

## ТЕХНОЛОГІЯ КЛАСТЕРИЗАЦІЇ НА БАЗІ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ ОПТИМІЗАЦІЇ ІНФОРМАЦІЙНОЇ БАЗИ ТЕСТІВ

Костікова М. В., Скрипіна І. В.,  
ХНАДУ

**Анотація.** З метою оптимізації формування тестових завдань подано можливу формалізацію даних, технологію кластеризації на базі штучних нейронних мереж. Визначено задачі, які виникають при створенні бази даних тестових завдань. Запропоновано процедуру навчання нейронної мережі з використанням пакета математичного моделювання *Matlab*.

**Ключові слова:** *Moodle*, інформаційна база тестів, самоорганізація, оптимізація, фактор, нейронна мережа, кластеризація, мережа Кохонена.

### Вступ

Сучасні інформаційні технології надають великі можливості одержувати освітні послуги, коли людина знаходиться вдома, в офісі та навіть на вулиці. Застосування цих технологій прибирає межі між дистанційним навчанням і безпосередньо всередині навчального закладу.

На сьогодні у більшості університетів пліч-о-пліч із заочною формою навчання існують дистанційні курси навчання з різних дисциплін за спеціальностями вищих навчальних закладів.

На кафедрі інформатики і прикладної математики ХНАДУ сконструйовано лінійку курсів з вивчення таких дисциплін як «Комп'ютерна техніка та програмування», «Сучасні інформаційні та комп'ютерні технології», «Інформатика», «Інформаційні системи і технології на транспорті». Як платформа в цих курсах була використана система *Moodle*. Ця система орієнтована передусім на організацію взаємодії між педагогом та членами академічної групи, хоча вона є придатною і для організації звичних дистанційних курсів, а також підтримки очного навчання. Вже декілька років ця система успішно використовується під час навчання студентів-заочників нашого університету. З другого боку, елементи цих курсів активно використовують і студенти денної форми навчання.

Система *Moodle* дозволяє розробляти електронні навчальні курси, проводити аудиторне навчання та навчання на відстані. Вказаний програмний продукт створений відповідно до стандартів інформаційних освітніх систем. Він відповідає таким характеристикам як інтеропераційність, багаторазове використання, адаптивність, довговічність, доступність, економічна ефективність.

Викладачі можуть застосовувати *Moodle* для створення дистанційних курсів, завантаження файлів і виготовлення уроків, ведення чатів та форумів, проведення тестувань та іспитів, швидкої і зручної зміни навчальних матеріалів; також вказана система дозволяє встановлювати термін виконання завдань та багато іншого. На нашій кафедрі викладачі вже не перший рік використовують цю систему в навчальному процесі й весь цей час шукають шляхів покращення його проведення, у тому числі під час проведення тестування з метою оцінки отриманих студентами теоретичних знань та практичних навичок роботи на персональному комп'ютері.

Обговоримо такий елемент дистанційного навчання як перевірка і контроль знань студентів. Одним з найважливіших і невід'ємних компонентів навчання є контролювання якості, оцінювання знань та умінь, отриманих студентами. Система тестування дозволяє виявляти рівень і якість засвоєння в системі дистанційного навчання. Розроблена система тестування використовується і для контролю, і для навчання студентів. Головними її складовими є бази тестових завдань та респондентів з результатами тестування і статистикою.

Потрібно врахувати, що дослідження стану контролю знань студентів із застосуванням тестових вимірників виявило певні проблеми при використанні тестів: недостатня якість і валідність змісту тестових завдань, ненадійність результатів тестування, недоліки обробки результатів за класичною теорією тестів, відсутність використання сучасної теорії обробки тестових матеріалів із застосуванням обчислювальної техніки. Висока похибка вимірювання тестових результатів

не дозволяє говорити про високу надійність результатів вимірювання.

Можливості системи *Moodle* для проведення тестування такі: у закритій формі (множинний вибір); на встановлення відповідності; на встановлення вірної послідовності; вірно / невірно – питання альтернативного типу (так / ні); коротка відповідь; числовий; вбудована відповідь (закритий); есе – питання відкритого типу, що вимагає розгорнутої відповіді: огляду, твору, звіту (таке питання оцінюється викладачем); опис; обчислюваний.

При використанні цієї системи доходимо висновку: значна частка тестових завдань – це завдання закритого типу, для яких є ймовірність вгадування правильних відповідей. У найбільш типових випадках, коли студенти повинні вибрати один правильний варіант із тридцяти п'яти запропонованих, ця ймовірність варіює від 25 % до 35 %. Навіть питання на відповідність допускають до 20 % можливості вгадування відповідей.

