

## І Н Ф О Р М А Т И К А

УДК 519.21

### КЛАСТЕРНИЙ АНАЛІЗ ОЦІНКИ СЕРЕДНЬОГО РІВНЯ ЗНАНЬ ЗА РЕГІОНАМИ УКРАЇНИ (РЕЗУЛЬТАТИ ЗОВНІШНЬОГО НЕЗАЛЕЖНОГО ОЦІНЮВАННЯ)

Н. Задорожня, Я. Єлейко, Ю. Щербина

*Львівський національний університет імені Івана Франка,  
вул. Університетська, 1, Львів, 79000*

Вивчено використання кластерного аналізу в оцінюванні освітнього рівня України. Проведено кластеризацію областей України за результатами зовнішнього незалежного оцінювання чотирьох предметів за 2009-2013 рр. Проаналізовано отримані групи однорідних областей та порівняно результати кластеризації за допомогою порівняльного графіка коефіцієнтів варіації.

*Ключові слова:* кластерний аналіз, кластер, метод  $k$ -середніх, ЗНО, коефіцієнт варіації.

#### 1. ВСТУП

У світі існує практика проведення тестового контролю та діагностики для оцінювання рівня знань випускників шкіл і тестування абітурієнтів для прийому у вищі навчальні заклади. Наприклад, у США це SAT (з 1901) і ACT (з 1959), більшість країн Європи застосовують тест Matura для отримання атестату зрілості. Результати таких тестувань активно застосовують для аналізу рівня знань школярів, а також для знаходження слабких і сильних місць у національній системі середньої освіти для подальшого підвищення надійності та валідності тесту.

Український аналог такого оцінювання рівня знань – зовнішнє незалежне оцінювання (ЗНО). У 2006 р. розпочав свою діяльність Український центр оцінювання якості освіти. Тоді ж вперше провели масштабне добровільне тестування школярів. З 2008 р. проходження ЗНО є обов'язковою умовою вступу до вищого навчального закладу. У 2010 р. кількість потрібних для складання предметів збільшилася. Крім обов'язкових, української мови та літератури, а також історії України чи математики, наступного року абітурієнти проходитимуть випробування з ще одного предмета [1]. Досвід України в емпіричному дослідженні тестового контролю рівня знань учнів незначний порівняно з багаторічним досвідом країн Європи і США, тому система зовнішнього незалежного оцінювання, відповідно, перебуває на стадії корекції та удосконалення.

Сьогодні популярно використовувати матеріали Українського центру оцінювання якості освіти (регіональних даних) і різних аналізів результатів ЗНО, зокрема для статистичного аналізу тестів, для складання рейтингів середніх загальноосвітніх навчальних закладів.

Розглянемо класифікацію областей України за результатами ЗНО чотирьох основних предметів: української мови і літератури, історії України, математики та фізики. Проведемо аналіз результатів тестування кожного регіону України і класифікуємо їх за ступенем однорідності. Цю задачу можна розв'язати, використовуючи теорії розпізнавання образів, до якої належить кластерний аналіз.

## 2. ВИКЛАДЕННЯ ТЕОРЕТИЧНОГО МАТЕРІАЛУ

Кластерний аналіз – це сукупність методів класифікації багатомірних спостережень або об'єктів, заснованих на визначенні поняття відстані між об'єктами з наступним виділенням з них груп, “згустків” спостережень (кластерів, таксонів) [1].

На відміну від інших методів, цей вид аналізу дає змогу класифікувати об'єкти не за однією ознакою, а за кількома одночасно. Для цього вводять відповідні показники, які характеризують певну міру близькості за всіма класифікаційними параметрами. Дія кластерного аналізу полягає у привнесенні структури в досліджувані об'єкти. Це означає, що методи кластеризації потрібні для виявлення структури в даних, які нелегко знайти при візуальному обстеженні або за допомогою експертів.

Зауважимо, що різні методи кластеризації породжують різні кластерні рішення для тих самих даних. Залежно від вхідних даних і бажаних результатів варто наперед визначитися з методами. Основні методи такі.

*Метод деревоподібної (ієрархічної) кластеризації.* Призначення цього алгоритму – об'єднати об'єкти у досить великі кластери, використовуючи деяку міру подібності або відстань між об'єктами. Типовим результатом такої кластеризації є ієрархічне дерево. Результати роботи всіх ієрархічних процедур зазвичай оформляють у вигляді так званої дендрограми [3].

