

# РОЗВИТОК РЕГІОНАЛЬНОЇ ЕКОНОМІКИ

М.Д. БІЛИК,  
д.е.н., професор, КНЕУ,  
Н.В. СЕНЬКО,

Київський національний університет ім. Т. Шевченка

## Кластеризація регіонів України за показниками енергоспоживання

*У статті розглянуто технологію самоорганізованих карт Кохонена та зроблений кластерний аналіз даних для формування кластерів регіонів України відповідно до їхнього енергоспоживання.*

**Ключові слова:** самоорганізовані карти Кохонена, кластерний аналіз, енергоспоживання.

*В статье рассмотрены технологии самоорганизующихся карт Кохонена и сделан кластерный анализ данных для формирования кластеров регионов Украины в соответствии с их энергопотреблением.*

**Ключевые слова:** самоорганизующиеся карты Кохонена, кластерный анализ, энергопотребление.

*The article reviews the technology of self-organized Kohonen maps and made cluster analysis to form clusters of regions of Ukraine according to their energy consumption.*

**Постановка проблеми.** Пріоритетним завданням для країни є розвиток регіонів України. Стратегія регіонального розвитку не однотипна по відношенню до різних областей. Така тенденція зумовлена тим, що регіони дуже відрізняються один від одного за забезпеченням ресурсами, структурою їх господарства, досягнутим рівнем розвитку різних сфер економіки, темпами трансформації форм власності, конкурентними перевагами, виявленими в природно-кліматичних, демографічних, виробничих, географічних факторах. Ефективним методом територіально-економічної організації і інструментом зростання конкурентоспроможності регіонів країни виділяють кластерний підхід. У статті акцент

зроблено на створення регіональних кластерів, котрі представляють собою об'єднання областей з подібними підходами енергокористування, з метою включення регіонів у систему світових і регіональних ринків товарів, фінансових і трудових ресурсів, технології і інформації, що визначає їх конкурентоспроможність.

**Аналіз досліджень та публікацій з проблеми.** Проблеми, що пов'язані з організацією кластерів, їх ідентифікацією, функціонуванням та оцінюванням, тенденціями та перспективами розвитку розглядалися в працях С.С. Бакая, В.І. Бойка, А.Д. Войротовича, Ю.П. Воскобойніка, В.М. Кутяна, М.Г. Лобаса, І.І. Лукінова, Б.В. Погрішука, М.Я. Полоцького, П.Т. Саблука, Р.П. Саблука, О.К. Слюсаренка, О.О. Сторожука, І.І. Червенята та інших дослідників. Однак недостатньо розглянутою і новою є тема кластерів для України, враховуючи необхідність підвищення економічного розвитку різних напрямів економічної діяльності.

Передумовами підвищення економічного розвитку України є ефективне управління енергоносіями країни, нарощення можливостей та співпраця між регіонами. Досвід розвинених країн підтверджує багатофункціональну роль кластерного підходу в забезпеченні умов формування і реалізації конкурентних переваг регіонів. У сучасних умовах кластери визнані однією з найбільш ефективних форм інтеграційних систем.

**Метою статті** є кластеризація даних для формування кластерів регіонів України відповідно до їх енергоспоживан-

ня та виявлення особливостей всередині кластерів за допомогою методів самоорганізованих карт Кохонена.

**Виклад основного матеріалу.** Нейронна мережа – це величезний розподілений паралельний процесор, що складається із елементарних одиниць обробки інформації, які накопичують експериментальні знання і представляють їх для подальшої обробки.

Нейронні мережі цікаві не самі по собі, а як інструмент для розв'язання практичних задач, які так чи інакше пов'язані з обробкою образів:

- апроксимація функції за набором точок (регресія);
- класифікація даних на основі заданого набору класів;
- кластеризація даних із виявлення раніш невідомих кластерів-прототипів;
- стиснення інформації;
- відновлення втрачених даних;
- асоціативна пам'ять;
- оптимізація, оптимальне управління.

