

УДК 330.4:[330.101.8:519.237.8]:37.014.54(477.8)

Мищук Н. В.

Троханяк С. Р.

Львівський національний університет імені Івана Франка

НЕЧІТКИЙ КЛАСТЕРНИЙ АНАЛІЗ ОСВІТНЬОГО ПОТЕНЦІАЛУ ЗАХІДНОГО РЕГІОНУ УКРАЇНИ В ЧАСОВОМУ ВИМІРІ

У статті викладено проблемні аспекти освітнього потенціалу України, зокрема її Західного регіону. Окреслено основні показники, що визначають освітній потенціал. Запропоновано модель кластерного аналізу освітнього потенціалу Західного регіону України на часовому проміжку. Для реалізації моделі пропонується використати теорію нечітких множин. Під час розрахунків застосовано пакет Fuzzy Logic Toolbox програми MatLab як найбільш популярного та універсального засобу для побудови експертних систем на основі нечіткої логіки.

Ключові слова: освітній потенціал, кластеризація, лінгвістична змінна, нечітка множина, нечіткі логічні висновки, функція належності, терм-множина.

Постановка проблеми. Через своє геополітичне розташування Західний регіон України займає особливе місце у системі її соціально-економічного розвитку. Наявність розвиненої транспортної мережі міжнародного значення, близькість до країн ЄС, сусідство із шістьма іншими країнами визначають особливості розвитку цього регіону.

У Західному регіоні порівняно з іншими регіонами України спостерігається значна освітня еміграція української молоді за кордон. Основними чинниками цієї еміграції до країн Європи є висока якість освітніх послуг ВНЗ, привабливі умови навчання та проживання, вища ймовірність працевлаштування, кар'єрного зростання та гідної оплати праці в майбутньому. Це негативно впливає на освітній потенціал нашої країни та Західного регіону зокрема.

Водночас спостерігаються високі обсяги освітньої еміграції молоді з інших регіонів України до ВНЗ Західного регіону, основними чинниками якої, на нашу думку, є висока престижність певних ВНЗ цього регіону, окремих спеціальностей, за якими ведеться підготовка фахівців, можливість участі у програмах міжнародного стажування, а також бойові дії на сході країни.

Знижує рівень освітнього потенціалу Західного регіону України не тільки еміграція української молоді за кордон. Особливо негативним чинником розвитку освіти у цьому регіоні є вплив інтелектуального капіталу через зростання освітньої міграції науково-дослідних працівників та викладачів ВНЗ, тому основним завданням уряду для виправлення цієї ситуації є проведення ефективних реформ в економіці, освіті, науці та соціальній сфері. Це неможливо здійснити без аналізу в часовому вимірі освітнього потенціалу регіонів України, зокрема і Західного її регіону.

Аналіз останніх досліджень і публікацій. Дослідженню формування та розвитку освітнього потенціалу країни присвятили свої праці такі науковці, як В. Андрущенко, В. Астахова, О. Грішнова, В. Кремень, В. Куценко, М. Степко, І. Мицишин, В. Філіппова та ін., в яких розглядається сутність державної освітньої політики та визначаються деякі напрями її модернізації. Завдання кластеризації вперше було розглянуто в 1930-х роках. Цю проблему вивчали такі дослідники, як Д. МакКуїн, У. Ланс, Д. Уільямс, Д. Хартиган, М. Вонг, Т. Кохонен, Б. Фрицке, А.Н. Колмогоров та ін.

Виділення невирішених раніше частин загальної проблеми. Однак низка питань, пов'язаних з аналізом освітнього потенціалу в часовому вимірі,

не була опрацьована достатньою мірою і потребує більш детального дослідження.

Мета статті полягає у проведенні нечіткої кластеризації Західного регіону України в часовому вимірі для останніх 10 років залежно від рівня його освітнього потенціалу.

Для досягнення цієї мети потрібно вирішити такі завдання:

- проаналізувати та виокремити основні вагомні чинники, які формують освітній потенціал регіону;

- побудувати модель, яка б дала можливість на основі значень первинних показників отримати величини проміжних і остаточного інтегральних показників нечіткої кластеризації освітнього потенціалу регіону;

- побудувати п'ять кластерів, віднісши до кожного з них Західний регіон України за відповідний рік залежно від рівня ОП у ньому у цьому році.

Виклад основного матеріалу дослідження. Проблеми розвитку освітнього потенціалу особливо гостро проявляються сьогодні на регіональному рівні, оскільки саме на рівні регіону можуть бути створені оптимальні умови для формування освітнього потенціалу країни у цілому.