Отже створюється необхідність враховувати це як під час визначення підсумкової оцінки, так і у процесі оцінювання психометричних показників окремих завдань.

Під час проведення тестування студентів ми зіткнулися з тим, що частина з них робить вибір з усіх наявних варіантів, а частина відкидає варіанти, які вважає неприйнятними, і робить вибір з варіантів, що залишилися. Оцінювання успішності виконання кожного завдання потребує аналізу статистики розподілу вибору варіантів відповідей. Такі дані можна одержати, якщо у статистиці тесту вибрати інформацію щодо конкретного завдання. Втім для аналізу потрібні чималі вибірки результатів, отриманих в однакових умовах, що не завжди можливо під час проведення тестування в *Moodle*.

У нашій системі ми вирішили вилучити з бази тести вірно / невірно, тести в закритій формі (множинний вибір), тести на відповідність.

Тестування ускладнюється також можливістю простого заучування вірних відповідей. Досвід засвідчив, що студент здатний завчити правильні відповіді на 30–40 питань, при цьому не вдумуючись в їх смислову оцінку.

На перших стадіях розробки інформаційної бази тестів ми (щоб виключити заучування і заучування) збільшували кількість тестів, що невідворотно збільшило стомлю-

ваність студентів і не привело до бажаного результату.

Для вирішення цих проблем стали використовувати можливість системи створення питань випадкового вибору одного з декількох аналогічних відповідей.

Для вдосконалення системи тестування застосовуються генератори тестових даних. У випадку автоматичної генерації тестів забезпечується унікальність будь-якого отриманого тесту.

У свою чергу перед нами постало завдання оптимізації інформаційної бази. Набори тестів, створені в автоматичному режимі з використанням випадкових питань і відповідей, виявилися неоднаковими за складністю. Їх використання суттєво впливало на правильність оцінювання якості знань студентів. За однакової кількості тестових завдань складність набору тестів виявлялася різною.

### Аналіз публікацій

Для управління складністю треба визначити функцію залежності складності тестових завдань від вхідних характеристик цих завдань.

Ми зіткнулися із проблемою, що оптимізація побудови аналітичної залежності традиційними методами не можлива. В наш час вирішується задача самоорганізації інформаційної бази з використанням нейромережових технологій [1].

У матеріалах [2] подано найважливіші розділи теорії штучних нейронних мереж. У книзі дається детальний огляд і опис найважливіших методів навчання мереж різної структури.

У джерелі [3] розглянуто основні типи нейронів, архітектор алгоритмів навчання штучних нейронних мереж. Особлива увага приділяється задачам обробки інформації в реальному часі. Поряд із традиційними, книга містить оригінальні результати, пов'язані з оптимізацією процесів навчання.

Інформаційна база тестових завдань складає вектор полів  $\bar{X} = (X_1, X_2, \dots, X_n)$ . Ці фактори впливають на складність тестів. Після відповідних перетворень і зведення інформації в якісну форму (так для тесту в закритій формі присвоюємо коефіцієнт 1, тест на відповідність – 2, тест із вбудованою відповіддю – 3, тест який обчислюється – 4, коротка відповідь – 5 і так далі).

Після виконання такої формалізації буде створено максимізовану ентропію закодованих тестових завдань. Найбільшу ентропію

серед статистичних функцій розподілу має рівномірний розподіл. Отже, після нормування тестові завдання рівномірно заповнюють одиничні інтервали. При модифікації всі фактори мають однакове інформаційне навантаження. До вектора  $\overline{X}$  ще треба додати час, який буде головним фактором при встановленні тенденцій і прогнозування.

Першу модель ідентифікації залежності складності тестового завдання, що є вихідною характеристикою вхідних факторів, формально можна поти як визначення функції

$$B = F(X_1, X_2, \dots, X_n). \quad (1)$$

Такі задачі розв'язуються методом найменших квадратів; задачі самоорганізації моделей із типовим представником розв'язуються методом групового врахування аргументів, трендовим аналізом та апроксимацією рядами Фур'є.

Друга модель визначення коефіцієнтів чутливості тестових завдань до зміни значення вхідних факторів, іншими словами, визначення

$$c_k = \frac{\partial B}{\partial X_k}, d_k = \frac{c_k \overline{X_k}}{\overline{B_k}}, k = \overline{1, n}, \quad (2)$$

де  $c_k$  – абсолютний коефіцієнт чутливості;  $d_k$  – відносний коефіцієнт чутливості;  $\overline{X_k}$ ,  $\overline{B_k}$  – середні значення  $k$ -го фактора вихідної характеристики відповідно в певному класі тестів.