Нейронні мережі Кохонена характеризуються тим, що в них нейрони представлені у вигляді двомірного масиву, кожному вузлу  $i$  якого поставлено у відповідність вектор  $m_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{in}]^T$ , який має розмірність рівну розмірності простору вхідних векторів  $R^n$ . Тип масиву може бути будь-яким – прямокутний, гексагональний тощо. Як відгук мережі на вхідний вектор  $x \in R^n$  розуміється вузол-переможець  $C$ , відстань (яка може бути визначена будь-яким чином, звичайно Евклідова  $|x - m_c|$ ) до якого від  $x$  мінімальна. Для вибору вузла  $C$  вектор  $x$  порівнюється зі всіма вузлами  $m_i$ ,  $i = \arg \min_i \{ |x - m_i| \}$ , тобто [5]:

$$|x - m_c| = \min_i \{ |x - m_i| \} \quad (1)$$

Навчання починається з ініціалізації початкових векторів  $m_i(0)$ . Як правило, вони визначаються як випадкові значення з рівномірним розподілом в діапазоні значень відповідних компонент вхідних векторів. Якщо є апріорні знання про розподіл значень векторів, то їх можна використовувати, задаючи початкові значення векторів, що сприяє забезпечення подібності методу навчання [8]. У процесі навчання вектора  $m_i$  відповідні до вузлів  $i$  змінюють свої значення відповідно до поданого в даний момент часу  $t$  на вхід вектора  $x(t)$  згідно з правилом:

$$m_i(t+1) = m_i(t) + h_{ci}(t) [x(t) - m_i(t)], \quad (2)$$

де  $t$  – дискретна координата часу;

$h_{ci}(t)$  – функція сусідства, яка грає в процесі навчання головну роль.

Для подібності методу навчання необхідно, щоб  $h_{ci}(t) \rightarrow 0$  при  $t \rightarrow \infty$ . Звичайно [5]:

$$h_{ci}(t) = h(|r_c - r_i|, t), \quad (3)$$

де  $r_c, r_i \in R^2$  – вектори вузлів  $c$  та  $i$  відповідно.

Зі збільшенням  $|r_c - r_i|$ ,  $h_{ci} \rightarrow 0$ . Найчастіше використовуються два простих варіанти для  $h_{ci}(t)$ . Найпростіший з них –

«множина сусідства» точок масиву навколо вузла-переможця  $c$ . Нехай множина їх індексів буде позначена як  $N_c$ , і  $h_{ci}(t) = \alpha(t)$ , якщо  $i \in N_c$  та  $h_{ci}(t) = 0$ , якщо  $i \notin N_c$ :

$$m_i(t+1) = m_i(t) + \alpha(t) [x(t) - m_i(t)], i \in N_c \quad (4)$$

$$m_i(t+1) = 0, i \notin N_c \quad (5)$$

Величина  $\alpha(t)$  називається фактором навчання (швидкістю навчання) ( $0 < \alpha(t) < 1$ ).  $\alpha(t)$  і радіус  $N_c(t)$ , як правило, монотонно спадають з часом [2].

Другий широко розповсюджений варіант задавання  $h_{ci}(t)$  – у вигляді функції Гаусса:

$$h_{ci}(t) = \alpha(t) * \exp\left(\frac{-|r_c - r_i|^2}{2\sigma^2(t)}\right), \quad (6)$$

де  $\alpha(t)$  – скалярний фактор навчання (швидкість навчання);

$\sigma(t)$  – параметр;

$\alpha(t), \sigma(t)$  – деякі монотонно спадні функції від часу [4].

Навчання можна розділити на два етапи – грубого та тонкого налаштування координат векторів, відповідних вузлів карти. В цьому випадку якщо за першого етапу вектори впорядковуються, то за другого – вектори в кожній групі деталізуються. Другий етап, як правило, має меншу швидкість навчання і більшу кількість ітерацій. Обидва етапи навчання здійснюється протягом заданого числа кроків  $T$ , а  $\alpha(t)$ , як правило [7], підбирають так, щоб  $\alpha(T) = 0$ , наприклад

$$\alpha(t) = \alpha * \left(1 - \frac{t}{T}\right), \text{ де } \alpha - \text{ задана константа.}$$

Після закінчення навчання ми отримуємо звичайну карту, яка має вигляд впорядкованого двовимірного масиву, вектори, відповідні вузлам якого, розподілилися в просторі нейронів  $R^n$  відповідно до множини векторів  $X$ , які подаються на вхід. Даний метод нагадує натягування еластичної мембрани на дану множину векторів  $X$ , при чому її еластичність поступово збільшується в процесі натягування (зменшується  $h_{ci}(t)$ ) для все більш тонкого налаштування.

Впорядкування векторів у вигляді двовимірної карти виражається в тому, що чим ближче координати двох векторів на карті, тим ближче вони і в просторі  $R^n$ , але не навпаки [13].

