

В. І. ЗИМОВЕЦЬ, О. С. ПРИХОДЧЕНКО, М. І. МИРОНЕНКО

Сумський державний університет, Україна

ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНИЙ КЛАСТЕР-АНАЛІЗ ВХІДНИХ ДАНИХ ПРІ ФУНКЦІОНАЛЬНОМУ ДІАГНОСТУВАННІ

Метою дослідження є підвищення функціональної ефективності машинного навчання системи функціонального діагностування (СФД) багатоканатної шахтною підйомною машиною (ШПМ) шляхом кластер-аналізу діагностичних ознак. Для досягнення мети необхідно було вирішити такі **завдання**: формалізувати постановку задачі інформаційного синтезу здатної навчатися СФД, яка функціонує в режимі кластер-аналізу діагностичних ознак; запропонувати категорійну модель і на її основі розробити алгоритм інформаційно-екстремального кластер-аналізу діагностичних ознак, в процесі інформаційно-екстремального машинного навчання СФД здійснити фазифікацію вхідних нечітких даних шляхом оптимізації геометричних параметрів гіперсферичних контейнерів класів розпізнавання, які характеризують можливі технічні стани об'єкту діагностування; запропонований алгоритм реалізувати на прикладі інформаційного синтезу СФД багатоканатної ШПМ. **Об'єктом** дослідження є процесі інформаційного синтезу здатної навчатися (СФД, інтегрованої в автоматизовану систему керування багатоканатної ШПМ. **Предметом** дослідження є категорійні моделі, алгоритм інформаційно-екстремального машинного навчання СФД, яка функціонує в режимі кластер-аналізу діагностичних ознак та вирішальні правила. **Методи** дослідження базуються на ідеях і методах інформаційно-екстремальної інтелектуальної технології аналізу даних, теоретико-інформаційного підходу до оцінки функціональної ефективності машинного навчання та на геометричному підході теорії розпізнавання образів. Отримано такі **результати**: запропоновано категорійну модель і на її основі розроблено та програмно реалізовано алгоритм інформаційно-екстремального машинного навчання СФД багатоканатної ШПМ, який дозволяє автоматично формувати вхідну класифіковану нечітку навчальну матрицю, що суттєво зменшує часові та матеріальні втрати при створенні вхідного математичного опису. Отриманий результат було досягнуто шляхом кластер-аналізу структурованих векторів діагностичних ознак, отриманих за архівними даними для трьох класів розпізнавання, з використанням процедури k -середніх. Як критерій оптимізації параметрів машинного навчання розглядалася модифікована міра Кульбака у вигляді функціонала від точнісних характеристик діагностичних рішень і дистанційних критеріїв близькості класів розпізнавання. За отриманими в процесі машинного навчання оптимальними геометричними параметрами контейнерів класів розпізнавання побудовано вирішальні правила, які дозволили класифікувати вектори діагностичних ознак класів розпізнавання із достатньо високою повною ймовірністю прийняття правильних діагностичних рішень. **Висновки**. Наукова новизна отриманих результатів полягає в розробленні нового методу інформаційного синтезу СФД багатоканатної ШПМ, яка функціонує в режимі кластер-аналізу, що дозволило автоматично формувати вхідну класифіковану нечітку навчальну матрицю з наступною її дефазифікацією в процесі інформаційно-екстремального машинного навчання системи.

Ключові слова: інформаційно-екстремальна інтелектуальна технологія; машинне навчання; кластер-аналіз; система функціонального діагностування; навчальна матриця; інформаційний критерій оптимізації; багатоканатна шахтна підйомна машина.

Вступ

Одним із важливих завдань експлуатації електромеханічного обладнання є оцінка його технічного стану. При цьому перспективним шляхом підвищення функціональної ефективності системи діагностування вважається перехід від технічного обслуговування за регламентом до обслуговування за фактичним станом вузлів і пристроїв складної машини. Для організації такого обслуговування актуа-

льною задачею є створення науково-методологічних основ інформаційного синтезу системи функціонального діагностування (СФД), яка дозволяє оцінювати технічний стан обладнання в поточний момент часу, знаходити місця дефекту та визначати причини його появи. Іншою важливою задачею СФД є накопичення статистичного матеріалу для прогнозування зміни технічного стану вузлів і пристроїв обладнання та профілактичних заходів, удосконалення технології виготовлення і режимів експлуатації електромеханічного обладнання тощо. При цьому ос-

новні ускладнення при реалізації функціонального діагностування пов'язані з необхідністю обробки великого обсягу інформації при багато вимірності алфавіту класів розпізнавання, які характеризують можливі дефекти обладнання, та прийняття діагностичних рішень в реальному темпі часу. Подолання цих ускладнень можливе шляхом використання сучасних інтелектуальних інформаційних технологій аналізу даних на основі машинного навчання та розпізнавання образів.

У статті на прикладі багатоканатної шахтної підйомної машини (ШПМ) розглядається задача машинного навчання системи функціонального діагностування, яка функціонує в режимі кластер-аналізу діагностичних ознак.

Огляд літератури

На практиці БШПМ знайшли широке застосування для шахт з великою глибиною залягання корисних копалин. Через високі вимоги до безпечної роботи та надійності шахтні підйомні машини відносяться до об'єктів критичного призначення. Згідно з працею [1] поточний ресурс роботи найбільш важливих компонентів багатоканатної ШПМ визначається, як правило, шляхом підрахунку загального часу напрацювання машини в годинах. При цьому такою критерією використовується для відстеження ступеня спрацювання вузлів і компонентів, але не враховує вплив фактичних умов експлуатації та оточуючого середовища, які призводять до пошкоджень.

У праці [2] досліджувалися різні фактори, які викликають зміни в технічному стані ШПМ і з часом призводить до відповідних відмов. Суттєвою особливістю цих факторів є їх випадковий (стохастичний) характер. До чинників, які мають найбільш істотний вплив на швидкість зміни технічного стану деталі, відносяться: технологічні навантаження, що мають варіативний характер, зміна температури під час роботи, проблеми футеровки канатів тощо. Суттєвим недоліком переважної більшості сучасних систем функціонального діагностування БШПМ є диференційована оцінка технічного стану лише окремих елементів, підсистем і вузлів без взаємозв'язку з іншими підсистемами. У результаті такі системи лише фіксують локальні прояви відмов і не аналізують їх причини [3 – 5]. Крім того, збільшення обсягу інформації, що аналізується про стан обладнання веде до значних змін в методах роботи і вимагає автоматизації процесів обробки та інтелектуального аналізу діагностичних ознак. При цьому основними ускладненнями інформаційного синтезу інтелектуальних систем функціонального діагностування є [1]:

– наявність невизначеності вхідних даних, а також присутність неоднозначності (багатоваріантності) процесу пошуку рішення;

– необхідність розробки і обґрунтування рішень проблеми за умови жорстких часових обмежень, які визначаються ходом керованих процесів;

– необхідність коригування і введення додаткової інформації в процес пошуку рішень;

– інтерактивний характер логічного висновку рішень.