Цей метод вважають найбільш показовим і зрозумілим, використовують для формування кластерів. Ці відстані можна визначити в одновимірному або багатовимірному просторі. Вибираємо спосіб обчислення відстаней між об'єктами серед основних мір відстаней [2-4].

Відстанями в цьому випадку можуть бути:

*евклідова відстань*

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

вважається найчастіше уживаною відстанню. Має певні переваги (наприклад, відстань між двома об'єктами не змінюється при введенні в аналіз нового об'єкта, який може виявитися викидом). Недоліком є те, що у випадках, коли ознаки виміряні у різних одиницях, зміна масштабу одиниць вимірювання може призвести до істотної зміни результатів класифікації;

*манхеттенівська відстань*

$$d(x, y) = \sum_{i=1}^n |x_i - y_i|.$$

Обчислювати цю відстань найпростіше, бо здебільшого призводить до таких самих результатів, як і евклідова відстань, але зменшується вплив окремих великих різниць (викидів) через те, що відстань обчислюють за простими різницями координат;

*відстань Чебишева*

$$d(x, y) = \sup\{|x_i - y_i|\}$$

застосовують, якщо хочуть позначити два об'єкти як різні, позаяк вони відрізняються довільною однією координатою (довільним одним виміром);

*степенева відстань*

$$d(x, y) = \sqrt[r]{\sum_{i=1}^n (x_i - y_i)^p}.$$

Тут  $r$  і  $p$  – параметри, визначені користувачем.

Використовують, коли хочуть прогресивно збільшити або зменшити вагу, що належить до розмірності, для якої відповідні об'єкти сильно відрізняються. Тут параметр  $p$  відповідальний за поступове зважування різниць по окремих координатах, а параметр  $r$  – за прогресивне зважування великих відстаней між об'єктами [2];

*відсоток невідповідності*

$$d(x, y) = \frac{\text{Кількість}(x_i \neq y_i)}{i}.$$

Таку відстань використовують у випадку, коли дані належать до номінативної шкали (це шкала найменувань, до вимірів належать, наприклад, національності, професії тощо).

*1-коефіцієнт кореляції Пірсона (1-Pearson  $r$ )* використовують, коли абсолютні значення та різниці між об'єктами несуттєві, а важливішою є наявність зв'язку між ними  $d(x, y) = 1 - r$ . Така міра чутлива до подібності профілів об'єктів. Отримують результати близькі до факторного аналізу – кластери наближаються до факторів.

Коли вибрали одну з відстаней, то можна продовжити деревоподібний кластерний аналіз, використавши одну зі стратегій кластеризації, які ще називаються правилами об'єднання. Вони, переглядаючи таблицю подібностей об'єктів, на кожному кроці послідовно об'єднують пару найбільш схожих об'єктів, доки на останньому кроці не утвориться один великий кластер, до якого увійдуть усі об'єкти.

Основні стратегії ієрархічної кластеризації [3-4].

*Стратегія найближчого сусіда (стратегія одиночного зв'язку)* відстань між двома кластерами визначають як відстань між двома найближчими об'єктами (найближчими сусідами). Стратегія пов'яже два кластери разом, коли будь-які два об'єкти в цих кластерах ближче один до одного за усі інші.

*Стратегія найвіддаленішого сусіда (стратегія повного зв'язку)* відстань між кластерами визначають як найбільшу відстань між двома об'єктами з різних кластерів (між найвіддаленішими сусідами). Стратегія добре працює, коли об'єкти реально належать до різних класів.

*Стратегія незваженого попарного середнього*, в якій відстань між двома кластерами визначають як середню відстань між всіма парами об'єктів у них. Метод ефективний у випадку реального об'єднання об'єктів у "кущі" та в "ланцюжки".

*Стратегія зваженого попарного середнього* відрізняється від попередньої тим, що при обчисленнях розмір відповідного кластера (кількості об'єктів, які він містить) використовують як ваговий коефіцієнт. Тому цю стратегію застосовують тоді, коли передбачають появу кластерів нерівного розміру.

*Стратегія Варда* використовує методи дисперсійного аналізу для оцінки відстаней між кластерами. Цей метод мінімізує суму квадратів для двох гіпотетичних кластерів, які можуть бути сформовані на кожному кроці процесу кластеризації. Вважають ефективною, але утворює кластери малого розміру.