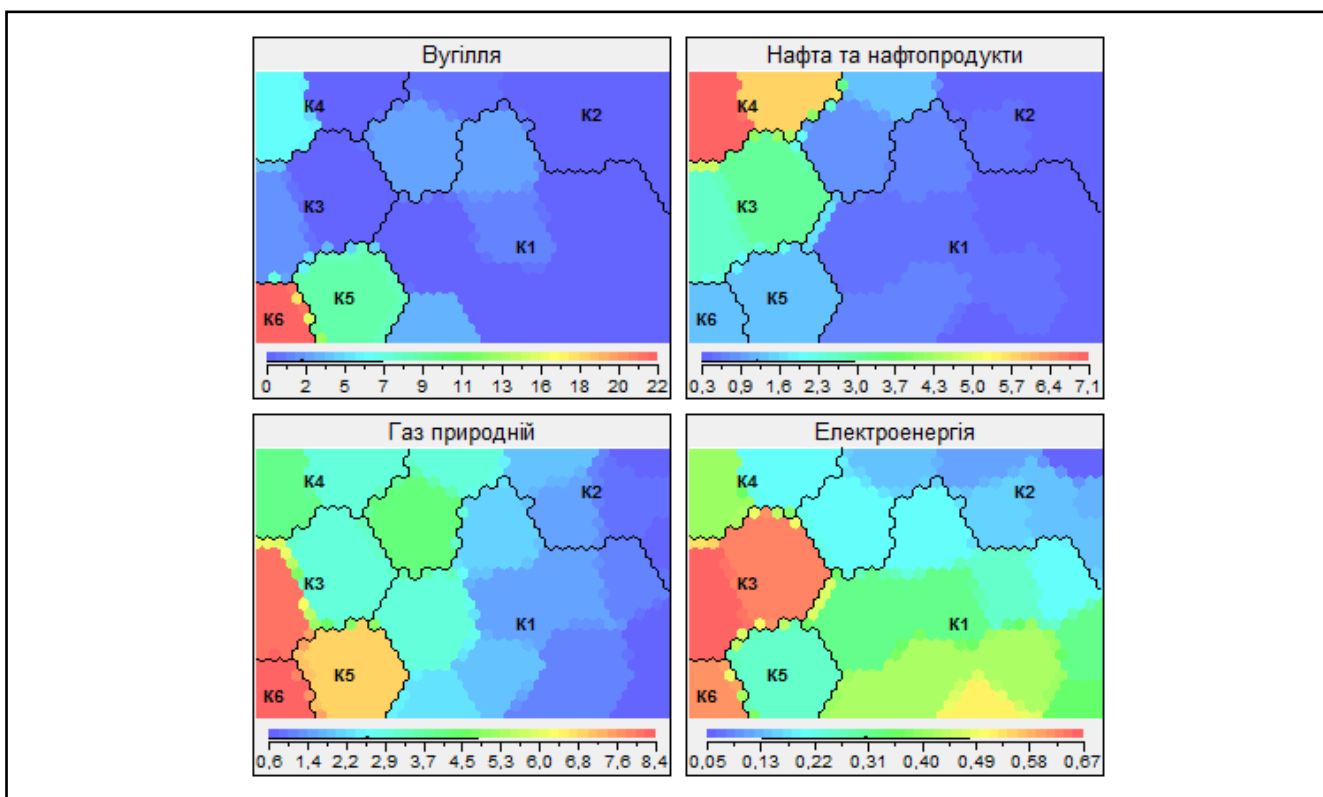
Для кластеризації областей України в роботі використовується база даних, розроблена автором на основі даних Державного комітету статистики. Всі розрахунки проводилися за допомогою комп'ютерної програми обробки статистичних даних Statistica, програми для роботи з самоорганізованими картами ViscoverySOMline 5.2 та програми для створення й обробки електронних таблиць Microsoft Excel.

Процес кластеризації починається з навчання карт. Для цього визначимо основні параметри тренування: щоб роз-



**Рисунок 1. Карта регіонів України**

\* *Власні розрахунки авторів.*



**Рисунок 2. Атлас показників енергоспоживання регіонів**

\* *Власні розрахунки авторів.*

межування на кластери було якісним кількість нейронів встановимо на рівні 1000, а «натяг карти» (міра здатності карти підстроюватись до вхідних даних) дорівнюватиме 0,5.

У результаті застосування технології самоорганізованих карт вся база по регіонах України була перетворена у чотирьохвимірний простір, який карта розподілила на шість економічних кластерів (рис. 1). Різнокольоровість даної карти ні про що не говорить, бо це лише умовність.