Третя модель будується при прогнозуванні складності бази тестових завдань за її зміни з часом. Вона залежить від значень вхідних факторів та їх композиції

$$B = G(X_i, t), B = G(X_{i_1}, X_{i_2}, \dots, X_{i_k}, t), \quad (3)$$

де  $\overline{X_i} = (X_{i_1}, X_{i_2}, \dots, X_{i_k}, t)$  –  $i$ -й вектор факторів. Необхідно визначити динаміку впливу цього фактора на  $B$ .

Значна кількість вхідних факторів, наявність додаткових умов у початкових даних, значна наявність нетривіальних залежностей між вихідною характеристикою і композицією вхідних факторів, необхідність попередньої обробки початкових даних – все це є особливостями моделей (1)–(3).

Побудова моделей (1)–(3) і пошук необхідних результатів при їх реалізації є потріб-

ними для розробки інформаційно-аналітичного забезпечення бази тестових завдань.

На точність реалізації моделей (1)–(3) впливає інформативність тестових завдань, які містяться в базі. Складність тестів залежить як від об'єктивних, так і від суб'єктивних ситуацій. Вони визначаються мультиколінеарними факторами. Наявність великої кількості тестових завдань, різноманітні типи потребують зменшення множини вхідних факторів. Для цього необхідно виявити найменш значущі і зробити вилучення. Здійснити це традиційними методами без побудови аналітичної залежності дуже важко. Якщо функцію одержано, але композиційна складність залежності не дозволить отримати вірний результат.

### Мета і постановка завдання

Використання нових інформаційних технологій стирає границі між дистанційним навчанням і освітою безпосередньо в школі. Контроль якості, оцінка знань і навичок, що здобуваються студентами, є одним з найважливіших і невід'ємних компонентів освіти. Система тестування дозволяє визначити рівень і якість навчання в системі дистанційного навчання. Генератори тестових завдань використовуються для поліпшення системи керування.

Метою дослідження є удосконалення інформаційної бази тестових завдань.

Об'єкт дослідження – база даних тестових завдань. Предмет дослідження – тестові завдання.

### Основний матеріал досліджень

Активна інтеграція України до європейської спільноти в наш час ставить перед вищою школою нові завдання. Актуальність тестування знань студентів (отримання освітніх послуг без участі викладача, за допомогою сучасних інформаційно-освітніх технологій і систем телекомунікації) визначається тим, що традиційні методи і форми навчання сьогодні вже не можуть повністю задовольнити потребу в послугах навчання для всіх категорій населення.

Провідні науковці вважають, що за допомогою штучних нейронних мереж можна розв'язувати майже всі задачі. Проблема полягає у правильній їх формалізації, визначенні необхідної потужності множини початкових даних і виборі архітектури та принципів функціонування мереж, оскільки її кількість становить декілька тисяч. Задачі

кластеризації та класифікації з успіхом вирішують нейронні мережі.

База даних тестових завдань є тим об'єктом дослідження, для якого основними процедурами є впорядкування, систематизація та класифікація. Значна кількість об'єктивних факторів, які впливають на ефективність її використання, створюють перепони до аналізу та прогнозування тенденцій усієї бази даних тестових завдань.

Самоорганізацію інформаційної бази тестових завдань здійснимо з використанням нейромережових технологій [2, 3], які дозволяють полегшити викладачу процес прийняття важливих і неочевидних рішень в умовах невизначеності, дефіциту часу і обмежених інформаційних ресурсів. Нейронні мережі дозволяють побудувати якісну інформаційну систему для оцінки знань студентів.

Необхідно розв'язати задачу кластеризації, яка полягає у визначенні груп (кластерів) векторів входу, які мають певні спільні властивості. Вхідні фактори і вихідні характеристики належать до векторів входу.

До одного і того ж кластеру належать образи, відстань між якими є менше за деяке додатне число. Образи подібного кластера належать гіперсфері. Значення дрібних факторів можуть виходити за межі гіперсфери по одній чи декількох осях, тому що вони не впливатимуть на належність образу до класу. Визначення цих факторів та їх видалення з інформаційної бази дозволить зменшити наявність шумових ефектів, скоротити час на-

вчання нейронної мережі, а також збільшити точність ідентифікації.