Крім того, важливою задачею функціонального діагностування є формування бази даних для прогнозування деградаційних процеси зміни технічного стану складних механічних систем, що експлуатуються. При функціональному діагностуванні ШПМ найбільш широко використовуються тепловий, акустико-емісійний, та вібраційний методи контролю [6, 7]. Ці методи дозволяють оцінювати поточний технічний стан окремих вузлів машини, але без врахування умов її експлуатації та аналізу причин виникнення дефектів, що унеможливує на практиці розв'язання задачі прогнозування технічних відмов. З іншого боку існуючі методи контролю та сучасні інформаційно-вимірювальні засоби дозволяють отримувати в процесі експлуатації ШПМ структуровані вектори діагностичних ознак для формування навчальної матриці і реалізації машинного навчання системи функціонального діагностування [8, 9]. При інтелектуальному аналізі даних найбільше поширеним є застосування нейроподібних структур [10 – 12]. Але інформаційний синтез систем функціонального діагностування на основі штучних нейронних мереж є проблематичним через великий обсяг діагностичних даних і велику потужність класів розпізнавання, які характеризують можливі технічні стани вузлів і деталей ШПМ. При цьому застосування багатосарових згорткових структур для стиснення вхідних даних призводить до втрати інформації і пониження оперативності прийняття діагностичних рішень. Крім того, необхідно враховувати ускладнення, які виникають при перенавчанні штучних нейронних мереж при розширенні словника діагностичних ознак і алфавіту класів розпізнавання.

Перспективним напрямком підвищення функціональної ефективності СФД є застосування ідей і методів так званої інформаційно-екстремальної інтелектуальної технології (ІЕІ-технології) аналізу даних, яка ґрунтується на максимізації інформаційної спроможності системи в процесі її машинного навчання [13, 14]. У праці [15] запропоновано алгоритм інформаційно-екстремального машинного навчання СФД ШПМ за навчальною матрицею, формування якої здійснювалося шляхом аналізу архівних даних. Оскільки функціональне діагностування складних машин і систем пов'язано з аналізом вели-

ких обсягів діагностичних ознак, то з метою зменшення часових витрат на формування вхідного математичного опису інтелектуальних систем доцільно використовувати методи кластер-аналізу [16, 17]. При цьому переважна більшість методів автоматичної класифікації, побудованих на процедурі к-середніх та її модифікаціях, використовують дистанційні критерії близькості. Тому на практиці такі методи автоматичної класифікації не дозволяють отримати чітке розбиття простору ознак на кластери за умови прийнятної з практичних міркувань гіпотези нечіткої компактності векторів-реалізацій класів розпізнавання.

Метою роботи є надання СФД багатоканатної ШПМ властивості самонавчання шляхом розроблення методу інформаційно-екстремального кластер-аналізу діагностичних ознак, який дозволяє автоматизувати формування вхідної навчальної матриці і забезпечує достатньо високу достовірність діагностування.

Постановка задачі

Розглянемо формалізовану постановку задачі інформаційно-екстремального машинного навчання СФД, яка функціонує в режимі кластер-аналізу діагностичних ознак. Нехай дано масив структурованих векторів діагностичних ознак $\{Y_{0,i}^{(j)} \mid i = \overline{1, N}; j = \overline{1, n}\}$, які відбивають властивості M класів розпізнавання. При цьому N – кількість діагностичних ознак в структурованому векторі, а n – загальна кількість векторів.

Крім того, задано структурований вектор параметрів машинного навчання:

$$g = \langle g_1, \dots, g_{\xi_1}, \dots, g_{\Xi_1} \rangle \quad (1)$$

з відповідними обмеженнями.

Необхідно в процесі машинного навчання з кластер-аналізом вхідних даних:

1) сформуванати апріорно класифіковану нечітку навчальну матрицю для M класів розпізнавання;

2) оптимізувати параметри вектору (1) за умови, що усереднений за алфавітом класів розпізнавання інформаційний критерій оптимізації параметрів машинного навчання набуває глобального максимуму в робочій області визначення його функції:

$$E^* = \frac{1}{M} \sum_{m=1}^M \max_G E_m, \quad (2)$$

де E_m – інформаційний критерій оптимізації параметрів навчання СФД розпізнавати вектори-

реалізації класу X_m^0 ; G — область допустимих значень параметрів машинного навчання;

3) за отриманими в процесі машинного навчання оптимальними геометричними параметрами контейнерів класів розпізнавання побудувати вирішальні правила;

4) в режимі екзамену перевірити функціональну ефективність машинного навчання.

Категорійні моделі машинного навчання

Категорійну модель інформаційно-екстремального машинного навчання СФД, яка функціонує в режимі кластер-аналізу діагностичних ознак, представимо у вигляді орієнтованого графу відображення множин, що застосовуються в процесі навчання. Вхідний математичний опис СФД подано у вигляді структури

$$\Delta = \langle G, T, \Omega, Z, Y, X; f_1, f_2 \rangle,$$

де G – фактори, що впливають на СФД; T – множина моментів часу формування векторів-реалізацій класів розпізнавання; Ω – простір діагностичних ознак; Z – простір технічних станів системи; Y – вхідна нечітка класифікована навчальна матриця, вектори якої складаються із структурованих за часом зчитування діагностичних ознак із значеннями, заданими в евклідовому просторі; X – класифікована нечітка бінарна навчальна матриця; f_1 – оператор формування вхідної класифікованої навчальної матриці Y ; f_2 – оператор формування вхідної нечітко класифікованої бінарної навчальної матриці X . При цьому декартовий добуток $G \times T \times \Omega \times Z$ утворює універсум випробувань, який є джерелом інформації для формування вхідної навчальної матриці Y .

Категорійну модель інформаційно-екстремального машинного навчання з кластер-аналізом діагностичних ознак показано на рис. 1.

На рис. 1 терм-множина E , елементами якої є обчислені на кожному кроці машинного навчання значення інформаційного критерію (2), згідно з принципом повної композиції є загальною для всіх контурів оптимізації параметрів навчання. Оператор $\tau: E \rightarrow \tilde{\mathfrak{R}}^M$ буде на кожному кроці навчання покриття $\tilde{\mathfrak{R}}^M$, яке відображається оператором θ на нечіткий розподіл класифікованих бінарних векторів діагностичних ознак. Далі оператор $\psi: X \rightarrow I^{S_1}$, де I^{S_1} – множина гіпотез, перевіряє

основну статистичну гіпотезу $\gamma_1 : x_{m,i}^{(j)} \in X_m^o$. Оператор γ визначає множину точнісних характеристик $\mathfrak{S}^{|Q|}$, де $Q = S^2$, а оператор ϕ обчислює множину значень E інформаційного критерію оптимізації, якій є функціоналом від точнісних характеристик. Контур оптимізації контрольних допусків на діагностичні ознаки замикається через терм-множину D , елементами якої є значення системи контрольних допусків. Оператор u регламентує процес машинного навчання.

