УДК 519.832.3::519.876.2

В. В. Романюк, канд. техн. наук, доц.

МОДЕЛЬ ОПТИМІЗАЦІЇ БАГАТОЕТАПНОГО ПРОЦЕСУ НЕЙРОМЕРЕЖЕВОГО НАВЧАННЯ З ФУНКЦІЄЮ ОБМЕЖЕННЯ ВТРАТ У ЗАДАЧІ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

На прикладі нейромережевого дослідження задачі розпізнавання кольорових зображень розглянуто проблему якнайшвидшого виведення системи обробки інформації на нормальний режим функціонування за умов заздалегідь невідомого розподілу пріоритетів її характеристик. Формалізовано відповідну задачу прийняття рішень, у якій альтернативами є методи навчання нейронної мережі, а її характеристики сприймаються як наслідки. Виходячи з відсутності імовірнісних мір на множині невизначених характеристик системи обробки інформації, за цією задачею прийняття рішень розв'язано матричну гру, у якій оптимальна стратегія другого гравця використовується для комбінованого вибору методів навчання нейронної мережі, незалежно від вагомості її оцінюваних характеристик. Продемонстровано перевагу моделі оптимізації багатоетапного процесу нейромережевого навчання з функцією обмеження втрат перед класичною моделлю реалізації оптимальної стратегії, а також методом навчання, за якого досягається мінімальна відстань до точки утопії.

Проблема якнайшвидшого виведення системи обробки інформації на нормальний режим функціонування

Керована система обробки інформації (система керування базою даних, експертна система, нейронна мережа, статистичний модуль) має в реальному потоці часу забезпечувати швидку та якісну обробку даних. Для цього вона під час перехідного процесу від ініціалізації до початку своєї сталої роботи має бути виведена на нормальний режим функціонування. Якщо проектувальник володіє інформацією про будь-яку імовірнісну міру на множині невизначених властивостей або атрибутів системи обробки інформації, то отримуємо задачу прийняття рішень, де вироблення множини оптимальних стратегій відбувається з використанням відомих критеріїв формування множини оптимальних альтернатив [1]. Коли ж такої інформації у нього немає, то ніяким чином звести модель конфліктного явища, породжуваного нестохастичною невизначеністю, до задачі прийняття рішень, оминувши матричну гру, не вдасться: проектувальник має мінімізувати втрати від незнання дійсних значень атрибутів системи та співвідношень між ними для усіх варіантів, з найгіршими включно [1, 2]. У випадку нейромережевої системи обробки інформації, для якої нормальний режим функціонування настає після закінчення процесу її навчання, проектувальник може керувати надзвичайно великою кількістю параметрів, що впливають на тривалість цього процесу, однак у першу чергу необхідно розглядати більш глобальні. Такими параметрами є методи навчання нейронної мережі.

Методи навчання нейромережевої системи обробки інформації

Методи навчання нейромережевої системи обробки інформації залежать від класу розв'язуваних нею задач [3]. Якщо розглядати проблему розпізнавання зображень, то для її вирішення зазвичай проектують нейронну мережу з алгоритмом зворотного розповсюдження похибки (АЗРП), який дозволяє досягти максимальної продуктивності багатошарового персептрона [4, 5] за прийнятну тривалість навчання. При цьому АЗРП як модифікація класичного алгоритму градієнтного спуску має низку програмних реалізацій (методів), кожна з яких спрямована на удосконалення [3, 6] одного-двох виокремлених параметрів АЗРП (параметри швидкості навчання, коефіцієнт інерційності градієнтного спуску, ресурси пам'яті, параметри пакетного режиму). В інструмент Neural Network Toolbox 7.0.3 середовища MATLAB 7.14 вбудовано 13 таких методів. І за кожного з них після процесу навчання нейронна мережа набуває певних характеристик (атрибутів чи властивостей), середньостатистичні кількісні оцінки яких можуть визначатись на фіксованому класі розв'язуваних нейромережевою системою задач. Зрозуміло, що вибір конкретного методу навчання залежить від цих оцінок, за якими нейронна мережа має швидко навчатись й ефективно функціонувати. Однак зазвичай на кінцевій стадії процесу здійснення цього вибору знаходяться щонайменше два взаємовиключні методи навчання: один швидко навчає, але надає системі обробки інформації гіршої функціональності, а за іншого — висока функціональність системи набувається повільніше [3, 7].

