

ВИКОРИСТАННЯ КЛАСТЕРНОГО АНАЛІЗУ ЗА ДОПОМОГОЮ КАРТ КОХОНЕНА, ЩО САМООРГАНІЗУЮТЬСЯ, ДЛЯ ДОСЛІДЖЕННЯ ІННОВАЦІЙНОЇ АКТИВНОСТІ ПІДПРИЄМСТВ УКРАЇНИ

ДВНЗ “Український державний хіміко-технологічний університет”, м. Дніпро

Запропоновано та обґрунтовано нейромереживий підхід до кластеризації підприємств України за рівнем їх інноваційної активності. Кластерний аналіз являється однією із задач класу Data Mining і дозволяє виявляти в наборі необроблених даних приховані закономірності з метою відстежування тенденцій, зокрема, у інноваційній політиці промислового підприємства. Проаналізовано переваги нейромережевого методу кластеризації даних за допомогою карт Кохонена, що самоорганізуються, у порівнянні з класичними статистичними методами. Визначено кластери за видами інновацій, на які можливо розподілити підприємства України різних видів економічної діяльності. Показано, що основним видом інновацій, які впроваджують підприємства, є технологічні, маркетингові та організаційні інновації. Виявлено тенденції щодо спрямування інновацій на підприємствах певних видів економічної діяльності. У якості вихідних даних було обрано дані обстеження інноваційної діяльності в економіці України, здійсненого за міжнародною методологією. У якості інструмента нейромережевого підходу до аналізу інноваційної активності підприємств застосовано аналітичну платформу DEDUCTOR Academic версії 5.0 компанії BaseGroup_Labs.

Ключові слова. Інноваційна активність підприємств; нейронні мережі, карти Кохонена, що самоорганізуються, кластеризація, Data Mining.

Вступ та актуальність дослідження

Інноваційна активність має на сьогодні вирішальне значення в економічному розвитку сучасної економіки. Результати оцінювання інноваційної активності підприємств країни є необхідними як будь-якому інвестору, так, власне, і для підприємства. Основою існуючого інструментарію оцінювання інноваційної активності підприємств на даний час є статистичні та експертні методи аналізу даних. Через притаманні їм недоліки дані методи приводять до високого ступіня суб'єктивізму та часткової неспроможності вирішення задач моделювання складних соціально-економічних явищ. До основних недоліків даних методів можливо віднести багатокритеріальність, неповноту та неоднорідність вхідної інформації. Математична статистика в основному корисна при перевірці наперед сформульованих гіпотез (*verification-driven data mining*) [1].

Аналіз досліджень і публікацій

Використання статистичних методів аналізу інноваційної активності підприємств поши-

рене у сучасних наукових роботах. Серед останніх можемо відмітити роботу К.О. Бояринової, присвячену дослідженню інноваційної діяльності підприємств машинобудування в переробній промисловості [2]. В даній роботі побудовані регресійні моделі впливу дохідної та витратної частин інноваційної діяльності підприємств машинобудування на валовий прибуток підприємств переробної промисловості. У той же час потрібно відмітити появу наукових праць, де можливо спостерігати спроби застосування існуючих методів інтелектуального аналізу даних, зокрема, методу нечіткої логіки [3,4,5].

Виклад основного матеріалу

Виокремлюють наступні основні методи інтелектуального аналізу даних: нейронні мережі, генетичні алгоритми, нечітка логіка, дерева рішень, еволюційне програмування та ін. Метод нейронних мереж використовується для класифікації, кластеризації, прогнозування та розпізнавання образів. Серед моделей нейронних мереж існує три основних типи, а саме:

1) найбільш розповсюджені мережі прямо-

го розповсюдження (backpropagation), їх використання обмежено прогнозуванням економічних показників;

2) мережі зі зворотним зв'язком, наприклад, модель Хопфілда, які використовуються для оптимізації розрахунків;

3) мережі, що самоорганізуються (наприклад, модель Кохонена), які використовуються для кластерного аналізу.