Освітній потенціал (ОП) – сукупність реальних ресурсів освіти (наявних і нагромаджених: знань, навичок, досвіду та ін.), які має суспільство для забезпечення соціально-економічного розвитку, і можливостей їх нарощування у майбутньому шляхом провадження раціональної державної політики в освітній сфері. Кількісні параметри ОП визначають за допомогою низки абсолютних показників (наявність навчальних закладів, обсяг витрат, кількість учнів і студентів) і відносних (частка витрат держави на освіту, частка учнів і студентів у загальній кількості населення, середня тривалість навчання одного працівника тощо) [1].

Освітній потенціал визначають багато первинних показників, які неможливо обчислити з достатньою точністю. Неможливо дати кількісну оцінку на основі слабоструктурованої системи з неповною та неточною інформацією, тому для кластерного аналізу оцінки освітнього потенціалу доцільно застосувати модель на основі нечітких множин та нечіткої логіки.

Нечітка кластеризація – це синтез методів кластерного аналізу та теорії нечітких множин. Суть цієї теорії полягає у тому, що замість звичайного розбиття об'єктів на кластери розглядаються їх нечіткі варіанти. Тобто будуються функції належності об'єктів для кожного зі створених кластерів, кожна з яких змінюється на інтервалі

[0, 1], що дає змогу оцінити належність об'єкту до того чи іншого кластеру [2].

Для моделювання оцінки освітнього потенціалу регіону нечітка модель може бути представлена у вигляді нечіткої продукційної мережі (НПМ), елементи та сукупності елементів якої реалізують різні компоненти нечітких моделей та етапи нечіткого висновку [2]. Для побудови НПМ оцінки ОП потрібно визначити простір передумов $X = \{x_i, i = \overline{1, n}\}$ – первинних показників, які визначають ОП, та повний простір висновків $Y = \{y_j, j = \overline{1, m}\}$ – проміжних і остаточного інтегральних показників оцінки рівня ОП.

Для задання лінгвістичних змінних показників, які визначають ОП, можуть використовуватися такі терм-множини (табл. 1, табл. 2). Ці терм-множини визначають рівні показників [2; 3].

Вказані в табл. 2 проміжні інтегральні показники $Y = \{y_j, j = \overline{1, 3}\}$ визначаються такими первинними показниками: людський потенціал визначається чисельністю економічно активного населення, міграційним приростом, рівнем безробіття та рівнем зайнятості населення; рівень економічного розвитку – валовим регіональним продуктом, індексом реальної заробітної плати, середньою заробітною платою та індексом споживчих цін на товари і послуги; показник «навчальні заклади» – кількістю ВНЗ I–IV рівнів акредитації, кількістю прийнятих на навчання до ВНЗ та кількістю випущених із ВНЗ.

Взаємозв'язок між первинними показниками (антецедентами) та проміжними й остаточним інтегральними показниками (консеквентами) – це, по суті, бінарне нечітке відношення на декартовому добутку відповідних нечітких множин. Нечітке причинно-наслідкове відношення між антецедентом та консеквентом задається у вигляді нечіткої продукції. Процес формування бази нечіткого логічного виведення (НЛВ) являє собою формальне представлення емпіричних знань експертів у досліджуваній предметній сфері за схемою «якщо ..., то ...» [2].

НЛВ моделі кластерної оцінки освітнього потенціалу подано в табл. 3.

Інформаційну базу досліджень становить офіційні статистичні дані показників за 2008–2017 рр., від яких залежить ОП Західного регіону України (Львівська, Волинська, Рівненська, Тернопіль-

Таблиця 2

Проміжні та остаточний інтегральні показники, які визначають освітній потенціал

Позначення	Найменування лінгвістичної змінної
y_1	Людський потенціал
y_2	Рівень економічного розвитку
y_3	Навчальні заклади
y_4	Потреба у працівниках (тис. осіб)
K	Кластери оцінки ОП в часовому вимірі

ська, Івано-Франківська, Закарпатська, Чернівецька області) (табл. 4) [4].

Для автоматизації розрахунку моделі оцінки ОП було вибрано пакет Fuzzy Logic Toolbox програми MatLab як найбільш популярного й універсального засобу для побудови експертних систем на основі нечіткої логіки.

Розрахунки проводилися згідно з таким алгоритмом.

Крок 1. Формування нечіткої продукційної моделі оцінки освітнього потенціалу (рис. 1).