Для кластеризації використаємо самоорганізуючу мережу, вона називається мережею Кохонена. Мережа реалізує принцип навчання без вчителя, і результатом її функціонування буде формування класів та віднесення до них дослідних образів. В системі моделювання *Matlab* алгоритм навчання має такий вигляд

```
net=newc(pr,s,klr),
net.trainParam.epoch s = 1000
net=train(net,P),
a=sim(net,P)
ac=vec2ind(a).
```

Мережа Кохонена є одношаровою. Така мережа створюється після виконання функції *newc(pr,s,klr)*, де *pr* – матриця мінімальних і максимальних елементів векторів входу, *s* – кількість кластерів (нейронів), *klr* – коефіцієнт навчання. Функція *train(net,P)* здійснює навчання мережі. Визначимо таблицю початкових даних як *P*, а кількість циклів навчання установимо 1000.

Функції *a = sim(net,P)*, *ac = vec2ind(a)* виконують моделювання роботи мережі після навчання. Елементи вектора *ac* містять номери класів вхідних образів.

На рис. 1 показані архітектура системи і функції, що використовуються при побудові мережі Кохонена.

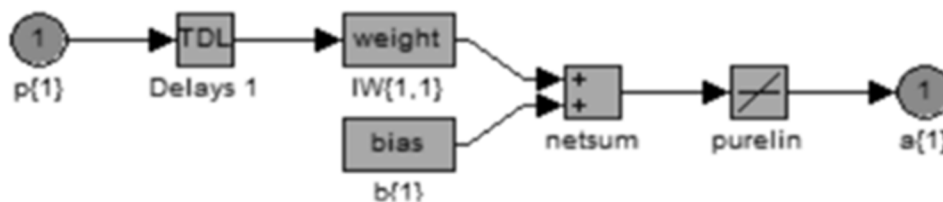


Рис. 1. Архітектура і функції мережі Кохонена

Щоб навчити мережу Кохонена, необхідно певним чином розташувати вагові коефіцієнти. Слід використовувати функції: *negdist* – обчислення від'ємної евклідової відстані вагових коефіцієнтів від вхідного образу  $z = -\sqrt{\sum(w-x)^2}$ ; *netsum* – обчислення активації; *compet* – визначення головного нейрона.

Допустимо, що кількість вхідних образів являє *m*, а кількість кластерів – відповідно

*k*. В результаті моделювання отримаємо вектор  $Q = (q_1, q_2, \dots, q_m)$ , де  $q \in \{1, 2, \dots, k\}$ ,  $i = 1, m$  – номер кластера *i*-го вхідного образу.

Зрозуміло, що кількість кластерів мусить бути більшою чи рівною 2. У задачі вона так само не перевищуватиме кількості вхідних факторів  $n+1$ . Якщо фактор є неважливим, то, незалежно від кількості кластерів, він не буде здійснювати вплив на належність образу до певного кластеру.

Для того, щоб перевірити вартість факторів  $X_1, X_2, \dots, X_n$ , треба виконати алгоритм, наведений вище для первинних образів  $p$  за різноманітних значень  $s$ . Кожному із розглянутих образів буде визначено у відповідність число – номер класу, інакше кажучи, зроблено відображення  $P_i \rightarrow K_i^j$ ,  $i = \overline{1, m}$ ,  $j = \overline{2, n+1}$ , де  $i$  – це номер образу, а  $j$  – кількість кластерів.

Наалі віднаходимо кореляційну матрицю  $R$  векторів  $X_1, X_2, \dots, X_n, K^2, K^3, \dots, K^{n+1}$ .

Для нового аналізу матриці  $R = (r_{ij})_{i,j}^{2n+1}$  будуть потрібні елементи  $r_{ij}$ ,  $i > n+1$ ,  $j \leq n+1$ .

Абсолютні значення за стовпчиками знаходяться як:  $S_i = \sum_{j=1}^{n+1} |r_{ij}|$ ,  $j = \overline{1, n+1}$ .

Для створення вектора значущих факторів викладач, який приймає рішення, повинен задати певне додатне число

$C \in \left( \min_j S_j, \max_j S_j \right)$  і виключити всі фак-

тори, відповідні значення яких  $S_j < C$ .

Втрачена точність у результаті виконання такої процедури покривається зростанням швидкості навчання нашої нейронної мережі та скороченням наявності шумових ефектів. Коли визначається вектор важливих входних факторів, тоді й закінчується перший етап самоорганізації бази даних БТЗ (бази даних тестових завдань).

