розроблено адаптивні алгоритми класифікації (диференціювання) учнів на класи (рівні, підгрупи) на основі корекції функції втрат залежно від нелінійних функцій дисперсій центрів класів у результуючій гетероскедастичній суміші розподілів (25). Чисельним дослідженням алгоритмів (25) на еталонних гетероскедастичних сумішах установлено, що оцінки центрів класів для неперетинних і слабко перетинних класів є незсунутими. Для сильно перетинних класів спостерігається зсув оцінок центрів класів, що не перевищує 6,35% від меж надійних інтервалів. Для виродженого випадку сильно перетинних класів з однаковими математичними сподіваннями й різними дисперсіями алгоритми (25) дають зсунуті оцінки для параметрів суміші.

2. Установлено, що час навчання (адаптації) t_n^a залежно від міри перетинання класів в еталонних сумішах змінюється від 7 до 86 кроків, а помилки класифікації (диференціювання) учнів на класи (рівні, підгрупи) $N_{er\%}$ змінюються від 0% (для неперетинних класів) до 14,3% (для сильно перетинних класів), що є цілком прийнятним при практичному використанні алгоритмів.

3. Показано, що розроблені алгоритми (25) доцільно використовувати в автоматизованих навчальних системах, де потрібна класифікація (диференціювання) учнів на класи за одним з параметрів, які складають образ учня.

The author of the work considers the modified algorithms of students' level (class) differentiation according to initial parameters which form student's model. The results of their numerical research are given.

Література

1. Зархин В.Г. Психофизиологические различия учащихся в процессе обучения с помощью ЭВМ // Вопросы психологии. – 1979. – № 1. – С. 79–83.
2. Хоменко Л.Я. Нові методи навчання на Заході // Іноземні мови. – 1998. – № 1. – С. 32–35.
3. Логвіна-Бик Т.А. Педагогічне керівництво диференційованим навчанням учнів середніх і старших класів (на прикладі предметів біологічного циклу): Автореф. дис… канд. пед. наук / Нац. пед. ун-т ім. М.П.Драгоманова. – Київ, 1999. – 20 с.
4. Есипова Н.Д. Дифференцированный подход в обучении информатике // Информатика и образование. – 1996. – № 6. – С. 27–34.
5. Меняйленко О.С. Розробка і дослідження субоптимальних адаптивних алгоритмів диференціювання учнів на рівні (класи) // Вісн. Східноукр. нац. ун-ту ім. В.Даля. – 2004. – № 4(74). – С. 38–49.
6. Menaylenko A.S. Computer-aided instruction and test of knowledge // New Media and Telematic Technologies for Education in Eastern European Countries. Twente University Press, Enschede. – 1997. – С. 314–318.
7. Меняйленко О.С. Про один підхід до побудови імовірнісної моделі учня в автоматизованих навчальних системах // Вісн. Східноукр. нац. ун-ту ім. В.Даля. – Луганськ, 2002. – № 3. – С. 138–141.
8. Цыпкин Я.З. Основы обучающихся систем. – М.: Наука, 1970. – 252 с.
9. Цыпкин Я.З. Основы информационной теории идентификации. – М.: Наука, 1984. – 320 с.
10. А. с. 1327067 СССР, МКИ4 G 05 В 23/02. Устройство для идентификации параметров возмущающих воздействий / А.С.Меняйленко, В.А.Ульшин (СССР). – № 3930458/24-24; Заявлено 12.07.85. Опубл. 30.07.87, Бюл. № 28.
11. Меняйленко О.С. Автоматизовані педагогічні навчальні системи: Монографія. – Луганськ: Альма-матер, 2003. – 272 с.
12. Дисперсионная идентификация / Под ред. Н.С. Райбмана. – М.: Наука, Главная ред. физ.-мат. литературы, 1981. – 336 с.
13. Дорофеюк А.А. Алгоритмы автоматической классификации // Автоматика и телемеханіка. – 1971. – № 12. – С. 78–113.

14. Айвазян С.А., Бежаева З.И., Староверов О.В. Классификация многомерных наблюдений. – М.: Статистика, 1974. – 240 с.

Одержано 03.10.2005 р.