Формування моделі передбачає задання лінгвістичних змінних $(x_1, x_2, \dots, x_{11}, y_1, y_2, \dots, y_4, K)$, які б формалізували первинні та інтегральні показники, що визначають ОП, а також бази нечітких продукційних правил $(БП_1, БП_2, \dots, БП_4)$ [2; 3].

Крок 2. На цьому кроці потрібно задати функції належності для терм-множин вхідних та вихідних лінгвістичних змінних $(x_1, x_2, \dots, x_{11}, y_1, y_2, \dots, y_4, K)$ [2; 5]. Для реалізації моделі для лінгвістичних змінних x_1, x_2, \dots, x_{11} вибрано трапецієвидний тип функцій належності:

$$\mu(x) = \begin{cases} 0, & x \leq a \text{ або } x \geq d; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & b \leq x \leq c; \\ \frac{d-x}{d-c}, & c \leq x \leq d, \end{cases} \quad (1)$$

де: $[a, d]$ – песимістична оцінка (рівень упевненості експертів про наявність в елементів універсальної множини властивостей нечіткої множини є нижчим від 1); $[b, c]$ – оптимістична оцінка (рівень упевненості експертів про наявність в елементів універсальної множини властивостей нечіткої множини рівний 1).

Таблиця 1

Первинні показники, які визначають освітній потенціал

Позначення	Найменування лінгвістичної змінної	Тип терм-множини та інтерпретація рівнів показників
x_1	Чисельність економічно активного населення (у віці 15-70 років), тис. осіб	Н – низький; С – середній; В – високий.
x_2	Рівень безробіття населення у віці 15-70 років (за методологією МОП), % до економічно активного населення відповідної вікової групи	Н – низький; С – середній; В – високий.
x_3	Рівень зайнятості населення, % до економічно активного населення відповідної вікової групи	Н – низький; С – середній; В – високий.
x_4	Міграційний приріст/скорочення, тис. осіб	Н – низький; С – середній; В – високий.
x_5	Валовий регіональний продукт, млн. грн.	Н – низький; С – середній; В – високий.
x_6	Середня заробітна плата (у розрахунку на одного штатного працівника) грн.	Н – низький; С – середній; В – високий.
x_7	Індекс споживчих цін на товари та послуги, % за рік	Н – низький; С – середній; В – високий.
x_8	Індекс реальної заробітної плати, % до попереднього року	Н – низький; С – середній; В – високий.
x_9	Кількість ВНЗ I–IV рівнів акредитації, одиниць	Н – низький; С – середній; В – високий.
x_{10}	Кількість прийнятих на навчання до ВНЗ, осіб	Н – низький; С – середній; В – високий.
x_{11}	Кількість випущених із ВНЗ, осіб	Н – низький; С – середній; В – високий.

Таблиця 3

Нечіткі логічні висновки (фрагменти)

Позначення правила	Антецедент	Консеквент
База правил П1		
П 1.1	$x_1 = B \cap x_2 = H \cap x_3 = B \cap x_4 = B$	$y_1 = B$ (високий)
П 1.2	$x_1 = B \cap x_2 = H \cap x_3 = C \cap x_4 = C$	$y_1 = BC$ (вище середнього)
П 1.3	$x_1 = C \cap x_2 = C \cap x_3 = C \cap x_4 = C$	$y_1 = C$ (середній)
П 1.4	$x_1 = C \cap x_2 = C \cap x_3 = H \cap x_4 = H$	$y_1 = HC$ (нижче середнього)
База правил П2		
П 2.1	$x_5 = H \cap x_6 = H \cap x_7 = H \cap x_8 = C$	$y_2 = H$ (низький)
П 2.2	$x_5 = H \cap x_6 = C \cap x_7 = H \cap x_8 = C$	$y_2 = HC$ (нижче середнього)
П 2.3	$x_5 = C \cap x_6 = B \cap x_7 = C \cap x_8 = H$	$y_2 = HC$ (нижче середнього)
П 2.4	$x_5 = C \cap x_6 = B \cap x_7 = C \cap x_8 = B$	$y_2 = C$ (середній)
База правил П3		
П 3.1	$x_9 = B \cap x_{10} = B \cap x_{11} = B$	$y_3 = B$ (високий)
П 3.2	$x_9 = C \cap x_{10} = B \cap x_{11} = B$	$y_3 = B$ (високий)
П 3.3	$x_9 = B \cap x_{10} = C \cap x_{11} = B$	$y_3 = BC$ (вище середнього)
П 3.4	$x_9 = B \cap x_{10} = B \cap x_{11} = C$	$y_3 = BC$ (вище середнього)
База правил П4		
П 4.1	$y_1 = C \cap y_2 = C \cap y_3 = C \cap y_4 = B$	$K = C$ (середній)
П 4.2	$y_1 = B \cap y_2 = B \cap y_3 = B \cap y_4 = B$	$K = B$ (високий)
П 4.3	$y_1 = H \cap y_2 = H \cap y_3 = H \cap y_4 = H$	$K = H$ (низький)
П 4.4	$y_1 = C \cap y_2 = B \cap y_3 = B \cap y_4 = C$	$K = BC$ (вище середнього)

