

УДК 004.93.1'

А. В. ВАСИЛЬЄВ, А. С. ДОВБИШ, В. О. ЛЮБЧАК, А. С. ОСАДЧИЙ

Сумський державний університет, Україна

## ІНФОРМАЦІЙНИЙ СИНТЕЗ ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ ОЦІНКИ ЯКОСТІ НАВЧАЛЬНОГО КОНТЕНТУ

*Запропоновано алгоритм оптимізації параметрів функціонування здатної навчатися інформаційно-аналітичної системи оцінки якості навчального контенту випускової кафедри вищого навчального закладу з метою його адаптації до потреб ринку праці. Машинне навчання системи здійснено в рамках інформаційно-екстремальної інтелектуальної технології аналізу даних, що ґрунтується на максимізації інформаційної спроможності системи в процесі її навчання. Як параметри функціонування системи, що оптимізуються на етапі машинного навчання, розглядаються контрольні допуски на значення ознак розпізнавання та геометричні параметри розбиття простору ознак на класи якості навчального контенту. Алгоритм машинного навчання полягає в реалізації ітераційної процедури наближення глобального максимуму інформаційного критерію функціональної ефективності системи до його граничного значення в допустимій області визначення його функції. Результати реалізації запропонованого алгоритму машинного навчання системи отримано на прикладі оцінки змістовних модулів навчальних дисциплін бакалаврського рівня за спеціальністю «Комп'ютерні науки та інформаційні технології».*

**Ключові слова:** машинне навчання, інформаційно-екстремальний алгоритм, розпізнавання, ознака розпізнавання, вкладений контейнер класу розпізнавання, інформаційний критерій, навчальний контент.

### Вступ

Якість навчального процесу визначається користю, яку отримують випускники університету та їх роботодавці [1]. Тому розробка інформаційно-аналітичної системи для встановлення постійного зворотного зв'язку між роботодавцями та навчальним закладом, який дозволяє виявити ефективні форми навчання, корегувати навчальні програми та позбутися неінформативного та застарілого навчального контенту, є актуальною науково-технічною задачею. Для переважної більшості навчальних закладів основним методом адаптації навчального контенту випускових кафедр до вимог ринку праці все ще залишається анкетування, аналіз результатів якого потребує значних матеріально-часових витрат. Тому важливою функцією інформаційно-аналітичної системи є автоматизація оброблення результатів опитування респондентів на основі інтелектуального аналізу даних [2, 3]. При цьому одним із напрямків підвищення функціональної ефективності комп'ютерних систем контролю якості навчального контенту є їх інформаційний синтез на основі машинного навчання та розпізнавання образів. Проте науково-методологічні питання створення таких систем все ще залишаються недостатньо дослідженими через багатofакторність і довільні поча-

ткові умови процесу оцінювання та вплив неконтрольованих факторів на організацію і керування навчальним процесом.

Одним із перспективних шляхів подолання вказаних вище ускладнень науково-методологічного характеру є застосування ідей і методів інформаційно-екстремальної інтелектуальної технології (ІЕІ-технології), що ґрунтується на максимізації інформаційної спроможності системи в процесі її машинного навчання [4, 5]. В праці [6] розроблено інформаційно-екстремальний алгоритм машинного навчання інформаційно-аналітичної системи оцінки якості навчального контенту для полімодальної структури гіперсферичних контейнерів класів розпізнавання, яка передбачає наявність декількох центрів розсіювання векторів-реалізацій образів. Але у випадках прийняття рішень за оціночними шкалами алфавіт класів розпізнавання, які характеризують відповідні рівні якості, є впорядкованим. Це дозволяє для побудови вирішальних правил застосовувати так звану ієрархічну структуру контейнерів класів розпізнавання, які відновлюються в процесі машинного навчання в радіальному базисі простору ознак.

В статті розглядається метод інформаційно-екстремального навчання інформаційно-аналітичної системи адаптації навчального контенту випускової кафедри до вимог ринку праці з оптимізацією (тут і

далі в інформаційному розумінні) геометричних параметрів гіперсферичних контейнерів класів розпізнавання з ієрархічною структурою.