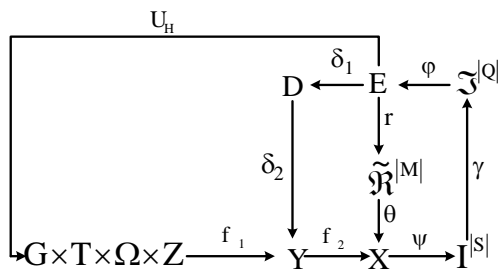


Рис. 1 Категорійна модель інформаційно-екстремального машинного навчання

Показаний на рис. 1 оператор

$$f_1 : G \times T \times \Omega \times Z \rightarrow X^*$$

реалізує безпосередньо процедуру кластер-аналізу діагностичних ознак, задачею якої є формування нечіткої класифікованої вхідної навчальної матриці Y .

На рис. 2 показано категорійну модель інформаційно-екстремального кластер-аналізу діагностичних ознак, який реалізує оператор f_1 .

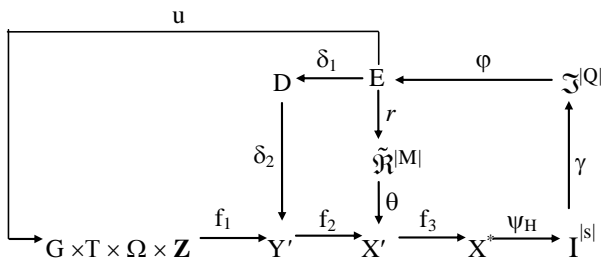


Рис. 2. Категорійна модель інформаційно-екстремального кластер-аналізу

На рис. 2 оператор $f_3 : X' \rightarrow X^*$ формує за процедурою k -середніх навчальну матрицю X^* із двійкових векторів, які знаходяться в контейнерах класів розпізнавання, що відновлюються в процесі машинного навчання в радіальному базисі простору діагностичних ознак. Процес формування вхідної класифікованої нечіткої бінарної матриці X^* здійс-

нюється до моменту забезпечення її мінімального репрезентативного обсягу за умови, що значення інформаційного критерію (2) не менше допустимого в робочій області визначення його функції.

Запропоновані категорійні моделі відбивають притаманні людині перетворення інформації та інформаційні потоки, які мають місце при когнітивних процесах прийняття класифікаційних рішень. Тому вони розглядаються як узагальнені структурні схеми алгоритму інформаційно-екстремального машинного навчання з кластеризацією вхідних даних.

Опис алгоритму машинного навчання з кластеризацією вхідних даних

При функціональному діагностуванні складних машин виникає необхідність аналізу великих обсягів діагностичних ознак при великій потужності алфавіту класів розпізнавання. Для розв'язання цієї задачі використовуються методи кластер-аналізу (автоматичної класифікації) [19, 20]. При цьому переважна більшість відомих методів кластер-аналізу базується на дистанційних критеріях близькості. Тому побудовані при їх реалізації кластери є як правило нечіткими, оскільки на практиці класи розпізнавання перетинаються в просторі ознак розпізнавання. Крім того, функціональна ефективність відомих методів кластер-аналізу суттєво залежить від потужності алфавіту класів розпізнавання, яка в задачах функціонального діагностування є як правило достатньо високою. Тому пропонується використовувати алгоритм кластер-аналізу як допоміжний для автоматичного формування вхідної навчальної матриці з наступною в процесі інформаційно-екстремального машинного навчання її адаптацією до максимальної повної ймовірності прийняття правильних класифікаційних рішень. При цьому для автоматичного формування класифікованої, але нечіткої вхідної навчальної матриці будемо використовувати алгоритм кластер-аналізу на основі процедури k -середніх, яка характеризується високою оперативністю і малою обчислювальною трудомісткістю. При цьому формування класифікованої навчальної матриці $\{x_{m,i}^{(j)} | m = \overline{1, M}, j = \overline{1, n_m}, i = \overline{1, N}\}$ здійснюється за умови $n_m \geq n_{\min}$, тобто кожний відновлений кластер досягає заданий мінімальний репрезентативний обсяг n_{\min} векторів-реалізацій. Це обмеження дозволяє відфільтрувати реалізації, які знаходяться на периферії їх розподілу і тому мають малу ймовірність належності до відновлюваних кластерів.

Як загальний критерій валідності розбиття простору діагностичних ознак будемо розглядати мо-

дифіковану інформаційну міру Кульбака у вигляді [16]

$$J_m^{(k)} = \log_2 \left(\frac{2 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)) + 10^{-p}}{\alpha_m^{(k)}(d) + \beta_m^{(k)}(d) + 10^{-p}} \right) * [1 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d))], \quad (3)$$

де $\alpha_m^{(k)}(d)$ – помилка першого роду прийняття рішення на k -му кроці машинного навчання; $\beta_m^{(k)}(d)$ – помилка другого роду; d – дистанційна міра, яка визначає радіуси гіперсферичних контейнерів, побудованих в радіальному базисі простору Хеммінга; 10^{-p} – достатньо мале число, яке введено для уникнення поділу на нуль.

Нормований критерій (3) має вигляд

$$E_m^{(k)} = \frac{J_m^{(k)}}{J_{MAX}}, \quad (4)$$

де J_{MAX} – максимальне значення критерію, одержане при підстановці у формулу (3) нульових значень помилок $\alpha_m^{(k)}(d)$ і $\beta_m^{(k)}(d)$.

Вектор параметрів машинного навчання (1), які оптимізуються, задамо у вигляді

$$g_m = \langle x_m, d_m, \delta \rangle,$$

де x_1 – статистично усереднений двійковий вектор-реалізація, який визначає геометричний центр контейнера класу розпізнавання X_m^o ; d_m – радіус контейнера класу X_m^o , який відновлюється в радіальному базисі простору діагностичних ознак; δ – параметр, який дорівнює половині симетричного поля контрольних допусків на діагностичні ознаки.

Крім того, задано обмеження на радіуси контейнерів класів розпізнавання:

$$d_m \leq d(x_m \oplus x_c) - 1,$$

де $d(x_m \oplus x_c)$ – кодова відстань між центрами найближчих сусідніх класів X_m^o і X_c^o , і на параметр поля контрольних допусків:

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

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

Таким чином, глибина інформаційно-екстремального машинного навчання у нашому випадку обмежується оптимізацією системи контрольних допусків на діагностичні ознаки.