Таблиця 4

Величини первинних показників, які визначають освітній потенціал*

Показник	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
x_1	4245	4238,2	4256,2	4280,3	4301,5	4324,3	4170,4	4163	4141,6	4107,5
x_2	8,03	10,03	9,09	9,19	8,43	8,01	9,40	9,51	9,83	10,13
x_3	56,76	55,47	56,36	56,74	57,41	58,14	55,16	54,94	54,50	53,79
x_4	-4858	-3976	-3016	-2554	3120	-745	3944	2168	127	389
x_5	112934	111089	130329	162111	188575	192393	224330	281598	326160	358776
x_6	1454,86	1547,43	1828,14	2083,86	2413,71	2612,57	2777	3333,86	4141,86	5979,52
x_7	120,47	111,26	108,51	103,36	99,00	99,71	125,6	143,99	111,60	113,70
x_8	106,71	89,44	119,60	109,93	113,27	106,13	88,33	84,41	104,71	109,80
x_9	170	159	158	156	152	145	140	138	137	135
x_{10}	93461	86143	92112	86916	78358	79543	76906	72015	68728	70048
x_{11}	103406	102287	106165	118890	102225	94308	92368	86587	74325	85894
y_4	19,90	7,85	8,38	11,63	9,90	10,00	7,63	8,18	9,90	15,43

*Розмірність показників (див. табл. 1).

Для лінгвістичних змінних y_1, y_2, \dots, y_4, K було вибрано дзвоноподібний тип функцій належності:
Для терму «Низький» (Н):

$$\mu(x) = \frac{1}{1+x^2} \tag{2}$$

Для термів «Нижче середнього» (НС), «Середній» (С), «Вище середнього» (ВС):

$$\mu(x) = 0,5 + \frac{1}{1+x^2} \tag{3}$$

Для терму «Високий» (В):

$$\mu(x) = 1 - \frac{1}{1+x^2} \tag{4}$$

Крок 3. Задання нечітких правил. Правила моделі формуються на основі загальних закономірностей поведінки досліджуваної системи (табл. 3).

Крок 4. Дефазифікація – перетворення нечіткої множини в чітке число. Існує кілька класичних

алгоритмів нечіткого висновку. Для досліджень було вибрано алгоритм Мамдані (Mamdani). Цей алгоритм використовує схему «два входи – база правил – один вихід» та легко може бути модифікованим для схеми з багатьма входами [3; 5].

Для перетворення дискретної множини значень функцій належності в чітке число використовується метод центру тяжіння:

$$\hat{y} = \frac{\sum_{r=1}^{Y_{\max}} y_r \mu_{M'}(y_r)}{\sum_{r=1}^{Y_{\max}} \mu_{M'}(y_r)} \tag{5}$$

де Y_{\max} – кількість елементів y_r в дискретизованій для обчислення області Y ;

M' – нечітка множина, визначена на Y ;

$\mu_{M'}(y_r) \in [0, 1]$ – функція належності нечіткої множини M' .

У результаті моделювання було отримано оцінки показників, які впливають на ОП (рис. 2)