## 1. Формалізована постановка задачі

При розробленні в рамках геометричного підходу структури вирішальних правил, побудова яких є головною задачею інформаційно-екстремального машинного навчання, приймемо гіпотезу про наявність базового класу  $X_1^0$ , який характеризує найбільш бажаний для студентів і респондентів-роботодавців рівень якості навчального контенту. Інші класи розпізнавання, які характеризують відповідні нижчі рівні якості контенту слід розглядати як відхилення від базового класу. Оскільки респонденти оцінюють якість навчального контенту за кількісною оціночною шкалою, то алфавіт класів розпізнавання є впорядкованим. У цьому випадку як експериментально доведено в праці [5] доцільно застосовувати вкладену структуру контейнерів класів розпізнавання, які мають єдиний центр розсіювання реалізацій образів. Наприклад, перший контейнер будується для базового класу  $X_1^0$ , який характеризує якість навчального контенту з оцінкою “відмінно” і є вкладеним у контейнер класу  $X_2^0$ , який характеризує рівень якості з оцінкою “добре”. Таким чином будується ієрархічна структура контейнерів для всіх упорядкованих класів розпізнавання.

Застосування унімодального класифікатора із вкладеними контейнерами класів розпізнавання суттєво зменшує обчислювальну трудомісткість алгоритму машинного навчання у порівнянні з полімодальним класифікатором, в якому, згідно з принципом компактності, кожний клас розпізнавання має свій центр розсіювання реалізацій. Наприклад, при вкладених гіперсферичних контейнерах в процесі навчання відпадає необхідність обчислення усереднених векторів-реалізацій класів розпізнавання, вершини яких визначають їх геометричні центри. При цьому специфіка навчання класифікатора з вкладеними контейнерами полягає у відсутності процедури визначення для кожного класу найближчого сусіда, оскільки класи розпізнавання апріорно є впорядкованими, що суттєво підвищує оперативність навчання. Крім того, не потрібно визначати зовнішній радіус найбільш віддаленого від центру розсіювання класу розпізнавання.

Розглянемо інформаційно-аналітичну систему оцінки якості навчального контенту, ядром якої є здатна навчатися система підтримки прийняття рішень (СППР) із вкладеною структурою контейнерів

класів розпізнавання. Нехай дано структурований алфавіт  $\{X_m^0 | m = \overline{1, M}\}$  класів розпізнавання, які характеризують рівні якості навчального контенту. За результатами опитування роботодавців та їх провідних фахівців сформовано вхідну багатовимірну (векторну) навчальну матрицю типу «об'єкт-властивість»

$$\{y_{m,i}^j, | i = \overline{1, N}, j = \overline{1, n}\},$$

де  $N$  – кількість ознак розпізнавання, які відповідають експертним оцінкам рівня якості змістовних модулів навчального контенту за кількісною шкалою,  $n$  – кількість векторів-реалізацій кожного класу.

Задано структурований вектор параметрів функціонування СППР

$$g = \langle x_1, d_m, \delta_i \rangle, m = \overline{1, M}; i = \overline{1, N}, \quad (1)$$

де  $x_1$  – статистично усереднений двійковий вектор-реалізація базового класу  $X_1^0$ , який визначає геометричний центр вкладених гіперсферичних контейнерів класів розпізнавання, що відновлюється в радіальному базисі простору ознак розпізнавання;  $d_m$  – зовнішній радіус контейнера класу  $X_m^0$ ;  $\delta_i$  – параметр, який дорівнює половині симетричного поля контрольних допусків на  $i$ -ту ознаку розпізнавання.

Крім того, задано обмеження на радіуси контейнерів класів розпізнавання:

$$d_1 > 0; d_{m-1} < d_m < d_{m+1},$$

і на параметр поля контрольних допусків на  $i$ -ту ознаку розпізнавання

$$\delta_i \in [0; \delta_{H,i} / 2],$$

де  $\delta_{H,i}$  – нормоване поле допусків на  $i$ -ту ознаку розпізнавання, яке визначає область значень параметра  $\delta_i$ .