Мета і завдання статті

Виходячи з принципової неможливості виділення єдиного методу навчання нейронної мережі за її середньостатистичними характеристиками (за критерієм максимізації показників ефективності функціонування) у більшості задач розпізнавання зображень, необхідно запропонувати такий критерій, який дозволив би здійснювати цей вибір однозначно та незалежно від вагомості цих характеристик. Для цього слід формалізувати відповідну задачу прийняття рішень (задачу вибору), з якої виключити строго доміновані альтернативи, і визначити за запропонованим критерієм коефіцієнти згортання для багатокритеріальної задачі оптимізації. Також на прикладі нейронної MATLAB-мережі для розпізнавання 40 кольорових зображень порівняємо результат такого згортання з відомими показниками і надамо супутні рекомендації.

Нейромережеве дослідження задачі розпізнавання кольорових зображень

Нехай система обробки інформації з F характеристиками $\left\{\vartheta_i\right\}_{i=1}^{F}$ може бути виведена на нормальний режим функціонування за одним із S_0 методів $\left\{m_j\right\}_{j=1}^{S_0}$, причому деякі з цих характеристик оцінюються під час перехідного процесу (зокрема, його тривалість) від ініціалізації системи до початку її сталої роботи. Позначимо через u_{ij} значення i-ї характеристики системи обробки інформації, якої вона набуває за j-м методом її виведення на нормальний режим функціонування. Для однозначності вважатимемо, що всі значення $\left\{\left\{u_{ij}\right\}_{i=1}^{F}\right\}_{j=1}^{S_0}$ характеризують ті властивості системи, які мають бути мінімізовані (наприклад, тривалість навчання нейронної мережі, рівень помилок у її роботі, об'єм споживаних нею ресурсів). Після нормування

$$w_{ij} = \frac{u_{ij}}{\max\limits_{k=\overline{1,\, S_0}} u_{ik}}, \text{ якщо } i=\overline{1,\, F} \text{ та } j=\overline{1,\, S_0} \tag{1}$$

отримуємо $S_0 \times F$-матрицю $\mathbf{W}_0^{\mathrm{T}} = \left(\left[w_{ij}\right]_{F\times S_0}\right)^{\mathrm{T}}$ прийняття рішень з елементами (1), де дані про імовірнісний розподіл на множині $\left\{\vartheta_i\right\}_{i=1}^{F}$ відсутні. Без втрати загальності можна вважати, що рядки матриці $\mathbf{W}_0^{\mathrm{T}} = \left(\left[w_{ij}\right]_{F\times S_0}\right)^{\mathrm{T}}$ переставлені (попередньо впорядковані) так, що

$$u_{ik} < u_{il} \text{ або } w_{ik} < w_{il}, \text{ якщо } k=\overline{1,\, S}\,;\; l=\overline{S+1,\, S_0} \text{ та } i=\overline{1,\, F}\,. \tag{2}$$

Оскільки альтернативи $\left\{m_j\right\}_{j=S+1}^{S_0}$ є строго домінованими, то задачу прийняття рішень з матрицею $\mathbf{W}_0^{\mathrm{T}} = \left(\left[w_{ij}\right]_{F\times S_0}\right)^{\mathrm{T}}$ можна редукувати до задачі прийняття рішень з матрицею $\mathbf{W}^{\mathrm{T}} = \left(\left[w_{ij}\right]_{F\times S}\right)^{\mathrm{T}}$. Зауважимо, що тоді множина $\left\{m_j\right\}_{j=1}^{S}$ є множиною Парето [1], утворюваною в результаті ігнорування строго домінованих стратегій проектувальника.