На рис. 2 відображені карти, побудовані за чотирма показниками, що використовувалися при побудові карти кластерів регіонів України. Ці карти є вимірами, за допомогою яких будується вся сукупність карт, яку назвемо «атласом

регіонів України». Темні області в деяких отриманих кластерах (рис. 1) свідчать про особливий характер певних регіонів, вони мають мінімальні або максимальні значення одного чи декількох показників, що добре видно на картах атласу, представлено на рис. 2. Це особливі області кластерів, вони будуть більш детально описані далі.

Проілюструємо значення, що відображені на чотирьох картах атласу, навівши таблицю статистики середніх значень отриманих економічних кластерів (табл. 1).

Для чіткішого аналізу опишемо склад та економічні характеристики отриманих кластерів за допомогою таблиці статистики середніх значень та карт атласу.

Таблиця 1. Статистика середніх значень кластерів

Кластер	1	2	3	4	5	6
Питома вага кожного кластера	44%	32%	8%	8%	4%	4%
Вугілля	0,76	0,44	0,94	2,92	8,44	22,46
Нафта та нафтопродукти	0,455	0,542	2,758	6,434	1,343	1,273
Газ природний	1,535	1,705	5,615	3,476	6,992	8,355
Електроенергія	0,344	0,138	0,652	0,314	0,256	0,612

\* Власні розрахунки авторів.

**Кластер 1.** В цьому кластері 11 з 25 регіонів України, а саме: Івано–Франківська, Житомирська, Вінницька, Сумська, Черкаська, Волинська, Чернігівська, Запорізька, Чернівецька, Хмельницька області та АР Крим. Незважаючи на те що це найбільший з кластерів, він знаходиться на передостанньому місці за обсягами споживання вугілля та на останньому – за споживанням природного газу. Середнє використання нафти та нафтопродуктів регіонами з даного кластеру є найнижчим.

**Кластер 2** складається з восьми областей: Львівська, Рівненська, Херсонська, Закарпатська, Миколаївська, Тернопільська, Харківська, Кіровоградська. Його особливість полягає в найменших показниках споживання вугілля (0,44 млн. т умовного палива (т.у.п.) у середньому по кластеру) та електроенергії (0,14 млн. т.у.п. у середньому по кластеру). Він знаходиться на п'ятому місці за середнім обсягом використання нафти та нафтопродуктів і природного газу.

**Кластер 3** є третім за розмірами. До його складу входять два регіони – це Одеська та Київська області. Для нього характерний найвищий середній показник споживання електроенергії 0,652 млн. т у.п. Займає другу сходинку за споживанням нафти та третю – за використанням природного газу.

**Кластер 4.** Має однакові розміри з попереднім кластером. Його представниками є Полтавська та Луганська області. Є лідером за споживанням сирової нафти та нафтопродуктів – 6,43 млн. т.у.п. у середньому по кластеру. Також займає третю сходинку у використанні вугілля. Кластер не виділяється показниками споживання електроенергії та природного газу.

Наступні два кластери представлені поодинокими регіонами.

**Кластер 5.** Кластер №5 представляє собою Дніпропетровську область. Даний регіон є другим за розмірами споживачем вугілля та природного газу. Знаходиться на п'ятому місці за показником використання електроенергії.

**Кластер 6.** Його представником є Донецька область, яка має найбільші обсяги споживання вугілля і природного газу. Кластер посідає другу сходинку у використанні електроенергії.

Оскільки області, що входять у два останні кластери, мають яскраво виражену або домінуючу ознаку над іншими, такі аномалії можуть вплинути на кластеризацію. Тому зробимо повторну кластеризацію за виключенням Дніпропетровської та Донецької області.

У результаті повторної кластеризації 23 регіонів отримали карту, яка складається з семи кластерів (рис. 3).

На рис. 4 наведені карти Кохонена, побудовані для кожної змінної.

Для характеристики кожного кластеру необхідно продемонструвати статистику середніх величин по кожному з семи кластерів (табл. 2).

Зупинимось детальніше на кожному кластері.

**Кластер 1.** До його складу увійшли: Вінницька, Запорізька, Івано–Франківська, Львівська, Харківська та Черкаська області. Кластер є третім за середніми обсягами споживання вугілля та п'ятим у використанні природного газу і нафти.