В процесі інформаційно-екстремального навчання відновлення контейнерів класів розпізнавання відбувається в радіальному базисі дискретного простору ознак розпізнавання. При цьому зовнішній радіус, наприклад, контейнера класу  $X_m^0$  в бінарному просторі Хеммінга вимірюється кодовою відстанню

$$d_m = d(x_1 \oplus \lambda_m) = \sum_{i=1}^N x_{1,i} \oplus \lambda_{m,i},$$

де  $\lambda_m$  – реалізація класу  $X_m^0$ , вершина якої знаходиться на роздільній гіперсфері з класом  $X_{m+1}^0$ ;

$x_{1,i}, \lambda_{m,i}$  –  $i$ -та ознака розпізнавання в реалізаціях;

$x_1$  і  $\lambda_m$  відповідно;  $\oplus$  – символ операції складання за модулем два.

Необхідно в процесі навчання СППР визначити оптимальні значення координат вектора (1), які забезпечують максимум інформаційного критерію функціональної ефективності (КФЕ) навчання СППР розпізнавати реалізації класу  $X_m^0$ , пошук якого здійснюється в робочій (допустимій) області визначення його функції.

На етапі екзамену, тобто при функціонуванні СППР безпосередньо в режимі моніторингу, необхідно прийняти рішення про належність сформованої респондентом реалізації до одного з класів алфавіту  $\{X_m^0\}$  і у разі необхідності видати рекомендації щодо його корекції.

## 2. Інформаційно-екстремальний синтез здатної навчатися СППР

Інформаційно-екстремальний синтез здатної навчатися СППР складається із двох етапів:

1) машинного навчання, на якому шляхом оптимізації за інформаційним критерієм будуються вирішальні правила;

2) машинного екзамену, тобто функціонування системи безпосередньо в робочому режимі прийняття класифікаційних рішень про належність реалізації, що розпізнається, до одного із класів заданого алфавіту.

Інформаційно-екстремальне машинне навчання полягає в реалізації ітераційної процедури пошуку глобального максимуму інформаційного КФЕ в робочій області визначення його функції. Для алгоритму навчання з оптимізацією системи контрольних допусків на ознаки розпізнавання така двохциклічна процедура має вигляд

$$\delta^* = \arg \max_{G_\delta} \{ \max_{G_E \cap \{k\}} \bar{E}^{(k)} \}, \quad (2)$$

де  $\bar{E}^{(k)}$  – значення обчисленого на  $k$ -му кроці навчання усередненого за алфавітом класів розпізнавання КФЕ;  $G_\delta$  – область допустимих значень параметра поля контрольних допусків на ознаки розпізнавання;  $\{k\}$  – множина кроків навчання.

У процедурі (2) основними задачами внутрішнього циклу є обчислення інформаційного КФЕ, пошук його глобального максимуму та визначення екстремального значення зовнішнього радіуса гіперсферичного контейнера, яке приймається за оптимальне.

Розглянемо алгоритм паралельної оптимізації контрольних допусків на ознаки розпізнавання, за яким параметр  $\delta$  поля контрольних допусків змінюється одночасно для всіх ознак. Така схема алгорит-

му (2) є доцільною у випадку, коли ознаки розпізнавання мають єдину шкалу вимірювань. Тоді нижній  $A_{НК,i}$  і верхній  $A_{ВК,i}$  контрольні допуски на  $i$ -ту ознаку розпізнавання обчислюються відповідно за формулами

$$A_{НК,i} = y_i - \delta; \quad A_{ВК,i} = y_i + \delta, \quad (3)$$

де  $y_i$  – усереднене значення  $i$ -ої ознаки розпізнавання в реалізаціях базового класу  $X_1^0$ .

Вхідними даними є тривимірний масив навчальної матриці  $\{y_{m,i}^j\}$  та поле  $\delta_H$  нормованих допусків, яке для всіх ознак розпізнавання є однаковим і визначає область значень системи контрольних допусків.