Завданням проектувальника є вибір на множині $\left\{m_j\right\}_{j=1}^{S}$ методу виведення системи на нормальний режим функціонування, який дозволив би мінімізувати значення (1) за якнайкоротший проміжок часу [6, 8]. Однак очевидно, що F-критеріальна задача

$$\left\langle \arg \min_{j=\overline{1,\, S}} w_{ij} \right\rangle_{i=1}^{F} \tag{3}$$

відносно оптимального методу $m_{j_{\text{opt}}}$ для $j_{\text{opt}} \in \arg \min\limits_{j=\overline{1,\, S}} w_{ij}$ не має розв'язку [1]. Звісно, опукле згортання F критеріїв в (3) як

$$\arg \min_{j=\overline{1,\, S}} \sum_{i=1}^{F} \lambda_i w_{ij} \tag{4}$$

за множиною коефіцієнтів

$$\left\{\lambda_i : \lambda_i \geqslant 0 \ \forall \ i = \overline{1, F},\ \sum_{i=1}^{F} \lambda_i = 1\right\} \tag{5}$$

дає однозначний розв'язок задачі (3), проте множина коефіцієнтів (5) для згортки в (4) є невідомою через те, що важливість (пріоритетність) характеристик $\left\{\vartheta_i\right\}_{i=1}^{F}$ для функціонування системи обробки інформації невизначена. З тієї самої причини задача

$$\arg \min_{j=\overline{1,\, S}} \sqrt{\sum_{i=1}^{F} \left(\lambda_i w_{ij}\right)^2} \tag{6}$$

з відстанню до початку координат в $\mathbb{R}^F$ як ідеальної і недосяжної на практиці точки (точки утопії) не розв'язується. Саме тому моделлю розв'язування задачі (3) є матрична гра з матрицею $\mathbf{W} = \left[w_{ij}\right]_{F \times S}$, яка виявляється фундаментальною при побудові ігрових моделей усунення нестохастичних невизначеностей й у складніших системах обробки інформації [2, 8].

У розв'язку матричної $\left[w_{ij}\right]_{F \times S}$-гри проектувальника цікавить тільки оптимальна стратегія другого гравця [8]:

$$\begin{aligned} \breve{\mathbf{Q}} &= \left[\breve{q}_1 \quad \breve{q}_2 \quad \cdots \quad \breve{q}_{S-1} \quad \breve{q}_S\right] \in \breve{\mathfrak{Q}} \subset \mathfrak{Q} = \\ &= \left\{\mathbf{Q} = \left[q_1 \quad q_2 \quad \cdots \quad q_{S-1} \quad q_S\right] \in \mathbb{R}^S : \sum_{j=1}^{S} q_j = 1,\ q_j \in [0;\,1] \ \forall \ j = \overline{1, S}\right\} \end{aligned} \tag{7}$$

із множини

$$\breve{\mathfrak{Q}} = \arg \min_{\mathbf{Q} \in \mathfrak{Q}} \max_{\mathbf{\Lambda} \in \mathfrak{L}} \left(\mathbf{\Lambda} \cdot \mathbf{W} \cdot \mathbf{Q}^{\mathrm{T}}\right) \subset \mathfrak{Q} \tag{8}$$

усіх його оптимальних стратегій,

де
$$\mathfrak{L} = \left\{\mathbf{\Lambda} = \left[\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_{F-1} \quad \lambda_F\right] \in \mathbb{R}^F : \sum_{i=1}^{F} \lambda_i = 1,\ \lambda_i \in [0;\,1] \ \forall \ i = \overline{1, F}\right\}. \tag{9}$$

Тоді j-й метод виведення системи обробки інформації на нормальний режим функціонування проектувальник застосовуватиме з імовірністю $\breve{q}_j$ у стратегії (7). Крім того, задача (6) вже матиме розв'язок

$$j_{\text{opt}} \in \arg \min_{j=\overline{1,\, S}} \sqrt{\sum_{i=1}^{F} \left(\widehat{\lambda}_i w_{ij}\right)^2}\,, \tag{10}$$