Черговий етап пошуку полягає у визначенні тих образів, які потрібні для розгляду та аналізу. Беручи до уваги те, що база тестових завдань містить значну кількість записів, притому, незважаючи на вилучення незначних факторів, кількість факторів, що залишились, становить кілька десятків, пошук необхідної інформації триватиме доволі довгий час. Також важливим є розв'язання задачі про те, чи належить нове завдання до класу, що цікавить викладача. Припустимо, що всі ТЗ (тестові завдання), інформація про які міститься в базі даних, належать до двох класів.

Запропонуємо таку процедуру класифікації та визначення потрібних тестових завдань з усієї генеральної сукупності даних по ТЗ – випадковим чином визначимо популяцію завдань. Якщо тестове завдання відноситься до теми модуля, то відносимо його до пер-

шого класу, інакше – до другого. Значимо, що надана популяція мусить бути репрезентативною. Без цього точність класифікації може бути низькою. Користуючись такою класифікацією та інформацією про класи, виконаємо процедуру навчання нейронної мережі в *Matlab*:

```
net=newlvq(pr,sl,lr),
```

```
P=[p1 p2 ... pn],
```

```
Tc=[2 1 ... 2],
```

```
net.trainParam.epochs = 2000,
```

```
net=trainParam.lr=0.05,
```

```
net=train(net,P,Tc).
```

Функція *newlvq* створює мережу для класифікації входних векторів. Найчастіше така мережа проводить кластеризацію і класифікацію векторів входу та є розвитком самоорганізуючих мереж Кохонена [2].

Навчання мережі виконується за допомогою функції *train*. Мережа має лінійні та конкуруючі шари. Після навчання мережі перевіряємо, чи правильно вона виконує класифікацію. Задаємо такі команди:

```
Y = sim(net,P),
```

```
Yc = vec2ind(Y).
```

Після виконання команд отримуємо вектор  $Y_c$ , що збігається з вектором  $T_c$ . Після такої перевірки мережу можна використати. Визначаємо належність тестового завдання відповідному класу. Для цього виконуємо послідовність команд:

```
P1=[p1]
```

```
Y1 = sim(net,P1),
```

```
Y1c = vec2ind(Y1),
```

де  $P1$  – контрольний образ.

Зважаючи, що на вхід нейронної мережі без перерви можна дати значення всіх факторів, що містять інформацію про всі розміщені в базі даних ТЗ, можна одержати її значення виходу, які вказуватимуть на приналежність тестового завдання до потрібного класу. Таким чином, задача класифікації буде розв'язана.

При виборі тестового завдання викладач звертає увагу на значення певних тестів з великою оцінкою, які він вважає головними.

Деякі фактори найчастіше лишаються поза його увагою. Так, при формуванні множини припустимих варіантів ТЗ найчастіше викладач орієнтується на значення індексних вхідних факторів. Вибір об'єкта із цієї множини здійснюється за оптимальною композицією інших факторів. Запропонований нами метод дозволяє оптимізувати цю процедуру за часом.

Застосування цього методу в реальних задачах є можливим при впровадженні спеціальної інтегрованої автоматизованої системи (SIAC). Вона орієнтована на базу даних з розробленими алгоритмами функціонування вказаних мереж.

### Висновки

Розглянута технологія кластеризації на базі нейромережових технологій є варіантом оптимізації великих баз даних. Головне, що нейронна мережа без втручання викладача з'ясовує, яку інформацію можна вилучити без збільшення ентропії, які фактори є другорядними при визначенні складності набору тестових завдань. При її використанні значно скорочується кількість розрахунків при визначенні записів бази тестових завдань, які відповідають вимогам викладача.

Для цього необхідно перевірити усі записи на відповідність кожного поля певному критерію. За новою технологією достатньо навчити нейронну мережу класифікувати тестові завдання, а після цього лише використати її у прямому режимі функціонування для визначення належності певним класам тестових завдань, що містяться у базі даних.

Впровадження запропонованої технології ефективно вплине на підвищення якості оцінки знань студентів.

### Література

1. Костикова М. В. Анализ эффективности тестирования при дистанционном обучении / М. В. Костикова, И. В. Скрипина // Дистанционная освіта у ВНЗ: інноваційні та психолого-педагогічні аспекти: збірник наукових праць за матеріалами Міжнародної науково-методичної конференції, м. Харків, 19–20 листопада 2015 р. – Харків: Міськдрук, 2015. – С. 126–128.
2. Осовский С. Нейронные сети для обработки информации / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.
3. Бодянский Е. В. Искусственные нейронные сети: архитектуры, обучение, применения / Е. В. Бодянский, О. Г. Руденко. – Харьков: Телетех, 2004. – 369 с.