Модель Кохонена належить до мереж, що навчаються без вчителя. Даний метод дослідження не потребує будь-яких апріорних припущень про розподіл даних. Важливі для нас області застосування мереж Кохонена – це розвідувальний аналіз даних і виявлення нових явищ. Окрім кластеризації даних мережі Кохонена вирішують задачі візуалізації даних, зниження їх розмірності та виявлення нових взаємозв'язків у великих наборах багатовимірних даних.

Першою фазою аналізу даних (*data minig*) є підготовка даних, яка, на думку багатьох дослідників, має вирішальне значення у всьому процесі аналізу даних [1]. Для нашого дослідження

у якості вихідних даних ми використовували результати обстеження інноваційної діяльності за 2012–2014 рр. [6,7,8], що містяться у щорічних статистичних збірниках, підготовлених фахівцями відділу статистики та інновацій Державної служби статистики (табл. 1). Основою наведених у цих збірниках даних є статистичні спостереження за формами: № 3 – наука “Звіт про виконання наукових та науково-технічних робіт”, № 1 – інновація “Обстеження інноваційної діяльності промислового підприємства”, № ІНН – “Обстеження інноваційної діяльності підприємства за період 2012–2014 рр.”.

На думку деяких дослідників, використання даних показників утруднене через їхню низьку достовірність та часткову відсутність даних на підприємствах. Окрім того, показники статистики інновацій непорівнянні з міжнародними стандартами та показниками, що публікуються статистичними відомствами розвинених країн. На підставі вітчизняного та закордонного досвіду, автори [9] пропонують застосовувати два методи – метод “чорного ящика” та метод, зап-

Таблиця 1

Розподіл кількості обстежених підприємств за видами економічної діяльності та за видами інновацій у 2012–2014 р. [6,7,8]

	загальна кількість обстежених підприємств	У тому числі								
		підприємства з інноваційною активністю	з них							підприємства тільки з маркетинговими та/або організаційними інноваціями
			підприємства з технологічними інноваціями	з них					підприємства з продовжувальною або перваною інноваційною діяльністю	
				підприємства з продуктовими інноваціями	підприємства з процесовими інноваціями	підприємства з продуктовими та процесовими інноваціями	підприємства з			
Усього	27992	4084	2663	446	1003	1008	206	1421	23908	
Промисловість	13529	2492	1888	360	646	731	151	604	11038	
Добувна промисловість і розроблення кар'єрів	413	49	38	2	27	4	5	11	364	
Переробна промисловість	11401	2221	1674	350	490	708	126	547	9180	
Постачання електроенергії, газу, пари та кондиційованого повітря	564	105	88	3	63	10	12	17	459	
Водопостачання, каналізація, поводження з відходами	1151	117	88	5	66	9	8	29	1034	
Послуги	14464	1592	775	86	357	277	55	817	12871	
Оптова торгівля, крім торгівлі автотранспортними засобами та мотоциклами	7592	850	308	38	154	90	26	542	6742	
Транспорт, складське господарство, поштова та кур'єрська діяльність	3646	267	160	6	101	42	11	107	3379	
Інформація і телекомунікації	1757	287	186	28	49	102	7	100	1470	
Діяльність у сфері архітектури та інжинірингу, технічні випробування та дослідження	1468	188	121	14	53	43	11	67	1280	

Відносна кількість обстежених підприємств, що займались інноваціями у 2012–2014 рр. за видами економічної діяльності та за видами інновацій, %

	підприємства з продуктивними інноваціями	підприємства з процесовими інноваціями	підприємства з продуктивними та процесовими інноваціями	підприємства з продовжувальною або перерваною інноваційною діяльністю	підприємства з продовжувальною або перерваною інноваційною діяльністю
Добувна промисловість і розроблення кар'єрів	0,48	6,54	0,97	1,21	2,66
Переробна промисловість	3,07	4,30	6,21	1,11	4,80
Постачання електроенергії, газу, пари та кондиційованого повітря	0,53	11,17	1,77	2,13	3,01
Водопостачання, каналізація, поводження з відходами	0,43	5,73	0,78	0,70	2,52
Оптова торгівля, крім торгівлі автотранспортними засобами та мотоциклами	0,50	2,03	1,19	0,34	7,14
Транспорт, складське господарство, поштова та кур'єрська діяльність	0,16	2,67	1,15	0,30	2,98
Інформація і телекомунікації	1,59	2,79	5,81	0,40	5,75
Діяльність у сфері архітектури та інжинірингу, технічні випробування та дослідження	0,95	3,61	2,93	0,75	4,56

ропований О.В. Нікітіною.