**Кластер 2.** Є найбільшим за розмірами і включає сім регіонів: АР Крим, Волинську, Житомирську, Сумську, Чернівецьку, Чернігівську та Хмельницьку області. Для представників цього кластеру характерним є шосте місце за середніми обсягами використання природного газу і нафти та нафтопродуктів.



Рисунок 3. Карта 23 регіонів України

\* Власні розрахунки авторів.

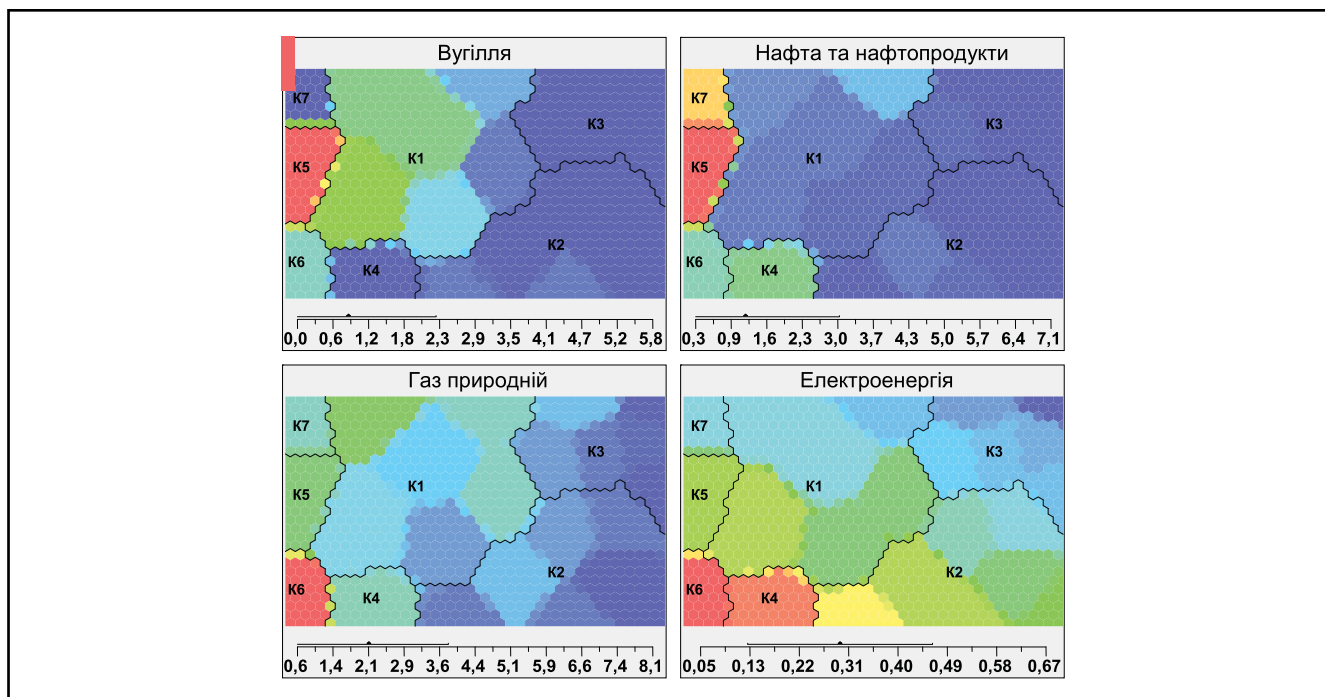


Рисунок 4. Карти Кохонена для окремих змінних

\* Власні розрахунки авторів

Таблиця 2. Середні величини по кластерах

Кластер	1	2	3	4	5	6	7
Питома вага кластерів	26,09%	30,43%	26,09%	4,35%	4,35%	4,35%	4,35%
Вугілля	1,754	0,138	0,062	0,087	5,819	1,803	0,031
Нафта та нафтопродукти	0,742	0,385	0,367	3,019	7,056	2,497	5,812
Газ природний	2,648	1,153	1,094	3,113	3,972	8,117	2,981
Електроенергія	0,2733	0,3586	0,1233	0,636	0,415	0,668	0,213

\* Власні розрахунки авторів

**Кластер 3.** Кластер №3 має однаковий з першим кластером розмір. Складається з таких областей, як Закарпатська, Кіровоградська, Миколаївська, Рівненська, Тернопільська і Херсонська. Його особливість полягає в тому, що середні обсяги спожитої електроенергії, природного газу та нафти є найнижчими серед кластерів.

Наступні чотири кластери складаються з одного регіону.