Згідно з категорійною моделлю (рис. 1) алгоритм інформаційно-екстремального машинного навчання представимо у вигляді ітераційної процедури пошуку глобального максимуму інформаційного критерію (4)

$$\delta^* = \arg \max_{G_\delta} \{ \max_{G_E \cap \{k\}} \bar{E}^{(k)} \}, \quad (5)$$

де $\bar{E}^{(k)}$ – значення обчисленого на k -му кроці навчання усередненого за алфавітом класів розпізнавання інформаційного критерію (4); G_δ – область допустимих значень параметра поля контрольних допусків на діагностичні ознаки; G_δ – робоча (допустима) область визначення критерію оптимізації; $\{k\}$ – множина кроків навчання.

Алгоритм інформаційно-екстремального машинного навчання СФД з автоматичним формуванням вхідної навчальної матриці складається з таких основних етапів:

1) формування класифікованої, але нечіткої вхідної навчальної матриці шляхом кластер-аналізу діагностичних ознак;

2) оптимізація параметрів машинного навчання СФД з метою побудови високодостовірних вирішальних правил;

3) перевірка функціональної ефективності машинного навчання СФД в режимі екзамену.

Алгоритм формування вхідного математичного опису СФД, яка функціонує в режимі інформаційно-екстремального кластер-аналізу діагностичних ознак з оптимізацією системи контрольних допусків згідно з формулою (5) має такі основні етапи:

1) вибір в просторі діагностичних ознак двох векторів-реалізацій, один із яких є найближчим до нульового вектору, а інший – найвіддаленішим;

2) відстань між цими векторами в кодових одиницях поділяється на $M - 1$ рівномірних відрізків, а вектори, вершини яких суміжні відріzkам, разом з векторами, обраними в пункті 1, приймаються як стартові геометричні центри кластерів;

3) обнуління лічильника зміни параметра δ поля контрольних допусків на діагностичні ознаки: $\delta := 0$;

4) $\delta := \delta + 1$;

5) обчислення нижніх $A_{H,i}$ і верхніх $A_{B,i}$ контрольних допусків на всі діагностичні ознаки:

$$A_{H,i} = y_i - \delta \frac{y_i}{100}; \quad A_{B,i} = y_i + \delta \frac{y_i}{100}, \quad (6)$$

де y_i – номінальне значення i -ї діагностичної ознаки;

б) обнуління лічильника кроків збільшення радіусів кластерів: $d := 0$;

7) $d := d + 1$;

8) обчислення на кожному кроці машинного навчання усередненого за кластерами нормованого інформаційного критерію $E(4)$;

9) реалізація процедури k -середніх:

– обчислення усередненого за кількістю реалізацій сформованого кластера двійкового вектору, вершина якого визначає центр ваги кластера;

– перенесення геометричного центра кластера в центр ваги;

10) якщо кількість реалізацій у будь-якому кластері менша за обсяг репрезентативної вибірки n_{\min} , то виконується пункт 7, інакше – пункт 10;

11) якщо усереднене за кластерами нормоване значення інформаційного критерію в робочій області $\bar{E}(d) < 0,5$ то виконується пункт 4, інакше – пункт 11;

12) поточне значення параметра поля контрольних допусків приймається за оптимальне δ^* і обчислюються за формулами (6) оптимальні нижні $A_{н,i}^*$ і верхні $A_{в,i}^*$ контрольні допуски на всі діагностичні ознаки;

13) із векторів-реалізацій, які належать фінальним кластерам формується нечітка класифікована бінарна навчальна матриця X^* ;

14) ЗУПИН.

Формування класифікованої нечіткої навчальної матриці здійснювалося з оптимізацією геометричних параметрів контейнерів класів розпізнавання і системи контрольних допусків на діагностичні ознаки. При цьому контрольні допуски оптимізувалися за паралельною процедурою, при якій вони змінювалися на кожному кроці машинного навчання одночасно для всіх діагностичних ознак. Тому отримані в процесі інформаційно-екстремального кластер-аналізу параметри слід розглядати як квазіоптимальні. У результаті сформована в процесі кластер-аналізу класифікована навчальна матриця X^* принципово не дозволяє побудувати безпомилкові вирішальні правила, оскільки вона є нечіткою через перетин класів розпізнавання в просторі діагностичних ознак. Адаптацію сформованого автоматично вхідного математичного опису до максимальної повної ймовірності правильного прийняття класифікаційних рішень будемо здійснювати шляхом оптимізації параметрів функціонування СФД в процесі інформаційно-екстремального машинного навчання

згідно з категорійною моделлю (рис. 1). Вхідними даними є тривимірний масив навчальної матриці $\{y_{m,i}^{(j)} | i = \overline{1, N}; j = \overline{1, n_m}\}$ та поле $\delta_{н}$ нормованих допусків, яке для всіх ознак розпізнавання є однаковим.

Розглянемо схему алгоритму оптимізації контрольних допусків на діагностичні ознаки в процесі інформаційно-екстремального машинного навчання за процедурою (5):

1) обнуління лічильника класів розпізнавання: $m := 0$;

2) $m := m + 1$;

3) обнуління лічильника зміни параметра δ поля контрольних допусків: $\delta := 0$;

4) $\delta := \delta + 1$;

5) обчислення за формулою (6) нижніх $A_{н,i}$ і верхніх $A_{в,i}$ контрольних допусків на всі діагностичні ознаки:

6) обнуління лічильника кроків зміни радіуса гіперсферичного контейнера: $k := 0$;

7) $k := k + 1$;

8) формується тривимірний масив бінарної навчальної матриці X , елементи якої обчислюються за правилом

$$x_{m,i}^{(j)}[k] = \begin{cases} 1, & \text{якщо } A_{нк,i}[k] < y_{m,i}^{(j)} < A_{вк,i}[k]; \\ 0, & \text{якщо інакше.} \end{cases}$$

9) формування масиву усереднених двійкових векторів-реалізацій $\{x_m\}$, елементи яких визначаються за правилом

$$x_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n} \sum_{j=1}^n x_{m,i}^{(j)} > \rho_m; \\ 0, & \text{if else,} \end{cases}$$

де ρ_m – рівень квантування координат двійкового вектору x_m , який за замовчуванням дорівнює 0,5.