*Стратегія незваженого центроїдного методу*. У цьому методі відстань між двома кластерами визначається як відстань між їхніми центрами тяжіння.

*Стратегія зваженого центроїдного методу (медіана)* – при обчисленнях використовують ваги для обліку різниці між розмірами кластерів (тобто числами об'єктів у них). Якщо є значні відмінності в розмірах кластерів, то надають перевагу зваженому методу.

*Кластеризацію методом  $k$ -середніх* використовують лише тоді, коли в дослідника вже є апріорні гіпотези стосовно кількості кластерів.

Дослідник наперед задає кількість кластерів  $k$  і метод кластеризації допомагає знайти ці кластери так, щоб вони максимально різнилися один від одного. Розв'язують задачу розбиттям  $n$  об'єктів на  $k$  ( $k < n$ ) однорідних у певному розумінні кластерів. На початковому етапі його реалізації вихідні об'єкти впорядковують і перші  $k$  об'єктів надалі розглядають як кластери, яким надають одиничні вагові коефіцієнти. Потім приймають  $(k+1)$ -й об'єкт і з'ясовують, до якого з наявних кластерів він найближчий. Цей кластер замінюють новим, розташованим у центрі ваги вихідного кластера й  $(k+1)$ -го об'єкта. Якщо точка  $(k+1)$ -й об'єкт рівновіддалена від декількох кластерів, то її поміщають до кластера з найменшим номером. Потім по чергово приєднують до наявних кластерів об'єкти, які залишилися [4].

Перевага цього методу – можливість перевірити статистичну значимість відмінностей між виділеними кластерами.

*Кластеризацію методом двоходового об'єднання* використовують найрідше, лише тоді, коли виникає потреба провести одночасну кластеризацію стовпчиків-об'єктів і рядків-спостережень [3].

### 3. ВИКЛАДЕННЯ ОСНОВНОГО МАТЕРІАЛУ

Проаналізуємо середній рівень знань за допомогою методу  $k$ -середніх, оскільки ця процедура кластеризації дає змогу розбити сукупність на визначену кількість груп (у цьому випадку розбивали сукупність на три групи: високого рівня, середнього рівня та низького рівня). Аналіз результатів тестування на підставі регіонів України виконали за допомогою програмного пакета для статистичного аналізу STATISTICA. Як критерій підбрано середній бал і дисперсію, що припадають на регіон. Усі розрахунки критеріїв проводили за даними Українського центру оцінювання якості освіти [3].

Проведемо аналіз результатів тестування ЗНО з обов'язкового предмета – української мови і літератури в 27 регіонах України за останні п'ять років і класифікуємо їх за ступенем однорідності. За результати кластеризації отримаємо об'єднання регіонів, які наведено в табл. 1.

Таблиця 1

Результат кластеризації результатів ЗНО з української мови та літератури за 2009 – 2013 роки

Кластер/ Рік	2009	2010	2011	2012	2013
Високий рівень	Закарпатський, Івано- Франківський, Львівський	Закарпатський, Івано- Франківський, Львівський	Закарпатський, Львівський, м. Київ	Закарпатський, Львівський, м. Київ	Закарпатський, Івано- Франківський, Львівський, Тернопільський, м. Київ, м. Севастополь

Середній рівень	Вінницький, Волинський, Житомирський, Київський, Полтавський, Рівненський, Сумський, Тернопільський, Хмельницький, Черкаський, Чернівецький, м. Київ	Вінницький, Волинський, Дніпропетр., Донецький, Житомирський, Запорізький, Луганський, Миколаївський, Одеський, Полтавський, Рівненський, Сумський, Тернопільський, Харківський, Хмельницький, Чернівецький, Чернігівський, м. Київ, м. Севастополь	АР Крим, Вінницький, Волинський, Донецький, Житомирський, Івано-Франківський, Київський, Одеський, Полтавський, Рівненський, Сумський, Тернопільський, Харківський, Хмельницький, Чернівецький, Чернігівський	Вінницький, Волинський, Дніпропетр., Донецький, Житомирський, Запорізький, Івано-Франківський, Київський, Луганський, Сумський, Тернопільський, Харківський, Хмельницький, Чернівецький, м. Севастополь	Вінницький, Волинський, Дніпропетр., Донецький, Київський, Одеський, Полтавський, Черкаський, Чернівецький, Чернігівський
Низький рівень	АР Крим, Дніпропетр., Донецький, Запорізький, Кіровоград., Луганський, Миколаївський, Одеський, Харківський, Херсонський, Чернігівський, м. Севастополь	АР Крим, Київський, Кіровоград., Херсонський, Черкаський	Дніпропетр., Запорізький, Кіровоград., Луганський, Миколаївський, Херсонський, Черкаський, м. Севастополь	АР Крим, Кіровоград., Миколаївський, Полтавський, Рівненський, Херсонський, Черкаський, Чернівецький	АР Крим, Житомирський, Запорізький, Кіровоград., Луганський, Миколаївський, Сумський, Херсонський