у якому множина коефіцієнтів $\left\{\widehat{\lambda}_i\right\}_{i=1}^{F}$ вибирається з точки [2]

$$\widehat{\mathbf{\Lambda}} = \left[\widehat{\lambda}_1 \quad \widehat{\lambda}_2 \quad \cdots \quad \widehat{\lambda}_{F-1} \quad \widehat{\lambda}_F\right] \in \widehat{\mathfrak{L}} = \arg \max_{\mathbf{\Lambda} \in \mathfrak{L}} \min_{\mathbf{Q} \in \mathfrak{Q}} \left(\mathbf{\Lambda} \cdot \mathbf{W} \cdot \mathbf{Q}^{\mathrm{T}}\right) \subset \arg \max_{\mathbf{\Lambda} \in \mathfrak{L}} \left(\mathbf{\Lambda} \cdot \mathbf{W} \cdot \breve{\mathbf{Q}}^{\mathrm{T}}\right) \subset \mathfrak{L}\,. \tag{11}$$

Виходячи з умов відсутності даних про імовірнісний розподіл на множині F характеристик

$\{\vartheta_i\}_{i=1}^{F}$ системи обробки інформації, її виведення на нормальний режим функціонування здійснюватиметься за Q етапів через практичну реалізацію оптимальної змішаної стратегії [9] проектувальника (7). Це завдання зводиться до досягнення проектувальником за Q етапів керування системою середніх втрат, що не перевищують значення $\left[w_{ij}\right]_{F\times S}$-гри $v_{\mathrm{opt}} = \widehat{\mathbf{\Lambda}} \cdot \mathbf{W} \cdot \breve{\mathbf{Q}}^{\mathrm{T}}$ з деяким спрощенням (релаксацією), де у граничному переході до $Q \to \infty$ ця релаксація має зникати. Нестохастичні невизначеності характеристик $\{\vartheta_i\}_{i=1}^{F}$ системи обробки інформації при цьому вже враховані в ігровій моделі з матрицею $\mathbf{W} = \left[w_{ij}\right]_{F\times S}$, а переходи в межах S методів $\{m_j\}_{j=1}^{S}$ керування системою мають відбуватись поетапно. Для нейронної мережі це еквівалентно перемиканню методів її навчання у процесі поетапного нарощування шумового навантаження на навчальні вибірки.

Як приклад розглянемо задачу розпізнавання прямокутних кольорових зображень нейронною мережею з прямим зв'язком, що використовує один прихований шар з 1250 нейронів. Розпізнавання ведеться серед класу 40 bmp-зображень форматом 1280×960 точок з глибиною кольору у 24 біти, але на вхід нейронної мережі надходитимуть дещо відмінні від оригіналів зображення, котрі формуються додаванням до 40 оригінальних зображень гаусового шуму з нульовим середнім і середнім квадратичним відхиленням, що не перевищує 0,5. Для забезпечення функціонування мережі під час її навчання на доступних об'ємах пам'яті застосовується проріджування зображень у формат $80 \times 60 \times 4$. Нейронна мережа спроектована у середовищі MATLAB 7.14 за допомогою інструмента MATLAB Neural Network Toolbox 7.0.3, що надає проектувальнику 13 методів її навчання за АЗРП [3]. Запуск, навчання і тестування нейронної мережі для розпізнавання кольорових зображень здійснюється на синхронізованому кластері 4 комп'ютерів з оперативною пам'яттю 16 гігабайт. Для цієї нейронної мережі у рамках задачі, що розглядається, виділено три ключових характеристики, середньостатистичні оцінки яких отримано за результатами 880 тестів і наведено у табл. 1: період навчання ϑ_1, середньостатистичний рівень помилок розпізнавання ϑ_2 і відносна частота ϑ_3 паралічу (зависання) мережі.