Розглянемо схему алгоритму інформаційно-екстремального машинного навчання СППР за процедурою (2) з паралельною оптимізацією системи контрольних допусків (СКД) на ознаки розпізнавання:

1) онулюється лічильник класів розпізнавання:  $m := 0$ ;

2)  $m := m + 1$ ;

3) онулюється лічильник кроків зміни параметра  $\delta$  поля контрольних допусків на ознаки розпізнавання:  $k := 0$ ;

4)  $k := k + 1$ ;

5) шляхом статистичного усереднення реалізацій вхідної навчальної матриці  $\{y_{1,i}^j\}$

базового класу  $X_1^0$  визначається вектор  $y_1$ , відносно координат якого на кожному кроці навчання задаються нижні та верхні контрольні допуски на ознаки розпізнавання;

6) за формулами (3) обчислюються нижній  $A_{НК,i}[k]$  і верхній  $A_{ВК,i}[k]$  контрольні допуски для всіх ознак розпізнавання;

7) формується тривимірний масив бінарної навчальної матриці  $\{x_{m,i}^j\}$ , елементи якої обчислюються за правилом

$$x_{m,i}^j[k] = \begin{cases} 1, & \text{якщо } A_{НК,i}[k] < y_{m,i}^j < A_{ВК,i}[k]; \\ 0, & \text{якщо інакше;} \end{cases}$$

8) шляхом статистичного усереднення реалізацій бінарної навчальної матриці  $\{x_{1,i}^j\}$

базового класу  $X_1^0$  визначається вектор  $x_1$ , вершина якого визначає центр розсіювання реалізацій всіх класів розпізнавання;

9) онулюється лічильник кроків зміни радіусу гіперсферичного контейнеру класу розпізнавання  $X_m^0$ ;  $d[m] := 0$ ;

10)  $d[m] := d[m] + 1$ ;

11) за навчальними матрицями класу  $X_m^o$  і суміжного класу  $X_{m+1}^o$  обчислюється інформаційний КФЕ  $E[m, k]$ ;

12) якщо  $d[m] < N$ , то виконується пункт 10, інакше – пункт 13;

13) в робочій області  $G_E$  визначається максимальне значення  $E^*[m, k]$  інформаційного критерію;

14) якщо  $\delta < \delta_H / 2$ , то виконується пункт 4, інакше – пункт 15;

15) визначається глобальний максимум  $E^*[m]$  інформаційного критерію в його робочій області;

16) обчислюється оптимальне значення зовнішнього радіусу контейнера класу  $X_m^o$

$$d^*[m] = \arg E^*[m];$$

17) якщо виконується умова  $m < M - 1$ , то реалізується пункт 2, інакше – пункт 18;

18) обчислюється усереднене значення інформаційного КФЕ за формулою

$$\bar{E}^* = \frac{1}{M-1} \sum_{m=1}^{M-1} E^*[m],$$

визначається оптимальне значення параметра поля контрольних допусків

$$\delta^* = \arg \bar{E}^*$$

і за формулою (3) обчислюються оптимальні нижні  $\{A_{НК,i}^*\}$  і верхні  $\{A_{ВК,i}^*\}$  контрольні допуски на ознаки розпізнавання;

19) будуються вирішальні правила у вигляді продукцій

if  $d[x_1 \oplus x^{(i)}] \leq d_1^*$  then  $x^{(i)} \in X_1^o$ ;

( $\forall m < M$ ) {if  $d_m^* \leq d[x_1 \oplus x^{(i)}] \leq d_{m+1}^*$  then  $x^{(i)} \in X_m^o$ };

if  $d[x_1 \oplus x^{(i)}] > d_{M-1}^*$  then  $x^{(i)} \in X_M^o$ , (4)

де  $x^{(i)}$  – реалізація образу, що розпізнається;

20) ЗУПИН.

Побудовані в процесі навчання вирішальні правила (4) характеризуються мінімальною обчислювальною трудомісткістю, що є перевагою геометричного підходу, і дозволяють приймати класифікаційні рішення в реальному темпі часу.