Інноваційну активність підприємств окремих видів економічної діяльності можливо оцінювати, поряд з іншими показниками, за розподілом підприємств за видами економічної діяльності та за видами інновацій. Статистичні дані, обрані нами як вхідні, були взяті зі статистичних збірників “Наукова та інноваційна діяльність в Україні” за 2014–2016 рр. [6,7,8]. Потрібно зауважити, що звітність, залучена нами дослідження, включала результати обстеження інноваційної активності підприємств, яке було проведено за міжнародною методологією.

За рівень інноваційної активності звичайно приймається кількість інноваційно активних підприємств. З метою співставлення інноваційної активності підприємств різних видів економічної діяльності на підставі даних табл. 1 була визначена відносна кількість підприємств, що займались інноваційною діяльністю у 2012–2014 рр., для кожного виду економічної діяльності та виду інновацій (табл. 2).

У якості програмного середовища для проведення досліджень використано аналітичну платформу *DEDUCTOR Academic* версії 5.0 компанії *BaseGroup_Labs*. Параметри навчання карти Кохонена надані у табл. 3.

Результати моделювання можна побачити на рисунку. Для аналізу результатів кластерування було використано не тільки карти вхідних векторів, а й матриця відстаней, матриця помилок квантування та матриця густини

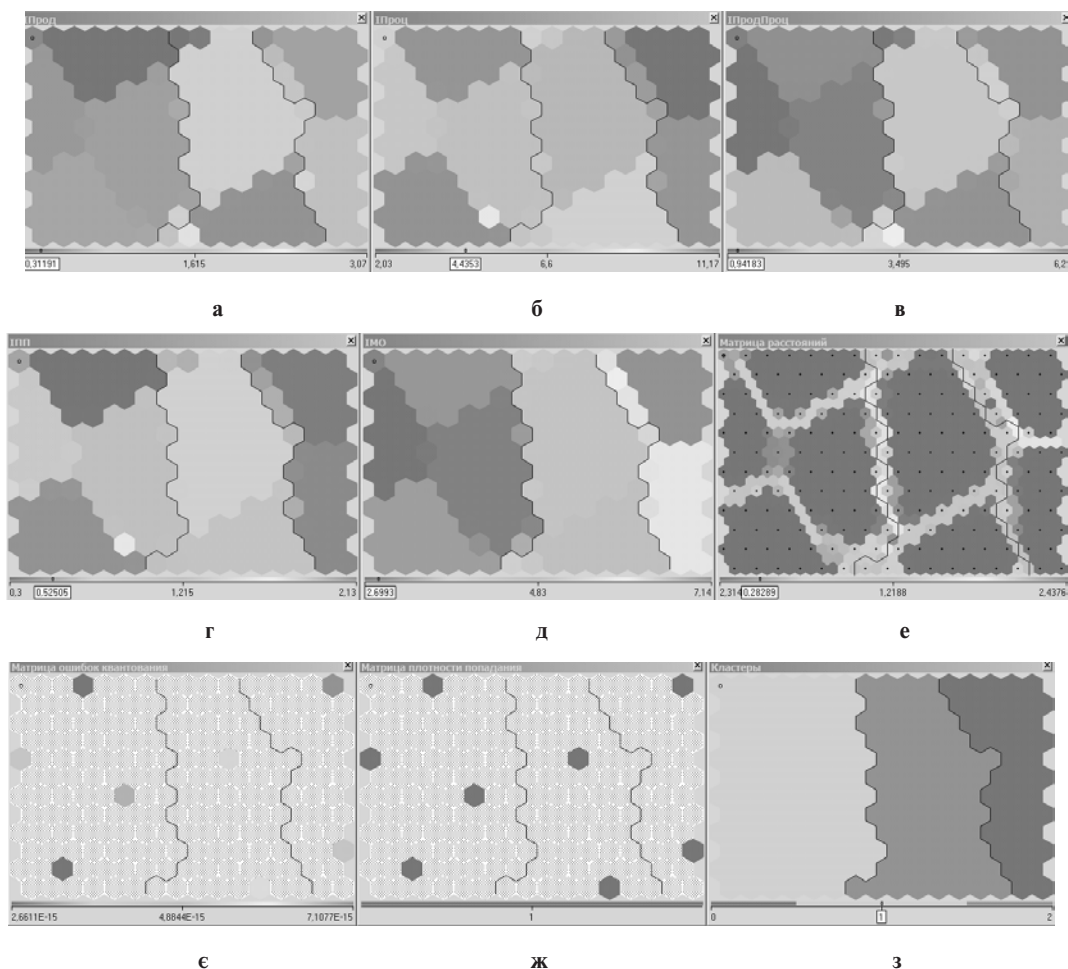
попадання.