Таблиця 1

Середньостатистичні оцінки характеристик нейронної мережі, отримані за результатами 880 тестів розпізнавання серед класу 40 прямокутних зображень формату $1280 \times 960 \times 24$

Номер методу навчання	Назва методу навчання у середовищі MATLAB 7.14	Оцінка періоду навчання ϑ_1 у секундах	Оцінка рівня помилок розпізнавання ϑ_2	Оцінка відносної частоти ϑ_3 паралічу мережі
1	trainbfg	95,4	0,8613	0,8443
2	trainbr	2541,6	0,1004	0,0716
3	traincgb	7030,8	0,6611	0,5239
4	traincgf	1778,4	0,8743	0,8102
5	traincgp	874,8	0,5396	0,3977
6	traingd	2565	0,1304	0,0239
7	traingda	392,4	0,1271	0,0705
8	traingdm	2597,4	0,1323	0,0295
9	traingdx	403,2	0,1271	0,0341
10	trainlm	466,2	0,2213	0,1477
11	trainoss	8031,6	0,5507	0,3727
12	trainrp	271,8	0,3791	0,2386
13	trainscg	120,6	0,1319	0,0352

З точки зору продуктивності така система обробки інформації не може виводитись на нормальний режим свого функціонування за методами, через які $\vartheta_2 > 0{,}15$. Тому множиною допустимих методів навчання є методи з номерами $\{2, 6, 7, 8, 9, 13\}$ у табл. 1. Після нормування типу (1) отримуємо задачу прийняття рішень з транспонованою для неї матрицею

$$\mathbf{W}_0 = \begin{bmatrix} 0{,}9785 & 0{,}9875 & 0{,}1511 & 1 & 0{,}1552 & 0{,}0464 \\ 0{,}7585 & 0{,}9854 & 0{,}9609 & 1 & 0{,}9602 & 0{,}997 \\ 1 & 0{,}3333 & 0{,}9841 & 0{,}4127 & 0{,}4762 & 0{,}4921 \end{bmatrix}. \qquad (12)$$

З матриці (12) з множиною номерів $\{2, 6, 7, 8, 9, 13\}$ випливає, що альтернатива m_8 строго домінується альтернативою m_6. Очевидно, що тут трикритеріальна задача (3) не має точного розв'язку, адже решта $\mathbb{R}^3$- точок в (12) утворюють множину Парето, оптимізація на якій вимагає оцінок коефіцієнтів $\{\lambda_i\}_{i=1}^{F}$ для $F = 3$. Щоправда, за критерієм Бернуллі [1], прийнявши $\left\{\lambda_i = F^{-1}\right\}_{i=1}^{F}$,

$$\arg\min_{j=\overline{1,\,6}} \sum_{i=1}^{3} w_{ij} = \arg\min\{2,737,\ \ 2,3062,\ \ 2,0961,\ \ 2,4127,\ \ 1,5917,\ \ 1,5355\} = \{6\};$$

$$\arg\min_{j=\overline{1,\,6}} \sqrt{\sum_{i=1}^{3}\left(w_{ij}\right)^2} = \arg\min\{2,5329,\ \ 2,0573,\ \ 1,9146,\ \ 2,1703,\ \ 1,1729,\ \ 1,2383\} = \{5\},$$

тобто у множині допустимих методів навчання $\{2, 6, 7, 8, 9, 13\}$ з табл. 1 оптимальними є 13-й та 9-й, відповідно. Та надшвидке навчання (упродовж двох хвилин) за методом trainscg призводитиме до значних помилок розпізнавання на рівні $\vartheta_2 = 0,1319$. А збільшення процесу виводу нейронної мережі на нормальний режим роботи у більш ніж утричі, використовуючи метод traingdx, не зменшить цей рівень навіть на 4 %. До того ж, не зважаючи на прийнятні шанси уникнути паралічу мережі, ці два методи навчання забезпечують низьку її продуктивність порівняно з методом trainbr, оскільки невідомо, коли доведеться нейромережу перенавчати (донавчати). Тому розв'язуючи задачі розпізнавання прямокутних кольорових зображень нейронною мережею з прямим зв'язком за оцінками трьох її характеристик у табл. 1, слід використати комбінацію допустимих методів навчання за оптимальною стратегією проектувальника (архітектора):

$$\mathbf{Q}_{\text{opt}} = \left[q_1^{\langle\text{opt}\rangle}\quad q_2^{\langle\text{opt}\rangle}\quad q_3^{\langle\text{opt}\rangle}\quad q_4^{\langle\text{opt}\rangle}\quad q_5^{\langle\text{opt}\rangle}\right] = [0,694\quad 0,116\quad 0\quad 0,19\quad 0] \in \boldsymbol{Q} \subset \mathbb{R}^5 \tag{13}$$