Як КФЕ інформаційно-екстремального навчання СППР може використовуватися будь-яка інформаційна міра [3]. Розглянемо модифіковану інформаційну міру Кульбака у вигляді добутку логарифмічного відношення повної ймовірності правильного прийняття рішень  $P_t$  до повної

ймовірності помилкового прийняття рішень  $P_f$  на їх різницю. Для рівноймовірних двох альтернативних гіпотез, що характеризує найбільш складний у статистичному розумінні випадок прийняття рішень, міру Кульбака подамо у вигляді

$$\begin{aligned} E_m^{(k)} &= \log_2 \frac{P_{t,m}^{(k)}}{P_{f,m}^{(k)}} [P_{t,m}^{(k)} - P_{f,m}^{(k)}] = \\ &= \left| \begin{array}{l} P_{t,m}^{(k)} = 0,5 \cdot D_{1,m}^{(k)} + 0,5 \cdot D_{2,m}^{(k)} \\ P_{f,m}^{(k)} = 0,5 \cdot \alpha_m^{(k)} + 0,5 \cdot \beta_m^{(k)} \end{array} \right| = \\ &= 0,5 [D_{1,m}^{(k)} + D_{2,m}^{(k)} - (\alpha_m^{(k)} + \beta_m^{(k)})] * \\ &\quad * \log_2 \left( \frac{D_{1,m}^{(k)} + D_{2,m}^{(k)}}{\alpha_m^{(k)} + \beta_m^{(k)}} \right) = \\ &= [D_{1,m}^{(k)} - \beta_m^{(k)}] \log_2 \left( \frac{1 + (D_{1,m}^{(k)} - \beta_m^{(k)})}{1 - (D_{1,m}^{(k)} - \beta_m^{(k)})} \right), \quad (5) \end{aligned}$$

де  $D_{1,m}^{(k)}$  – перша достовірність, обчислена на  $k$ -му кроці навчання СППР розпізнавати реалізації класу  $X_m^o$ ;  $D_{2,m}^{(k)}$  – друга достовірність;  $\alpha_m^{(k)}$  – помилка першого роду;  $\beta_m^{(k)}$  – помилка другого роду.

Нормовану модифікацію критерію (5) подамо у вигляді

$$E_m^{*(k)} = \frac{E_m^{(k)}}{E_{\max}^{(k)}},$$

де  $E_{\max}^{(k)}$  – значення критерію (5) при  $D_{1,m}^{(k)} = 1$  і  $\beta_m^{(k)} = 0$ .

Таким чином, процес інформаційно-екстремального навчання полягає в реалізації процедури пошуку глобального максимуму функції інформаційного критерію в робочій області її визначення і ітераційного наближення цього максимуму до його граничного максимального значення.

Алгоритм екзамену має такі вхідні дані:  $M$  – кількість класів, які СППР навчена розпізнавати;  $\{d_m^*\}$  – масиви оптимальних радіусів контейнерів класів розпізнавання; оптимальні нижні  $\{A_{НК,i}^*\}$  і верхні  $\{A_{ВК,i}^*\}$  контрольні допуски на ознаки розпізнавання, визначенні на етапі навчання;  $\{x^{(i)}\}$  – масив двійкових векторів-реалізацій образу, що розпізнається.

Розглянемо основні етапи реалізації алгоритму екзамену:

1. Формування лічильника класів розпізнавання:  $m := m + 1$ .
2. Формування лічильника числа реалізацій, що розпізнаються:  $j := j + 1$ .
3. Обчислення кодової відстані  $d(x_1 \oplus x^{(j)})$ .
4. Порівняння: якщо  $m \leq M$ , то виконується крок 1, інакше – крок 5.
5. Визначення за вирішальними правилами (4) класу  $X_m^0$ , до якого належить реалізація образу, що розпізнається.
6. ЗУПИН.

Після класифікації сформованих респондентами реалізацій задачею інформаційно-аналітичної системи є аналіз тенденції зміни запитів роботодавців шляхом оцінки частот віднесення реалізацій до відповідних класів розпізнавання.

### 3. Приклад реалізації алгоритму машинного навчання

Реалізацію запропонованого алгоритму розглянемо на прикладі інформаційно-екстремального навчання інформаційно-аналітичної системи оцінки якості навчального контенту випускової кафедри комп'ютерних наук Сумського державного університету. При опитуванні як респонденти виступали провідні фахівці десяти ІТ-компаній, серед яких переважну кількість склали випускники кафедри. Вхідна навчальна матриця формувалася шляхом накопичення оцінок за стобальною шкалою 50 змістовних модулів 10 навчальних дисциплін підготовки бакалаврів за спеціальністю «Комп'ютерні науки та інформаційні технології», розбиті на шість блоків: «Мови програмування та операційні системи», «Бази даних», «Web-програмування», «Теорія алгоритмів і структури даних», «Комп'ютерні мережі» і «Системний аналіз». Алфавіт складався із трьох класів розпізнавання: клас  $X_1^0$  відповідав навчальному контенту високої якості, клас  $X_2^0$  характеризував навчальний контент з оцінкою «Задовільно», а клас  $X_3^0$  – «Незадовільно». Обсяг навчальної вибірки для кожного класу становив  $n = 40$ , а потужність словника ознак дорівнювала кількості змістовних модулів дисциплін  $N = 50$ .