10) розбиття множини векторів $\{x_m\}$ на пари найближчих “сусідів” $\mathfrak{R}_m^2 = \langle x_m, x_c \rangle$, де x_c – усереднений вектор сусіднього класу X_c^0 ;

11) обчислюється інформаційний критерій (4);

12) якщо $k \leq N$, то виконується пункт 7, інакше – пункт 13;

13) якщо $\delta < \delta_{н}$, то виконується пункт 4, інакше – пункт 14;

14) визначається максимальне значення критерію (4) в робочій області визначення його функції, де перша і друга достовірності більше 0,5;

15) якщо виконується умова $m < M - 1$, то реалізується пункт 2, інакше – пункт 16;

16) визначається глобальний максимум усередненого інформаційного критерію \bar{E}^* в робочій області визначення його функції;

17) визначаються оптимальні значення параметра δ^* і відповідних нижніх $A_{н,i}^*$ і верхніх $A_{в,i}^*$ контрольних допусків на всі діагностичні ознаки ;

18) визначаються радіуси оптимальних контейнерів класів розпізнавання:

$$d_m^* = \arg \bar{E}^* ;$$

19) будуються продукційні вирішальні правила, які в рамках ІЕІ-технології для гіперсферичних контейнерів мають такий предикатний вираз

$$(\forall X_m^o \in \mathfrak{R}^{[M]})(\forall x^{(j)} \in \mathfrak{R}^{[M]})[\text{if } (\mu_m > 0) \& (\mu_m = \max\{\mu_m\}) \text{ then } x^{(j)} \in X_m^o \text{ else } x^{(j)} \notin X_m^o], \quad (7)$$

де $x^{(j)}$ – вектор, що розпізнається; μ_m – функція, яка визначає належності вектору $x^{(j)}$ контейнеру класу розпізнавання X_m^o .

У виразі (7) функція належності для гіперсферичного контейнера класу розпізнавання X_m^o визначається за формулою

$$\mu_m = 1 - \frac{d(x_m^* \oplus x^{(j)})}{d_m^*},$$

де x_m^* , d_m^* – отримані в процесі машинного навчання параметри оптимального контейнера класу X_m^o – усереднена двійкова реалізація і радіус гіперсферичного контейнера відповідно.

Таким чином, інформаційно-екстремальний кластер-аналіз відбувається шляхом поєднання алгоритму k-середніх, який дозволяє сформувати вхідну в загальному випадку нечітку класифіковану навчальну матрицю, з інформаційно-екстремальним алгоритмом машинного навчання, який дозволяє побудувати високо достовірні вирішальні правила.

Приклад реалізації алгоритму

Реалізація розглянутого вище алгоритму здійснювалася за некласифікованою навчальною матрицею, отриманою за результатами моніторингу технологічного процесу згідно тахограми на всіх циклах роботи багатоканатної ШПМ. Матриця була сформована за архівними даними, наданими підприємством «УЛІС Системс», яке займається модернізацією системи керування ШПМ в «ДТЕК Павлоградвугілля» (м. Павлоград, Україна). Для наочності некласифікована навчальна матриця складалася із

векторів-реалізацій трьох класів розпізнавання: клас X_1^o характеризував функціональний стан електроприводу «Норма», клас X_2^o – стан підвищеної температури підшипників електродвигуна і клас X_3^o – стан тормозних колодок «Більше норми». Кожний вектор-реалізація відповідного класу розпізнавання складався із 62 структурованих кількісних і категорійних діагностичних ознак, які характеризували електричні, температурні та механічні характеристики вузлів ШПМ. У початковій вибірковій множині кількість векторів-реалізацій для кожного класу була однаковою і дорівнював 60.

Формування вхідної класифікованої нечіткої робочої бінарної матриці X^* здійснювалося шляхом інформаційно-екстремального кластер-аналізу за алгоритмом, розробленим згідно з категорійною моделлю, показаною на рис. 2. На рис. 3 показано графік зміни усередненого для заданого алфавіту класів розпізнавання інформаційного критерію (4) від кількості кроків кластеризації діагностичних ознак.

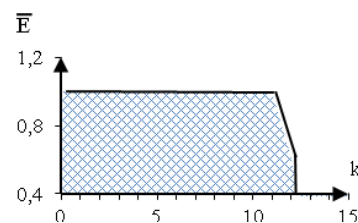


Рис. 3 Графік залежності інформаційного критерію від кроків кластеризації

Аналіз рис. 3 показує, що до 10-го кроку кластеризації значення усередненого нормованого інформаційного критерію (4) було максимальним, оскільки кластери не перетиналися в просторі діагностичних ознак. Далі при збільшенні радіусів контейнерів класів розпізнавання почали перетинатися, що призвело до зменшення значення критерію до значення $\bar{E} = 0,62$ через появу помилок першого та другого роду. На 12-му кроці процес кластеризації завершився, оскільки була виконана умова зупини: $\bar{E} > 0,5$ і навчальна вибірка досягнула мінімального репрезентативного обсягу, який визначався за методом, запропонованим в праці [19], і дорівнював $n_{min} = 56$.

Для підвищення повної ймовірності правильної класифікації було реалізовано згідно з категорійною моделлю, показаною на рис. 1, вище наведений алгоритм інформаційно-екстремального машинного навчання за автоматично отриманою вхідною навчальною матрицею. При цьому попередньо було штучно збільшено середню міжкласову відстань шляхом надлишкового кодування двійкових векто-

рів-реалізацій робочої навчальної матриці X за циклічним методом БЧХ. У результаті кодування мінімальна кодова відстань Хеммінга для усереднених двійкових векторів-реалізацій класів розпізнавання дорівнювала $d_{\min} = 7$.

На рис. 4 показано графік залежності усередненого інформаційного критерію (4) від параметра машинного навчання δ , отриманий за результатами машинного навчання з паралельною оптимізацією контрольних допусків на діагностичні ознаки.

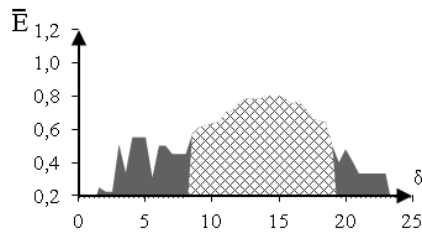


Рис. 4. Графік залежності інформаційного критерію від параметра поля контрольних допусків

На рис. 4 подвійною штриховкою позначено робочу область визначення інформаційного критерію, в якій перша і друга достовірності перевершують відповідно помилки першого та другого роду. Аналіз рис.4 показує, що оптимальне значення параметра поля контрольних допусків дорівнює $\delta^* = 14$ (у відносних одиницях) при максимальному значенні інформаційного критерію $\bar{E}^* = 0,80$.

Таким чином, в процесі машинного навчання СФД за сформованою за результатами кластер-аналізу навчальною матрицею Y відбулося збільшення значення інформаційного критерію оптимізації у порівнянні з його значенням при зупині кластеризації. Цей факт пояснюється двома причинами. По-перше, у навчальну матрицю Y не входять вектори ознак, найбільш віддалені від центрів відповідних кластерів, які могли належати іншим класам розпізнавання і навіть таким, що не входять в заданий алфавіт. По-друге, побудовані за результатами кластер-аналізу гіперсферичні кластери мали однакові радіуси, а отримані в процесі машинного навчання оптимальні радіуси більш точно апроксимують розподіли відповідних кластерів в просторі ознак, що підтверджується збільшенням інформаційного критерію оптимізації.