Параметри навчання карти Кохонена

Параметри	Значення параметрів
Розмір карти Кохонена	16×12
Форма осередків	стілники
Кількість епох (циклів) при наближеному налагодженні	500
Кількість епох (циклів) при тонкому налагодженні	500
Швидкість навчання при наближеному налагодженні	0,3
Швидкість навчання при тонкому налагодженні	0,006
Ініціалізація ваг	нормально розподілені випадкові величини
час навчання	5 хв
Помилка поділу даних	0,12
Топографічна помилка	0,21

Кластер являє собою номінальну частинку в атласі карт Кохонена і включає в себе одну чи кілька однакових інформаційних позицій [9,10,11]. Аналіз карти “Матриця відстаней” дозволив виділити кластери, які позначені цифрами 0; 1 та 2 (рис. 1,3).

За результатами розрахунків карти Кохонена побудована табл. 4, у якій можливо спостерігати розподіл галузей економіки за трьома кластерами. Найбільш чітко можливо ідентифі-



Атлас карт Кохонена з розподілом за кластерами підприємств, що впроваджують інновації: а) продуктові (ІПрод); б) процесові (ІПроц); в) продуктові та процесові (ІПродПроц); г) підприємства з перерваними або подовженими інноваціями (ІПП); д) маркетингові або організаційні (ІМО); характеристиками моделі у вигляді матриць: е) матриця відстаней; ж) матриця помилок квантування; з) матриця густини попадання та з) кластери вхідних значень

кувати кластер 1, потім кластер 2. Кластер 1 займає 46% всієї площі на рис. 1, з (88 комірок з 192 (12Ч16)), кластер 2 займає 34% площі (65 комірок з 192), найменший за розміром кластер 0 займає решту площі у 20% (39 комірок з 192).

Аналіз даних табл. 3 чітко, на наш погляд, показує розподіл підприємств України за видами економічної діяльності та за видами інновацій на три кластери. До найменшого кластеру 0 надійшли підприємства, які займаються оптовою торгівлею, окрім торгівлі автотранспортними засобами та мотоциклами, переробкою інформації та телекомунікаційними послугами. Підприємства цього кластеру віддають перевагу, насамперед, маркетинговим та організаційним інноваціям у порівнянні з процесовими інноваціями.

До найбільшого за обсягом кластеру 1 увійшли галузі, які займаються транспортними послугами, у тому числі послугами з транзиту електроенергії, газу та нафти, водопостачанням, транспортом, складським господарством, поштовою та кур'єрською діяльністю. Це найбільший

кластер за кількістю підприємств, що займаються інноваціями. Послуги підприємств цих галузей мають для країни стратегічне значення, підприємства потребують високотехнологічного переозброєння.

До кластеру 2 увійшли підприємства тих видів діяльності, які не увійшли до кластерів 0 та 1, і які випускають орієнтовану, насамперед, на експорт продукцію. Підприємства цих галузей мають кошти на оновлення застарілого устаткування та придбання устаткування, що випускає конкурентоспроможну продукцію, а також вкладають гроші в удосконалення існуючих технологій, здійснюють маркетингові та організаційні інновації.