Машинне навчання інформаційно-аналітичної системи здійснювалося за алгоритмом (2) з паралельною оптимізацією контрольних допусків на ознаки розпізнавання. На рис. 1 показано графік залежності усередненого за алфавітом класів розпізнавання потужності  $M-1$  інформаційного КФЕ (5) від параметра  $\delta$  поля контрольних допусків на ознаки

розпізнавання, одержаний в процесі інформаційно-екстремального машинного навчання інформаційно-аналітичної системи, яка отримала умовну назву ІАС «Випускник». На графіку штрихована ділянка позначає робочу область визначення функції інформаційного критерію (5), в якій одночасно виконуються умови:  $D_{1,m} > 0,5$ ,  $D_{2,m} > 0,5$ , і  $d_m > d_{m-1}$ , тобто перша і друга достовірності перевершують відповідно помилки першого і другого роду.

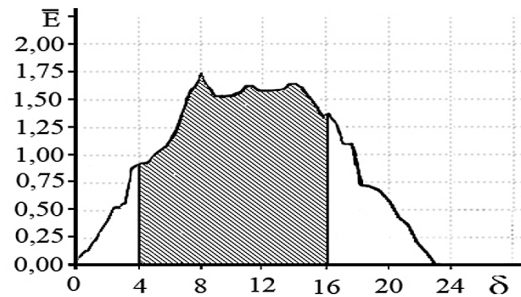


Рис. 1. Графік залежності КФЕ від параметра поля контрольних допусків

Аналіз рис. 1 показує, що оптимальний параметр поля контрольних допусків дорівнює  $\delta^* = \pm 8$  градацій стобальної оціночної шкали при максимальному значенні усередненого критерію  $\bar{E}^* = 1,75$ . При цьому нормований КФЕ (6) не досягає свого максимального граничного значення, але він суттєво перевершує наведений в праці [5] результат, отриманий при машинному навчанні СППР з полімодальною структурою вирішальних правил.

Оскільки метою інформаційно-екстремального навчання є відновлення в просторі ознак оптимальних контейнерів класів розпізнавання, то на рис. 2 наведено графіки залежності КФЕ (5) від радіусів контейнерів класів  $X_1^0$  і  $X_2^0$ .

Аналіз рис. 2 показує, що оптимальні значення зовнішніх радіусів гіперсферичних контейнерів класів  $X_1^0$  і  $X_2^0$  відповідно дорівнюють  $d_1 = 21$  і  $d_2 = 38$  (в кодових одиницях). При цьому зовнішній радіус класу  $X_3^0$  дорівнює розмірності бінарного простору, тобто  $d_3 = N$ .

Таким чином, вище наведений алгоритм машинного навчання дозволяє отримати квазіоптимальні параметри функціонування інформаційно-аналітичної системи, але не забезпечує досягнення граничного максимуму критерію (5). З метою побудови безпомилкових за

навчальною матрицею вирішальних правил слід згідно з принципом відкладених рішень А. Г. Івахненка оптимізувати інші параметри навчання, які впливають на функціональну ефективність інформаційно-аналітичної системи.