### References

1. Kostikova M. V. Analiz effektivnosti testirovaniya pri distantsionnom obuchenii [Analysis of efficacy testing at a distance education] / M. V. Kostikova, I. V. Skripina // Dystanteijnna osvita u VNZ: Innovatsijni ta psichologo-pedagogichni aspekty: zbirnyk naukovih prac za materialamy Mizhnarodnoyi naukovometodichnoyi konferenciyi, m. Xarkiv, 19 – 20 lystopada 2015 r. – Xarkiv: Miskdruk, 2015. – S. 126 – 128.
2. Osovskiy S. Neyronnyie seti dlya obrabotki informatsii [Neural networks for information processing] /S. Osovskiy. – M.: Finansyi i statistika, 2002. – 344 s.
3. Bodyanskiy E. V. Iskusstvennyie neyronnyie seti: arhitekturyi, obuchenie, primeneniya. [Artificial neural networks: architecture, training, applications] / E. V. Bodyanskiy, O. G. Rudenko. – Harkov: Teleteh, 2004. – 369 s.

**Костікова Марина Володимирівна**, канд. техн. наук, доцент, кафедра інформатики і прикладної математики, Харківський національний автомобільно-дорожній університет, вул. Ярослава Мудрого, 25, Харків, 61002, Україна, телефон +380577073774, [kmv\\_topaz@ukr.net](mailto:kmv_topaz@ukr.net)

**Скрипіна Ірина Валентинівна**, старший викладач, кафедра інформатики і прикладної математики, Харківський національний автомобільно-дорожній університет, вул. Ярослава Мудрого, 25, Харків, 61002, Україна, телефон +380577073774, [scriv@ukr.net](mailto:scriv@ukr.net)

**Technology of clusterization on the basis of artificial neural networks for optimization of the information basis of tests**

**Kostikova M. V., Skrypina I. V.**

***Abstract.** The use of new information technologies blurs the boundaries between distance learning and education directly within the school. One of the most important and integral components of education is the quality control, assessment of knowledge and skills acquired by students. The testing system allows determining the level and quality of learning in a distance learning system. Test task generators are used to improve the control system. **Goal.** To control the complexity of a set of tests, it is necessary to determine the function of dependence of the complexity of test tasks on the input characteristics of these tasks. Optimization of the construction of analytical dependence by traditional methods is impossible. It is necessary to*

solve the problem of self-organization of the information base using neural network technologies. This will allow developing information and analytical support for the database of test tasks. **Methodology.** The object of the study is a database of tests. The main procedures are streamlining, systematization and classification. A significant number of objective factors affecting the efficiency of its use, create obstacles to the analysis and prediction of trends in the entire database of test items. Kohonen self-organizing network was used for clustering. It implements the principle of learning without a teacher and the result of its functioning is the formation of classes and the assignment of research images to them. The algorithm for creating, learning and modeling is implemented in Matlab. Kohonen's network training consists in setting up weights in a certain way. **Results.** The proposed method allows to optimize the procedure for creating and classifying test tasks over time. It is used clustering based on neural network technology. **Originality.** The advantage of the proposed algorithm is that the neural network without the participation of the teacher determines what information can be extracted without increasing entropy and which factors are secondary in determining the complexity of the test. The number of calculations is significantly reduced. **Practical value.**

Using the obtained base of test items will improve the quality of student learning and more objectively assess the knowledge gained in the learning process. It will facilitate the work of the teacher in preparing the examination tasks.

**Key words:** Moodle, test database, self-organization, optimization, factor, neural network, clustering, Kohonen network.

#### Технология кластеризации на базе искусственных нейронных сетей для оптимизации информационной базы тестов

Костикова М. В., Скрипина И. В.

**Аннотация.** С целью оптимизации формирования тестовых заданий представлена возможная формализация данных, технология кластеризации на базе искусственных нейронных сетей. Определены задачи для решения при разработке информационно-аналитического обеспечения процессов в базе данных. Предложена процедура обучения нейронной сети с использованием пакета математического моделирования Matlab.

**Ключевые слова:** Moodle, информационная база тестов, самоорганизация, оптимизация, фактор, нейронная сеть, кластеризация, сеть Кохонена.

---