Можливо констатувати, що найбільшу інноваційну активність щодо заміни основних фондів та удосконалення технологій проявляють підприємства саме тих галузей, що видобувають та переробляють сировину.

Таким чином, неймережевий підхід за допомогою карт Кохонена дозволяє розподіли-

Параметри кластерів, отримані в результаті нейронної обробки вхідної інформації

Номер кластеру	Підприємства, що займаються певними видами економічної діяльності, які увійшли до кластеру	Параметри кластеру
0	Оптова торгівля, окрім торгівлі автотранспортними засобами та мотоциклами, інформація та телекомунікації	До кластеру входять підприємства, які характеризуються високим рівнем маркетингових та організаційних інновацій, низьким рівнем процесових інновацій та низьким рівнем подовженої або перерваної інноваційної діяльності
1	Постачання електроенергії, газу, пари та кондиційованого повітря, водопостачання, каналізація, з відходами, транспорт, складське господарство, поштова та кур'єрська діяльність	До кластеру входять підприємства, які одночасно характеризуються середнім рівнем маркетингових та організаційних інновацій, низьким рівнем продуктових інновацій, низьким рівнем процесових інновацій та низьким рівнем подовженої або перерваної інноваційної діяльності
2	Добувна промисловість і розроблення кар'єрів, переробна промисловість, діяльність у сфері архітектури та інжинірингу, технічні випробування та дослідження	До кластеру входять підприємства, що одночасно впроваджують процесові, продуктові, маркетингові та організаційні інновації,

ти підприємства України на кластери за видами економічної діяльності та за видами інновацій. Це допомагає розуміти специфіку впровадження інновацій на підприємствах кожного кластеру та розробляти для кожного підприємства відповідного кластеру стратегію інноваційного розвитку.

Висновки

Здійснене дослідження дозволило зробити наступні висновки:

1) нейромережвий підхід до кластеризації вихідних статистичних даних у порівнянні з класичними статистичними методами дозволяє виявляти приховані закономірності у зв'язках між вхідними даними;

2) досліджено параметри нейромережової моделі кластеризації підприємств різних видів діяльності та за видами інновацій і показано, що підприємства переробної промисловості та підприємства, що займаються постачанням електроенергії та газу, лідирують серед підприємств усіх галузей за рівнем впровадження інновацій (19,5 та 18,5%, відповідно) серед усіх підприємств відповідних галузей та активно впроваджують продуктові, процесові, маркетингові та організаційні інновації.

4) серед найбільш вагомих причин не здійснювати інновації протягом 2012–2014 рр. були: низький попит на інновації на ринку, впроваджені раніше інновації, низька конкуренція підприємства на ринку, відсутність фінансових можливостей для здійснення інновацій.

5) подальші розробки автора у цьому напрямку будуть спрямовані на удосконалення параметрів нейронної мережі з метою зниження рівня помилок і підвищення точності прогнозів.

СПИСОК ЛІТЕРАТУРИ

1. Манжула В.Г., Федяшев Д.С. Нейронные сети Кохонена и нечёткие нейронные сети в интеллектуальном анализе данных // Фундаментальные исследования. Технические науки. – 2016. – № 4. – С.108-114.
2. Бояринова К.О. Функціональність інноваційної діяльності підприємств машинобудування в переробній промисловості // Актуальні Проблеми Економіки. – 2016. – № 8(182). – С.352-360.
3. Modelling and forecasting of innovative development of entrepreneurial structures under the global competition // Tatyana L. Bezrukova, Aidarbek T. Giyazov, Aliia M. Bazieva // Actual problems of economic. – 2016. – № 8(82). – P.344-351.
4. М'ячин В.Г. Прогнозування інноваційного потенціалу промислових підприємств за допомогою гібридних інтелектуальних систем // Наук. вісник Ужгородського нац. ун-ту. Серія: Міжнародні економічні відносини та світове господарство. – Ужгород. – 2015. – Вип. 4. – С.51-56.
5. М'ячин В.Г., Куцинська М.В. Визначення пріоритетності показників фінансового стану, обраних для оцінки ризику банкрутства промислових підприємств методом нечітких множин // Наук. вісник Міжнар. гуманітарного ун-ту. Серія: Економіка і менеджмент. – 2015. – Вип.14. – С.159-163.
6. Наукова та інноваційна діяльність в Україні, 2015 рік. Статистичний збірник / за ред. О.О. Кармазіної. – К.: Державна служба статистики. – 2016. – 257 с.
7. Наукова та інноваційна діяльність в Україні у 2014 році. Статистичний збірник / за ред. О.О. Кармазіної. – К.: Державна служба статистики. 2015. – 255 с.
8. Інноваційна Україна 2020: національна доповідь / за заг. ред. В.М. Гейця та ін. – К.: НАН України, 2015. – 336 с.
9. Колос Е.А., Денисова О.К., Касенова А.Е. Выбор и апробация методик оценки инновационной активности на примере предприятий Восточно-Казахстанской области // Актуальні проблеми економіки. – 2016. – № 8(82). – С.124-