На рис. 5 показано графіки залежності інформаційного критерію оптимізації (4) від радіусів контейнерів відповідних класів розпізнавання, отримані при оптимальному параметрі $\delta^* = 14$ поля контрольних допусків на діагностичні ознаки.

Аналіз рис. 5 показує, що для класу розпізнавання X_1^0 оптимальне значення радіусу його кон-

тейнера дорівнює $d_1^* = 16$ (тут і далі в кодових одиницях) при максимальному значенні критерію $E_1^* = 0,78$, для класу X_2^0 – $d_2^* = 13$ при значенні критерію $E_2^* = 0,80$ і для класу X_3^0 – $d_3^* = 15$ при значенні критерію $E_3^* = 0,82$.

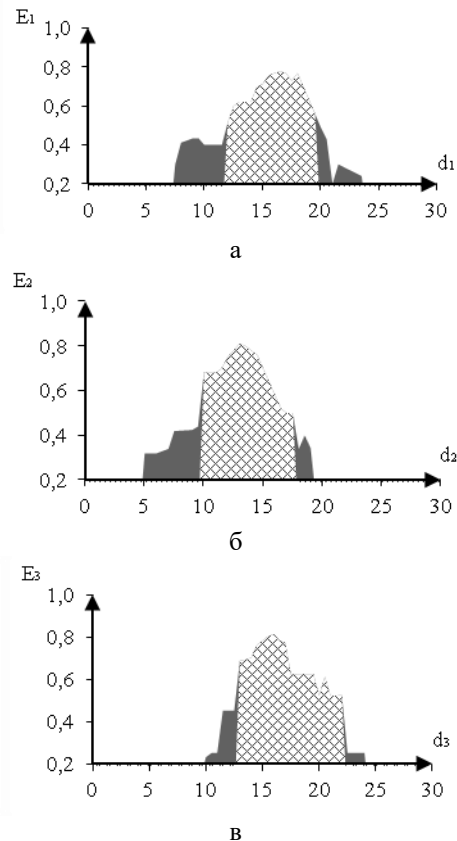


Рис. 5. Графіки залежності інформаційного критерію від радіусів контейнерів: а – клас X_1^0 ; б – клас X_2^0 ; в – клас X_3^0

За отриманими в процесі інформаційно-екстремального кластер-аналізу оптимальними геометричними параметрами контейнерів класів розпізнавання були побудовані продукційні вирішальні правила (7), які при функціонуванні СФД в режимі екзамену забезпечили для класу розпізнавання X_1^0 повну ймовірність правильних діагностичних рішень $P_{t,1} = 0,85$, для класу X_2^0 – $P_{t,2} = 0,86$ і для класу X_3^0 – $P_{t,3} = 0,83$, що є достатньо високими показниками. Разом з тим необхідно констатувати, що побудовані вирішальні правила не є безпомилковими за навчальною матрицею через наявність перетину класів розпізнавання в просторі діагностичних ознак. Тому для підвищення функціональної ефективності СФД необхідно збільшення глибини машинного навчання шляхом оптимізації інших пара-

метрів функціонування системи, що є предметом подальших досліджень.

Висновки

1. Запропоновано новий метод інформаційно-екстремального машинного навчання з кластер-аналізом діагностичних ознак, який дозволяє автоматично формувати вхідну нечітко класифіковану навчальну матрицю з подальшою її дефазифікацією шляхом оптимізації геометричних параметрів класів розпізнавання.

2. Особливість запропонованого методу полягає у використанні процедури k-середніх як допоміжної, призначенням якої є автоматичне формування вхідного математичного опису СФД. Тобто алгоритм класичного агломеративного кластер-аналізу є вкладеним в алгоритм інформаційно-екстремального машинного навчання.

3. Реалізація запропонованого методу на прикладі інформаційно-екстремального синтезу СФД багатоканатної ШПМ підтвердив достатньо високу достовірність діагностичних рішень. При цьому відкривається шлях для подальшого підвищення функціональної ефективності СФД через оптимізацію додаткових параметрів машинного навчання.

Література

1. Попов, Ю. В. Модернизация тормозной системы шахтной подъемной установки [Текст] / Ю. В. Попов, С. В. Кудрявцев, С. В. Степанов // Горный информационно-аналитический бюллетень (научно-технический журнал). – 2015. – № 9. – С. 195-197.

2. Герике, Б. Л. Динамическая диагностика машинных агрегатов горного оборудования [Текст] / П. Б. Герике, В. Н. Шахманов // Горный информационно-аналитический бюллетень (научно-технический журнал). – 2011. – № 55. – С. 80-89.

3. Sun, W. Fault diagnosis of rolling bearing based on wavelet transform and envelope spectrum correlation [Text] / W. Sun, G. A. Yang, Q. Chen, A. Palazoglu, K. Feng // Journal of Vibration and Control. – 2013. – Vol. 19, No. 6. – P. 924-941. doi: 10.1177/1077546311435348

4. Литвин, О. М. Про один підхід до тестування нових методів розв'язання нестационарної задачі теплопровідності [Текст] / О. М. Литвин, Л. С. Лобанова, Г. В. Залужна // Штучний інтелект. – 2012. – № 1. – С. 219-228.

5. Dong, S. Rotating Machine Fault Diagnosis Based on Locality Preserving Projection and Back Propagation Neural Network – Support Vector Machine Model [Text] / X. Xu, J. Liu, Z. Gao // Measurement and Control. – 2015. – Vol. 48, No. 7. – P. 211-212.

6. Wang, D. An enhanced Kurtogram method for fault diagnosis of rolling element bearings [Text] / D. Wang, P. W. Tse, K. L. Tsui // Mech. Syst. Signal Process. – 2013. – Vol. 35, No. 1. – P. 176-199.

7. Feng, Z. Joint envelope and frequency order spectrum analysis based on iterative generalized demodulation for planetary gearbox fault diagnosis

under nonstationary conditions [Text] / Z. Feng, X. Chen, M. Liang // Mech. Syst. Signal Process. – 2016. – Vol. 76. – P. 242-264.

8. Andrzej, O. Modern Methods Control and Diagnostics of Hoisting-Machines with Application of Artificial Intelligence Methods [Text] / O. Andrzej, S. Zygmunt // Archives of Mining Sciences. – 2010. – Vol. 1. – P. 217 – 231.

9. Сидоренко, В. Н. Мониторинг мощных электроприводов на основе технологии Data Mining [Текст] / В. Н. Сидоренко, А. П. Черный // Электротехнічні та комп'ютерні системи. – 2011. – № 3. – С. 414-416.

10. Jayaswal, P. Development of EBP-Artificial Neural Network Expert System for Rolling Element Bearing Fault Diagnosis [Text] / P. Jayaswal, S. Verma, A. Wadhvani // Journal of Vibration and Control. – 2011. – Vol. 17, No. 8. – P. 1131-1148, doi:10.1177/1077546310361858.