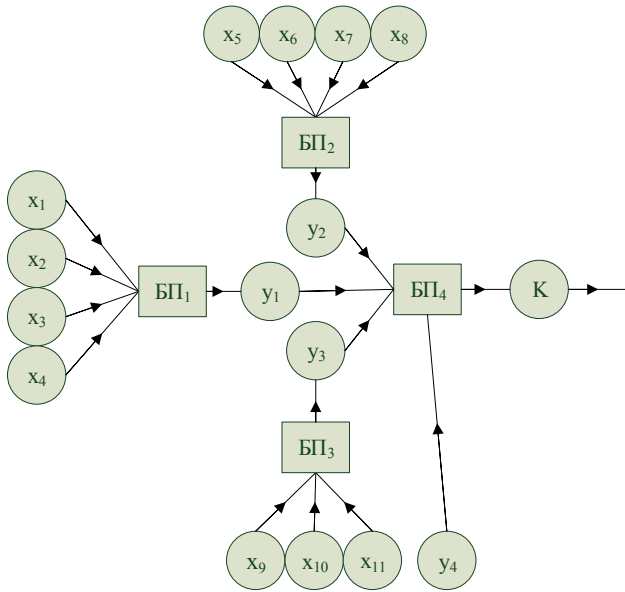


Рис. 1. Нечітка продукційна модель оцінки освітнього потенціалу

Таблиця 5

Кластери рівнів освітнього потенціалу Західного регіону України в 2008–2017 рр.

Роки	Числова оцінка	Кластери
2009	-40	Низький рівень
2015	-18	Нижче середнього
2016	-14	
2014	-9	Середній
2017	0	
2010	2	
2012	18	Вище середнього
2013	25	
2011	27	
2008	41	Високий

Значною мірою це спричинено зовнішньою агресією Росії (анексія Криму та війна на Донбасі). Також спостерігається значне падіння економіки. Суттєво знизився показник «навчальні заклади».

Слід зауважити, що показник «навчальні заклади» залежить не тільки від кількісного та якісного складу навчальних закладів, а й від кількості студентів, які в них навчаються, та фахівців, що ними випускаються. Періоди з 2014 по 2017 р. характеризуються значним зменшенням кількості студентів у навчальних закладах України. Основні причини зменшення студентів – це демографічна криза та навчання значної кількості студентів за кордоном (Польща та інші країни Європи).

Вже в 2017 р. спостерігається дещо краща ситуація з рівнем ОП. Причиною цього є зріст економічних показників за все ще низького рівня показника «людський потенціал». Також спостерігається поліпшення показника «навчальні заклади». Це пояснюється тим, що збільшилася кількість студентів у навчальних закладах Західного регіону, які переїхали зі східних районів України та Криму.

та кластери рівнів освітнього потенціалу (табл. 5). За найбільш оптимістичну оцінку взято число 50, середнє значення – 0, значення найбільш песимістичної оцінки – число -50.

Низький рівень освітнього потенціалу Західного регіону України припадає на 2009 р. У цьому році оцінки всіх проміжних інтегральних показників ОП порівняно з 2008 р. значно знизилися (рис. 2). Особливо стрімко знизилися рівень економічного розвитку та людський потенціал, що стало наслідком негативної демографічної ситуації та еміграції населення.

Нижче середнього рівень ОП спостерігається у 2015–2016 рр. У цей період відбулося значне зменшення показника «людський потенціал».

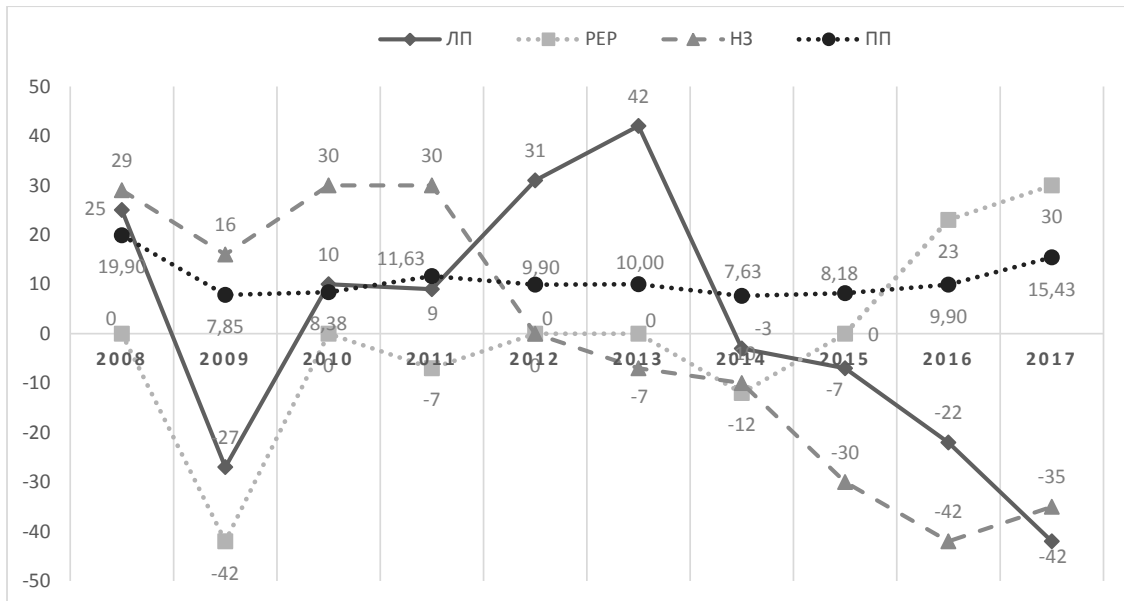


Рис. 2. Графіки проміжних інтегральних показників, які визначають освітній потенціал* Західного регіону України в 2008–2017 рр.