133.

10. Kohonen T. Essentials of the self-organizing map // Neural Networks. – № 37. – P.52-65.

11. Дебок К., Кохонен Т. Анализ финансовых данных с помощью самоорганизующихся карт: пер. с англ. – М.: Издательский Дом "Альпина", 2001. – 317 с.

12. Кохонен Т. Самоорганизующиеся карты. – Бином, 2014. – 659 с.

Надійшла до редакції 30.10.2016
Рецензент: д.е.н., доц. Л.Д. Гармідер

ПРИМЕНЕНИЕ КЛАСТЕРНОГО АНАЛИЗА С ПОМОЩЬЮ САМООРГАНИЗУЮЩИХСЯ КАРТ КОХОНЕНА ДЛЯ ИССЛЕДОВАНИЯ ИННОВАЦИОННОЙ АКТИВНОСТИ ПРЕДПРИЯТИЙ УКРАИНЫ

М'ячин В.Г.

Предложен и обоснован нейросетевой подход к кластеризации предприятий Украины по видам инновационной активности. Кластерный анализ является одной из задач класса Data Mining и позволяет выявлять в наборе необработанных данных скрытые закономерности с целью отслеживания тенденций, в частности, в инновационной политике государства. Проанализированы преимущества нейросетевого метода кластеризации данных с помощью самоорганизующихся карт Кохонена в сравнении с классическими статистическими методами. Определены кластеры по видам инноваций, благодаря чему стало возможным распределять предприятия Украины различных видов экономической деятельности. Показано, что основным видом инноваций, которые внедряют предприятия, являются процессные, маркетинговые и организационные инновации. Выявлены тенденции по внедрению инноваций на предприятиях разных видов экономической деятельности. В качестве исходных данных были выбраны данные обследования инновационной деятельности предприятий Украины, проведенного по международной методологии. В качестве инструмента нейросетевого подхода к анализу инновационной активности предприятий применена аналитическая платформа DEDUCTOR Academic версии 5.0 компании BaseGroup_Labs.

Ключевые слова. Инновационная активность предприятий, нейронные сети, самоорганизующиеся карты Кохонена, кластеризация, Data Mining.

THE USE OF CLUSTER ANALYSIS USING KOHONEN SELF-ORGANIZING MAPS FOR THE STUDY OF INNOVATION ACTIVITY OF ENTERPRISES IN UKRAINE

Myachin V.G.

It proposed and justified neural network approach to clustering of Ukrainian enterprises by their innovative activity. Cluster analysis is one of the tasks of Data Mining class. Cluster analysis allows you to identify a set of raw data hidden patterns in order to track trends, in particular in the innovation policy of the state. The advantages of neural network method of data clustering using Kohonen self-organizing maps in comparison with classical statistical methods. Clusters are defined by type of innovation. It is shown that the main type of innovation that implement enterprise is processing, marketing and organizational innovations. Trends in innovation have been identified in the enterprises of different economic activities. The initial data survey data innovation enterprises of Ukraine have been chosen, conducted by the international methodology. As a tool neural network approach to the analysis of innovation activity of enterprises applied analytical platform DEDUCTOR Academic version 5.0 of BaseGroup_Labs.

Keywords: innovative activity of enterprises, neural networks, Kohonen self-organizing maps, clustering, Data Mining.