11. Samanta, B. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features [Text] / B. Samanta, K. R. Al-Balushi // Mechanical systems and signal processing. – 2003. – Vol. 17, No. 2. – P. 317-328.

12. Subbotin, S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition [Text] / S. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2013. – Vol. 22, No. 2. – P. 97-103. DOI: 10.3103/s1060992x13020082.

13. Dovbysh, A. S. Information Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes [Text] / A. S. Dovbysh, M. M. Budnyk, V. V. Moskalenko // Journal of automation and information sciences. – 2012. – Vol. 44, No. 10. – P. 35-44. DOI: 10.1615/JAutomatInfScien.v44.i10.30.

14. Dovbysh, A. Information-extreme machine learning of the control system over the power unit of a terminal power main line [Text] / A. Dovbysh, D. Velikodnyi, I. Shelekhov, M. Bibyk // Eastern-European Journal of Enterprise Technologies. – 2017. – Vol. 5/4, No. 89. – P. 17-24. DOI: 10.15587/1729-4061.2017.112121.

15. Довбиш, А. С. Інформаційно-екстремальний алгоритм навчання системи функціонального контролю електроприводу шахтної підйомної машини [Текст] / А. С. Довбиш, В. І. Зимовець, З. С. Козлов // Радиозлектроника и информатика. – 2017. – No. 4. – С. 58-63.

16. Щербакова, Г. Ю. Оценка параметров DN-распределения с помощью помехоустойчивой кластеризации для автоматизированных систем технической диагностики [Текст] / Г. Ю. Щербакова // Штучний інтелект. – 2010. – No. 4. – С. 491-496.

17. Schonlau, M. Visualizing non-hierarchical and hierarchical cluster analyses with clustergrams [Text] / M. Schonlau // Computational Statistics. – 2004. – Vol. 19, No. 1. – P. 95-111.

References

1. Popov, Yu. V., Kudryavtsev, S. V., Stepanov, S. V. Modernizatsiya tormoznoy sistemy shakhtnoy pod'emnoy ustanovki [Modernization of the brake system of mine lifting installation] Gorniy informacionno-analiticheskiy bjulleten' (nauchno-

tehnicheskij zhurnal). – *Mining information and analytical bulletin (scientific and technical journal)*, 2015, vol. 9, pp. 195–197.

2. Gerike, B. L., Gerike, P. B., Shahmanov, V. N. Dinamicheskaja diagnostika mashinnyh agregatov gornogo oborudovaniya [Dynamic diagnostics of machine units of mining equipment]. *Gornyj informacionno-analiticheskij bjulleten' (nauchno-tehnicheskij zhurnal) – Mining information and analytical bulletin (scientific and technical journal)*, 2011, no. 5, pp. 80–89.

3. Sun, W., Yang, G. A., Chen, Q., Palazoglu, A., Feng, K. Fault diagnosis of rolling bearing based on wavelet transform and envelope spectrum correlation. *Journal of Vibration and Control*, 2013, vol. 19, no. 6, pp. 924–941. DOI: 10.1177/1077546311435348.

4. Lytvyn, O. M., Lobanova, L. S., Zaluzhna, H. V. Pro odyn pidkhid do testuvannya novykh metodiv rozv'yazannya nestatsionarnoyi zadachi teploprovodnosti [About one week before testing new methods for developing unsteady tasks and heating problems]. *Shtuchnyy intelekt – Artificial intelligence*, 2012, no. 1, pp. 219–228.

5. Dong, S., Xu, X., Liu, J., Gao, Z. Rotating Machine Fault Diagnosis Based on Locality Preserving Projection and Back Propagation Neural Network–Support Vector Machine Model. *Measurement and Control*, 2015, vol. 48, no. 7, pp. 211–216.

6. Wang, D., Tse, P. W., Tsui, K. L. An enhanced Kurtogram method for fault diagnosis of rolling element bearings. *Mech. Syst. Signal Process*, 2013, vol. 35, no. 1, pp. 176–199.

7. Feng, Z., Chen, X., Liang, M. Joint envelope and frequency order spectrum analysis based on iterative generalized demodulation for planetary gearbox fault diagnosis under nonstationary conditions. *Mech. Syst. Signal Process.*, 2016, vol. 76, pp. 242–264.

8. Andrzej, O., Zygmunt, S. Modern Methods Control and Diagnostics of Hoisting-Machines with Application of Artificial Intelligence Methods. *Archives of Mining Sciences*, 2010, no. 1, pp. 217–231.

9. Sidorenko, V. N., Chernyi, A. P. Monitoring of powerful electric drives on the basis of Data Mining technology]. *Elektrotehnichni ta komp'yuterni sistemi –*

Electrical and computer systems, 2011, no. 3, pp. 313–316.

10. Jayaswal, P., Verma, S. N., Wadhvani, A. K. Development of EBP-Artificial neural network expert system for rolling element bearing fault diagnosis. *Journal of Vibration and Control*, 2011, vol. 17, no. 8, pp. 1131–1148.

11. Samanta, B., Al-Balushi, K. R. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mechanical systems and signal processing*, 2003, vol. 17, no. 2, pp. 317–328.

12. Subbotin, S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition. *Optical Memory and Neural Networks (Information Optics)*, 2013, vol. 22, no. 2, pp. 97–103. DOI: 10.3103/s1060992x13020082.

13. Dovbysh, A. S., Budnyk, M. M., Moskalenko, V. V. Information Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes. *Journal of automation and information sciences*, 2012, vol. 44, no. 10, pp. 35–44. DOI: 10.1615/JAutomatInfScien.v44.i10.30.

14. Dovbysh, A., Velikodnyi, D., Shelehov, I., Bibyk, M. Information-extreme machine learning of the control system over the power unit of a thermal power main line. *Eastern-European Journal of Enterprise Technologies*, 2017, vol. 5/4, no. 89, pp. 17–24. DOI: 10.15587/1729-4061.2017.112121.

15. Dovbysh, A. S., Zimovets, V. I., Kozlov, Z. S. Information extreme learning algorithm system diagnostics about mine lifting machine. *Radioelektronika i informatika*, 2017, no. 4, pp. 58–63.

16. Scherbakova, G. Yu. DN-distribution Parameters Evaluation with Noise Stability Clustering for the Automated Systems of Technical Diagnostics. *Shtuchnyy intelekt – Artificial Intelligence*, 2010, vol. 4, pp. 491–496.

17. Schonlau M. Visualizing non-hierarchical and hierarchical cluster analyses with clustergrams. *Computational Statistics*, 2004, vol. 19, no 1, pp 95–111.