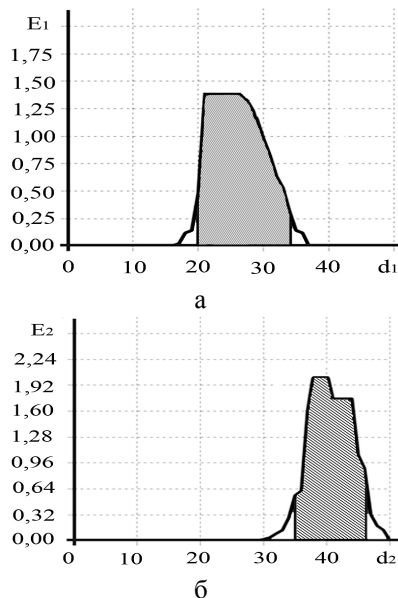


Рис. 2. Графіки залежності КФЕ навчання від радіусів контейнерів класів розпізнавання:

а – клас  $X_1^0$ ; б – клас  $X_2^0$

## Висновки

1. У рамках ІЕІ-технології розроблено алгоритм машинного навчання інформаційно-аналітичної системи оцінки відповідності навчального контенту випускової кафедри ВНЗ вимогам ринку праці. При цьому запропоновано при побудові вирішальних правил використовувати контейнери класів розпізнавання, що відновлюються в процесі машинного навчання в радіальному базисі простору ознак, із ієрархічною структурою. Одержані в наведеному прикладі результати реалізації запропонованого алгоритму машинного навчання підтверджують підвищення достовірності прийняття оціночних рішень у порівнянні з класифікаторами із полімодальною структурою контейнерів класів розпізнавання.

2. Оскільки вирішальні правила, отримані в процесі навчання, не є безпомилковими за навчальною матрицею, то для підвищення функціональної ефективності інформаційно-аналітичної системи необхідно оптимізувати додаткові параметри функціонування системи, наприклад, рівень селекції координат еталонних двійкових векторів класів розпізнавання або перейти до більш складної форми радіально-базисних контейнерів. Крім того, подальші дослідження слід спрямувати на розроблення алгоритмів кластер-

аналізу вхідних даних, що дозволить автоматизувати формування вхідної навчальної матриці, і факторного кластер-аналізу, що дозволить в режимі екзамену розширювати алфавіт класів розпізнавання та здійснювати перенавчання системи.

## Література

1. Агапова, М. О. Про оптимізацію навчального матеріалу [Текст] / М. О. Агапова, В. Г. Кремень // Теорія і практика управління соціальними системами. Науково-практичний журнал. – 2012. – № 1. – С. 34–39.
2. Xu, G. Applied Data Mining [Text] / G. Xu, Y. Zong, Z. Yang. – CRC Press, 2013. – 284 p.
3. Довбиш, А. С. Інтелектуальні інформаційні технології в електронному навчанні [Текст] / А. С. Довбиш. – Суми : Видавництво Сумського державного університету, 2013. – 172 с.
4. Dovbysh, A. S. Information-Extreme Method for Classification of Observations with Categorical Attributes [Text] / A. S. Dovbysh, V. V. Moskalenko, A. C. Rizhova // *Cibernetica and Systems Analysis*. – 2016. – V. 52. – № 2. – P. 45-52.
5. Інформаційно-екстремальне навчання системи оцінки якості навчального контенту випускової кафедри [Текст] / А. С. Довбиш, Є. С. Кулік, З. В. Козлов, А. С. Осадчий // *Радіоелектронні та комп'ютерні системи*. – 2016. – № 3 (77). – С. 71-77.

## References

1. Agapova, M. O., Kremen, V. G. Pro optimizaciju navchal'nogo material [About optimization of educational material]. *Teorija i praktyka upravlinnja social'nymy systemamy*. *Naukovo-praktychnyj zhurnal*. Kharkiv, 2012, no. 1, pp. 34-39.
2. Xu, G., Zong, Y., Yang, Z. *Applied Data Mining*. CRC Press, 2013. 284 p.
3. Dovbysh, A. S. Intel'ektual'ni informacijni tehnologii' v elektronnomu navchanni [Intellectual information technologies in electronic learning], Sumy, Sumy State University Publ., 2013. 172 p.
4. Dovbysh, A. S., Moskalenko, V. V., Rizhova, A. S. Information-Extreme Method for Classification of Observations with Categorical Attributes. *Cibernetica and Systems Analysis*, 2016, vol. 52, no. 2, pp. 45-52.
5. Dovbysh, A. S., Kulik, E. S., Kozlov, Z. V., Osadchij, A. S. Informatcii'no-ekstremal'ne navchannia sistemi otcinki iakosti navchal'nogo kontentu vipuskoivoi kafedri [Quality assessment of educational content in graduating departments using information-extreme intellectual technology of data analysis]. *Radioelektronni i komp'uterni sistemi - Radioelectronic and computer systems*, 2016, no. 3 (77), pp. 71-77.