у грі з підматрицею

$$\mathbf{W} = \begin{bmatrix} 0,9785 & 0,9875 & 0,1511 & 0,1552 & 0,0464 \\ 0,7585 & 0,9854 & 0,9609 & 0,9602 & 0,997 \\ 1 & 0,3333 & 0,9841 & 0,4762 & 0,4921 \end{bmatrix} \tag{14}$$

матриці (12). Методи trainbr, traingd і traingdx комбінуватимуться з імовірностями $q_1^{\langle\text{opt}\rangle} = 0,694$; $q_2^{\langle\text{opt}\rangle} = 0,116$ та $q_4^{\langle\text{opt}\rangle} = 0,19$, відповідно, упродовж процесу навчання, який для проектувальника розбитий на $Q = 20$ етапів (партій гри), на кожному з яких нейромережа навчається на навчальній вибірці з доданим фіксованим рівнем гаусового шуму, який еквівалентний відмінностям між реальними образами і 40 еталонними. Паралельно з цим, нормовані згідно з (1) втрати проектувальника не мають перевищувати значення гри v_{opt} з уступкою [7, 9]

$$\delta_2(g) = 0,4062 g^{-1,001} \quad \forall\ g = \overline{1,\,20} \tag{15}$$

на кожному g-му етапі. І стратегія (13) дозволить досягти оптимального значення $v_{\text{opt}} \approx 0,8232$ такої гри з можливою уступкою $\delta_2(20)$, тобто за цією стратегією відносні втрати архітектора не перевищуватимуть значення $0,8232 + \delta_2(20)$. Та очевидно, що нейромережа гарантовано [1, 2] набуває (теоретичних) характеристик

$$\vartheta_i = v_{\text{opt}} \cdot \max_{j=\overline{1,\,S_0}} u_{ij} \quad \forall\ i = \overline{1,\,3}, \tag{16}$$

тільки якщо $Q \to \infty$. Тому оцінки її характеристик $\{\vartheta_i\}_{i=1}^{3}$, отримані за класичною моделлю реалізації стратегії (13) та за моделлю з використанням функції допусків втрат (15), відрізняються (табл. 2) від теоретичних оцінок (16).

Як випливає з даних табл. 2, отримані за моделлю реалізації стратегії (13) з використанням функції (15) оцінки характеристик нейронної мережі є найбільш подібними до реально зафіксованих, де вона проходить процес навчання (ініціалізації) за менш ніж 36 хвилин з імовірністю $\vartheta_3 = 0,0542$ відмови (паралічу) і працює, забезпечуючи розпізнавання образів на рівні $\vartheta_2 = 0,1057$, що дуже близький до найменшого за методом trainbr у табл. 1.

Таблиця 2

Теоретичні та модельні оцінки характеристик нейронної мережі у порівнянні з реально зафіксованими середньостатистичними оцінками її характеристик

Вид оцінок характеристик нейронної мережі	ϑ_1	ϑ_2	ϑ_3
Теоретичні оцінки, отримані за стратегією (13)	2138,0818	0,1089	0,0589
Оцінки, отримані за класичною моделлю [10] реалізації стратегії (13)	2268,2184	0,1155	0,0625
Оцінки, отримані за моделлю [9] реалізації стратегії (13) з використанням функції допусків втрат (15)	2031,1561	0,1035	0,05598
Реально зафіксовані за результатами 240 тестів середньостатистичні оцінки характеристик нейромережі	2130,4871	0,1057	0,0542
Оцінки характеристик нейромережі за розв'язком задачі типу (10), де мінімальна відстань до точки утопії досягається на MATLAB-методі traingdx	403,2	0,1271	0,0341