Поступила в редакцію 21.10.2019, рассмотрена на редколлегии 10.12.2019

ИНФОРМАЦИОННО-ЭКСТРЕМАЛЬНЫЙ КЛАСТЕР-АНАЛИЗ ВХОДНЫХ ДАННЫХ ПРИ ФУНКЦИОНАЛЬНОМ ДИАГНОСТИРОВАНИИ

В. И. Зимовец, А. С. Приходченко, Н. И. Мироненко

Целью исследования является повышение функциональной эффективности машинного обучения системы функционального диагностирования (СФД) многоканатной шахтной подъемной машиной (ШПМ) путем кластер-анализа диагностических признаков. Для достижения цели необходимо было решить следующие **задачи**: формализовать постановку задачи информационного синтеза обучающейся СФД, функционирующей в режиме кластер-анализа диагностических признаков; предложить категориальную модель и на ее основе разработать алгоритм информационно-экстремального кластер-анализа диагностических признаков; в процессе информационно-экстремального машинного обучения СФД осуществить фазификацию входных нечетких данных путем оптимизации геометрических параметров гиперсферических контейнеров классов распознавания, характеризующих возможные технические состояния объекта диагностирования; разработать и программно реализовать алгоритм информационно-экстремального кластер-анализа на примере информационного синтеза СФД многоканатной ШПМ. **Объектом** исследования являются процессы информационного синтеза обучающейся СФД, интегрированной в автоматизированную систему управления многоканатной ШПМ. **Предметом** исследования являются категориальные модели, алгоритм информационно-экстремального машинного обучения СФД, функционирующей в режиме кластер-анализа диагностических признаков и построенные в процессе машинного обучения решающие правила. **Методы исследования** базируются на идеях и методах информационно-экстремальной интеллектуальной технологии анализа данных, теоретико-информационном подходе к оценке функциональной эффективности машинного обучения и на

геометрическом подходе теории распознавания образов. Получены следующие **результаты**: предложена категорийная модель, и на ее основе разработан и программно реализован алгоритм информационно-экстремального машинного обучения СФД многоканатной ШПМ, позволяющий автоматически формировать входную классифицированную нечеткую обучающую матрицу, что существенно уменьшает временные и материальные расходы при создании входного математического описания. Полученный результат был достигнут путем кластер-анализа структурированных векторов диагностических признаков, полученных по архивным данным для трех классов распознавания, с использованием процедуры k-средних. В качестве критерия оптимизации параметров машинного обучения рассматривалась модифицированная мера Кульбака в виде функционала от точностных характеристик диагностических решений и дистанционных критериев близости классов распознавания. По полученным в процессе машинного обучения оптимальным геометрическим параметрам контейнеров классов распознавания построены решающие правила, позволяющие классифицировать векторы диагностических признаков классов распознавания с достаточно высокой полной вероятностью принятия правильных диагностических решений. **Выводы.** Научная новизна полученных результатов заключается в разработке нового метода информационного синтеза СФД многоканатной ШПМ, функционирующей в режиме кластер-анализа, что позволило автоматически формировать входную классифицированную нечеткую обучающую матрицу с последующей ее дефазификацией в процессе информационно-экстремального машинного обучения системы.

Ключевые слова: информационно-экстремальная интеллектуальная технология; машинное обучение; кластер-анализ; система функционального диагностирования; учебная матрица; информационный критерий оптимизации; многоканатная шахтная подъемная машина.

INFORMATION-EXTREMAL CLUSTER-ANALYSIS OF INPUT DATA IN FUNCTIONAL DIAGNOSIS

V. I. Zimovets, O. S. Prykhodchenko, M. I. Myronenko

The study aims to increase the functional efficiency of machine learning of the functional diagnosis system of a multi-rope shaft hoist through cluster analysis of diagnostic features. To achieve the goal, it was necessary to solve the following tasks: formalize the formulation of the task of information synthesis, capable of learning a functional diagnosis system, which operates in the cluster-analysis mode of diagnostic signs; to propose a categorical model and, on its basis, to develop an algorithm for information-extreme cluster analysis of diagnostic signs in the process of information-extreme machine learning of a functional diagnostic system; carry out fuzzification of input fuzzy data by optimizing the geometric parameters of hyperspherical containers of recognition classes that characterize the possible technical conditions of the diagnostic object; to develop an algorithm and implement it on the example of information synthesis of the functional diagnostics system of a multi-rope mine hoisting machine. The object of the study is the processes of information synthesis of a functional diagnostic system capable of learning, integrated into the automated control system of a multi-rope mine hoisting machine. The subject of the study is categorical models, an information-extremal machine learning algorithm of a functional diagnostic system that operates in the cluster analysis model of diagnostic signs and constructs decision rules. The research methods are based on the ideas and methods of information-extreme intellectual data analysis technology, a theoretical-informational approach to assessing the functional effectiveness of machine learning and on the geometric approach of pattern recognition theory. As a result, the following results were obtained: a categorical model was proposed, and on its basis, an algorithm for information-extremal machine learning of the functional diagnostics system for a multi-rope mine hoist was developed and implemented, which allows you to automatically generate an input classified fuzzy training matrix, which significantly reduces time and material costs when creating incoming mathematical description. The obtained result was achieved by cluster analysis of structured vectors of diagnostic signs obtained from archival data for three recognition classes using the k-means procedure. As a criterion for optimizing machine learning parameters, we considered a modified Kullback measure in the form of a functional on the exact characteristics of diagnostic solutions and distance criteria for the proximity of recognition classes. Based on the optimal geometric parameters of the containers of recognition classes obtained during machine learning, decisive rules were constructed that allowed us to classify the vectors of diagnostic features of recognition classes with a rather high total probability of making the correct diagnostic decisions. Conclusions. The scientific novelty of the results obtained consists in the development of a new method for the information synthesis of the functional diagnostics system of a multi-rope mine hoisting machine, which operates in the cluster analysis model, which made it possible to automatically form an input classified fuzzy training matrix with its subsequent dephasification in the process of information-extreme machine learning system.

Keywords: information-extreme algorithm; training matrix; diagnostics; multichannel shaft lifting machine; clusterization; optimization; forecasting

Зимовець Вікторія Ігорівна – аспірант, Сумський державний університет, Суми, Україна.

Приходченко Олександр Сергійович – аспірант, Сумський державний університет, Суми, Україна.

Мироненко Микита Ігорович – аспірант, Сумський державний університет, Суми, Україна

Zimovets Victoria Ihorivna – PhD Student of Sumy State University, Sumy, Ukraine, e-mail: zc.vika@gmail.com, ORCID Author ID: 0000-0003-4604-4087, ResearcherID: H-5410-2017

<https://scholar.google.com.ua/citations?user=QasOWGQAAAAJ>

Prykhodchenko Oleksandr Sergiyovich – PhD student, Sumy State University, Sumy, Ukraine, e-mail: info.g@ukr.net.

Myronenko Mykyta Igorovich – PhD Student, Sumy State University, Sumy, Ukraine, e-mail: nikitam1996@ukr.net, ORCID Author ID: 0000-0001-5005-1672.