Надійшла до редакції 12.10.2016, розглянута на редколегії 09.12.2016

## ИНФОРМАЦИОННЫЙ СИНТЕЗ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКОЙ СИСТЕМЫ ОЦЕНКИ КАЧЕСТВА УЧЕБНОГО КОНТЕНТА

*А. В. Васильев, А. С. Довбыш, В. О. Любчак, А. С. Осадчий*

Предложен алгоритм оптимизации параметров функционирования способной учиться информационно-аналитической системы оценки качества учебного контента выпускной кафедры высшего учебного заведения с целью его адаптации к потребностям рынка труда. Машинное обучение системы осуществлено в рамках информационно-экстремальной интеллектуальной технологии анализа данных, которая основывается на максимизации информационной возможности системы в процессе ее учебы. Как параметры функционирования системы, которые оптимизируются на этапе машинной учебы, рассматриваются контрольные допуски на значение признаков распознавания и геометрические параметры разбиения пространства признаков на классы качества учебного контента. Алгоритм машинной учебы заключается в реализации итерационной процедуры приближения глобального максимума информационного критерия функциональной эффективности системы к его предельному значению в допустимой области определения его функции. Результаты реализации предложенного алгоритма машинной учебы системы получены на примере оценки содержательных модулей учебных дисциплин бакалаврского уровня за специальностью "Компьютерные науки и информационные технологии".

**Ключевые слова:** машинное обучение, информационно-экстремальный алгоритм, распознавание, признак распознавания, вложенный контейнер класса распознавания, информационный критерий, учебный контент.

## INFORMATIONAL SYNTHESIS OF INFORMATION AND ANALYTICAL SYSTEM OF ASSESSMENT QUALITY OF LEARNING CONTENT

*A. V. Vasilev, A. S. Dovbysh, V. O. Lubchak, A. S. Osadchiy*

In this paper the algorithm for parameters optimization of learning information-analytical system assessing the quality of educational content graduating department of higher education with the aim to adapt it to the needs of the labor market is proposed. Machine learning systems implemented under the informational and extremal intellectual technology of data analysis based on maximizing the capacity of information in the process of learning. The parameters of system settings that are optimized for machine learning stage, examined control tolerances mentioned signs of recognition and the geometrical parameters of the partition feature space into classes of quality of educational content as parameters of decision rules are considered. The algorithm of machine learning implements iterative approximation the global maximum informational criterion to its limit values in admissible domain of its function. The results of the proposed algorithm of machine learning system obtained by the example of assessment content modules disciplines of Bachelor in specialty "Computer science and information technology."

**Key words:** machine learning, informational and extremal algorithm, recognition, signs of recognition, nested a container class of recognition, information criteria, learning content.

**Васильев Анатолий Васильевич** – канд. техн. наук, профессор, ректор, Сумський державний університет, Суми, Україна, e-mail: rector@sumdu.edu.ua.

**Довбиш Анатолій Степанович** – д-р техн. наук, професор, завідувач каф. комп'ютерних наук, Сумський державний університет, Суми, Україна, e-mail: kras@id.sumdu.edu.ua.

**Любчак Володимир Олександрович** – канд. техн. наук, доцент, проректор, Сумський державний університет, Суми, Україна, e-mail: lub@sumdu.edu.ua.

**Осадчий Андрій Сергійович** – аспірант каф. комп'ютерних наук, Сумський державний університет, Суми, Україна, e-mail: lim\_33@mail.ru.

**Vasylyev Anatoliy Vasilyevych** – rector, Sumy State University, Sumy, Ukraine.

**Dovbysh Anatoliy Stepanovych** – head of computer science department, Sumy State University, Sumy, Ukraine.

**Lyubchak Volodymyr Olexandrovych** – Vice-rector, Sumy State University, Sumy, Ukraine.

**Osadchiy Andriy Sergijovych** – lecturer of Konotop institute of Sumy State University, Konotop, Ukraine.