* ЛП – людський потенціал; РЕР – рівень економічного розвитку; НЗ – навчальні заклади; ПП – потреба у працівниках.

Розвиток освітнього потенціалу Західного регіону України є одним із пріоритетних напрямів соціально-економічного розвитку країни. Однак слід урахувати особливості кожної з областей, що входить до складу регіону.

Державна освітня політика також потребує трансформаційних змін. Зусилля уряду доцільно спрямувати на корегування дисбалансу на ринку праці України. Для досягнення цієї мети було б доцільно проводити кластерний аналіз освітнього потенціалу регіонів із використанням запропонованої нечіткої моделі.

Висновки. Таким чином, проведене дослідження показало, що кількісні параметри освітнього потенціалу можна визначити за допомогою низки абсолютних і відносних показників. Однак

неможливо провести обчислення багатьох первинних показників із достатньою точністю, саме тому для оцінки освітнього потенціалу регіонів України запропоновано використовувати нечітку продукційну модель.

Провівши нечіткий кластерний аналіз освітнього потенціалу Західного регіону України в часовому вимірі, виявлено, що низький рівень освітнього потенціалу цього регіону припадає на 2009 р. У цьому році спостерігалось стрімке зниження всіх проміжних інтегральних показників освітнього потенціалу.

Подальших досліджень потребують механізми регулювання розвитку освіти як основного чинника формування конкурентоспроможності молоді на ринку праці України.

Список використаних джерел:

1. Економічний словник. URL: <https://predmety.in.ua/osvitnij-potencial/>.
2. Сявко М.С. Інтелектуалізована інформаційна система «Нечіткий експерт». Львів: ЛНУ імені Івана Франка, 2007. 320 с.
3. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 162 с.
4. Офіційний сайт Державної служби статистики України. URL: <http://www.ukrstat.gov.ua>.
5. Zimmerman H.J. Fuzzy Set Theory and Its Applications. Kluwer, Dordrecht, 1991. 315 p.

Мищук Н. В.
Троханяк С. Р.

Львовский национальный университет имени Ивана Франко

НЕЧЕТКИЙ КЛАСТЕРНЫЙ АНАЛИЗ ОБРАЗОВАТЕЛЬНОГО ПОТЕНЦИАЛА ЗАПАДНОГО РЕГИОНА УКРАИНЫ ВО ВРЕМЕННОМ ИЗМЕРЕНИИ

Резюме

В статье изложены проблемные аспекты образовательного потенциала Украины, в частности ее Западного региона. Обозначены основные показатели, определяющие образовательный потенциал. Предложена модель кластерного анализа образовательного потенциала Западного региона Украины на временном промежутке. Для реализации модели предлагается использовать теорию нечетких множеств. При расчетах был использован пакет Fuzzy Logic Toolbox программы MatLab как наиболее популярное и универсальное средство для построения экспертных систем на основе нечеткой логики.

Ключевые слова: образовательный потенциал, кластеризация, лингвистическая переменная, нечеткое множество, нечеткие логические выводы, функция принадлежности, терм-множество.

Mishchuk N. V.
Trokhanyak S. R.

Ivan Franko National University of Lviv

NON-CHILD CLUSTER ANALYSIS OF EDUCATIONAL POTENTIAL OF THE WESTERN REGION OF UKRAINE IN THE TIME METHOD

Summary

The article outlines the problematic aspects of the educational potential of Ukraine, in particular its western region. The main indicators defining the educational potential are outlined. The model of cluster analysis of educational potential of the western region of Ukraine for a time interval is proposed. To implement the model, it is proposed to use the theory of fuzzy sets. In the calculations, the Fuzzy Logic Toolbox package, MatLab, was used as the most popular and versatile tool for building expert systems based on fuzzy logic.

Keywords: educational potential, clusterization, linguistic variable, fuzzy set, fuzzy logical conclusions, membership function, term-set.