**Кластер 4** містить Одеську область. Обсяг споживання вугілля даного кластеру становить 0,087 млн. т.у.п. – це п'ятий результат. За використанням нафти та нафтопродуктів і природного газу є третім серед кластерів. Витрати електроенергії у обсязі 0,636 млн. т.у.п. дають кластеру другу сходинку.

**Кластер 5** – це Луганська область. Показники споживання вугілля та нафти є найвищими серед 23 регіонів. За обсягом використання природного газу поступається лише шостому кластеру.

**Кластер 6.** Представлений Київською областю. Кластер є лідером у споживанні природного газу та електроенергії, відповідні обсяги становлять 8,117 і 0,668 млн. т. у. п.

**Кластер 7.** До його складу входить лише Полтавська область. Серед кластерів демонструє найнижчий обсяг споживання вугілля (0,031 млн. т. у. п.). Використання елек-

тричної енергії теж є одним з найнижчих поряд із другим показником витрат нафти та нафтопродуктів.

Самоорганізована карта Кохонена виконує проектування багатовимірних даних у простір з меншою розмірністю (звичай двовимірний) і застосовується на практиці при візуалізації даних, щоб людина могла «побачити» наявність або відсутність кластерної структури в даних, число кластерів, закони сумісного розподілу ознак, залежності між змінними.

### Висновки

У результаті застосування технології самоорганізованих карт вся база регіонів України була перетворена у чотирих-вимірний простір, який карта розподілила на шість економічних кластерів. Серед них були області, що являлися єдиними представниками свого кластеру. В ході повторної кластеризації без особливих регіонів було виявлено сім кластерів, склад яких також відреагував на зміну бази даних. Тобто ці кластери не є однорідними і зв'язок у середині них не сталим.

Запропонована технологія самоорганізованих карт Кохонена дозволяє забезпечити у даному випадку ефективне використання енергоресурсів, реалізація цієї методики в інших галузях економіки відповідно інших видів ресурсів.

**Список використаних джерел**

1. Автономів А.Б. Світова енергетика: стан, масштаби, перспективи, стійкість розвитку, проблеми екології, цінова динаміка паливно-енергетичних ресурсів [Текст] / А.Б. Автономів // Електричні станції. – 2003. – №5. – С. 55–64.
2. Аракелов В.Е. Методичні питання економії енергоресурсів [Текст] / В.Е. Аракелов, А.И. Кремер. – М.: Энергоатомиздат, 2005. – 189 с.
3. Боровиков В.П. Прогнозирование в системе STATISTICA в среде Windows [Текст] / В.П. Боровиков. – М.: Финансы и статистика, 2004. – С. 35–37.
4. Войотович А.Д. Кластеризація регіонів за рівнем соціально-економічного розвитку [Текст] / А.Д. Войотович // Прикарпатський юридичний інститут Львівського державного університету внутрішніх справ. – 2009. – №2. – С. 20–29.
5. Дмитрівський А.Н. Природний газ в XXI столітті [Текст] / А.Н. Дмитрівський // Нафтове господарство – 2010. – №12. – С. 14–17.
6. Дюран Б.Д. Кластерный анализ [Текст] / Б.Д. Дюран. Пер. с англ. Е.З. Демиденко. Под ред. А.Я. Боярского. Предисловие А.Я. Боярского. – М., «Статистика», 1977. – 128 с.
7. Ежов А.А. Нейрокомпьютинг и его применение в экономике и бизнесе [Текст] / А.А. Ежов, С.А. Шумский. – М.: МИФИ.
8. Погрішук Б.В. Кластеризація регіонів за показниками інноваційно-інвестиційної діяльності зерно продуктового підкомплексу України [Текст] / Б.В. Погрішук. – 2010. – С. 34–45.
9. Сайт Алгоритмы кластеризации на службеDataMining. [Електрон. ресурс]. – Режим доступу: <http://www.basegroup.ru>
10. Сайт Интеллектуальный анализ данных. [Електрон. ресурс] – Режим доступу: <http://itnews.com.ua>
11. Сайт Кластеризация. Итеративные и плотностные алгоритмы. [Електрон. ресурс] – Режим доступу: <http://wiki.auditory.ru>
12. Український статистичний збірник. Статистичний збірник / Держкомстат України – К.: – 2010. – 414 с.
13. Хайкин С. Нейронные сети: полный курс [Текст] / С. Хайкин. – 2-е издание.: Пер. с англ. – М.: «Вильямс», 2006. – 1104 с.: ил. – Парал. тит. англ.