Використання методу traingdx, на якому досягається мінімальна відстань до точки утопії за розв'язком задачі типу (10) з коефіцієнтами $\left\{\hat{\lambda}_i\right\}_{i=1}^{F}$ в (11) для матриці (14), значно завищує рівень помилок розпізнавання. Зазначимо також близькість теоретичних оцінок (16) до зафіксованих середньостатистичних значень за результатами 240 тестів нейромережі.

Висновок

Запропонований критерій мінімізації втрат за оптимальною стратегією (7) другого гравця у матричній $\left[w_{ij}\right]_{F \times S}$-грі дозволяє здійснювати комбінований вибір методів навчання нейронної мережі однозначно та незалежно від вагомості її оцінюваних властивостей. Це наочно демонструє приклад оптимізації процесу навчання нейронної MATLAB-мережі для розпізнавання 40 кольорових зображень, де обмеження втрат архітектора задається функцією (15). Оптимальна стратегія проектувальника (7) разом з відповідною функцією обмеження його втрат може використовуватись і в інших системах обробки інформації, у яких перехідний процес від ініціалізації до початку сталої роботи розбивається на скінченну кількість етапів, під час яких комбінуються варіанти виводу системи обробки інформації на допустимий режим функціонування. Також коефіцієнти з точки (11) можуть використовуватись для згортання критеріїв у розв'язуванні F-критеріальної задачі (3) як (4) або (6).

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Трухаев Р. И. Модели принятия решений в условиях неопределённости / Р. И. Трухаев. — М. : Наука, 1981. — 258 с.
2. Теория игр : учеб. пос. для ун-тов / [Л. А. Петросян, Н. А. Зенкевич, Е. А. Семина]. — М. : Высшая школа, Книжный дом «Университет», 1998. — 304 с.
3. Хайкин С. Нейронные сети: полный курс / С. Хайкин. — 2-е изд. — М. : Издательский дом «Вильямс», 2006. — 1104 с.
4. Xudong Jiang. Constructing and training feed-forward neural networks for pattern classification / Xudong Jiang, Alvin Harvey Kam Siew Wah // Pattern Recognition. — 2003. — Vol. 36, Issue 4. — P. 853—867.
5. Goltsev A. Investigation of efficient features for image recognition by neural networks / A. Goltsev, V. Gritsenko // Neural Networks. — 2012. — Vol. 28. — P. 15—23.
6. McLoone S. Improving neural network training solutions using regularisation / S. McLoone, G. Irwin // Neurocomputing. — 2001. — Vol. 37, Issues 1—4. — P. 71—90.
7. Upper bound of the expected training error of neural network regression for a Gaussian noise sequence / [K. Hagiwara, T. Hayasaka, N. Toda and other] // Neural Networks. — 2001. — Vol. 14, Issue 10. — Pp. 1419—1429.
8. Romanuke V. V. Finding the probability distribution to minimax single-valued mapping for modeling the being controlled system parameter / V. V. Romanuke // XI International Conference «MCCS-2012», October 9—11, 2012, Vinnytsya : abstracts. — Vinnytsya : VNTU, 2012. — P. 7—8.
9. Романюк В. В. Метод реалізації оптимальних змішаних стратегій у матричній грі з порожньою множиною сідлових точок у чистих стратегіях з відомою кількістю партій гри / В. В. Романюк // Наукові вісті НТУУ «КПІ». — 2009. — № 2. — С. 45—52.
10. Томашевський В. М. Моделювання систем / В. М. Томашевський. — К. : вид. група BHV, 2005. — 352 с.

Рекомендована кафедрою комп'ютерних систем управління

Стаття надійшла до редакції 16.11.12
Рекомендована до друку 11.01.13

Романюк Вадим Васильович — доцент.

Кафедра прикладної математики та соціальної інформатики, Хмельницький національний університет, Хмельницький