Як видно з табл. 1, у всі роки до кластера високого рівня з середнім балом 153 протягом п'яти років входили Закарпатський і Львівський регіони, Івано-Франківський і м. Київ входили в цей кластер у 2009 – 2010, 2013 роках і 2011 – 2013 роках, відповідно. У кластері середнього рівня з оцінкою в 151 бал перебували без зміни свого рівня регіони Вінниці, Волині, Одеси, Сум, Тернополя і Хмельницька. До низького рівня протягом п'яти років входили Кіровоградська і Херсонська області, середня оцінка кластера становить 149 балів. Автономна республіка Крим і Миколаївський регіон тільки один раз, у 2011 і 2010 роках, відповідно, увійшли до середнього рівня, всі інші роки вони входили до кластера низького рівня.

Таблиця 2

Результат кластеризації результатів ЗНО з історії України за 2009 – 2013 роки

Кластер/ Рік	2009	2010	2011	2012	2013
Високий рівень	Закарпатський, Івано-Франківський, Львівський, Тернопільський, Чернівецький	Закарпатський, Івано-Франківський, Львівський, Тернопільський, Хмельницький, Чернівецький, м. Київ	АР Крим, Закарпатський, Івано-Франківський, Львівський, Тернопільський, Хмельницький, Чернівецький, м. Київ	Донецький, Закарпатський, Івано-Франківський, Львівський, Харківський, м. Київ	Закарпатський, Івано-Франківський, Львівський, м. Київ

Середній рівень	Вінницький, Волинський, Київський, Рівненський, Харківський, Хмельницький, Черкаський, Чернігівський, м. Київ	АР Крим, Вінницький, Волинський, Дніпропетр., Донецький, Житомирський, Запорізький, Київський, Кіровоград., Одеський, Полтавський, Рівненський, Харківський, Черкаський, Чернігівський	Вінницький, Дніпропетр., Донецький, Житомирський, Запорізький, Київський, Кіровоград., Луганський, Миколаївський, Одеський, Полтавський, Рівненський, Сумський, Харківський, Черкаський, Чернігівський	АР Крим, Вінницький, Дніпропетр., Житомирський, Запорізький, Київський, Кіровоград., Луганський, Сумський, Тернопільський, Хмельницький, Чернівецький, Чернігівський	Вінницький, Донецький, Київський, Кіровоград., Луганський, Полтавський, Рівненський, Сумський, Тернопільський, Харківський, Хмельницький, Чернівецький, Чернігівський
Низький рівень	АР Крим, Дніпропетр., Донецький, Житомирський, Запорізький, Кіровоград., Луганський, Миколаївський, Одеський, Полтавський, Сумський, Херсонський, м. Севастополь	Луганський, Миколаївський, Сумський, Херсонський, м. Севастополь	Волинський, Херсонський, м. Севастополь	Волинський, Миколаївський, Одеський, Полтавський, Рівненський, Херсонський, Черкаський, м. Севастополь	АР Крим, Волинський, Дніпропетр., Житомирський, Запорізький, Миколаївський, Одеський, Херсонський, Черкаський, м. Севастополь

Аналогічно проаналізуємо результати тестування знань з історії України. Результати кластеризації наведено в табл. 2.

Як бачимо, до кластера високого рівня увійшли Закарпатська, Івано-Франківська і Львівська області з середньою оцінкою кластера в 152 бали. Тернопільський і Чернівецький регіони до високого рівня входили в перші три роки. Київ лише у 2009 р. був внесений до середнього кластера, куди усі п'ять років входили лише Вінницька, Київська і Чернігівська області з середнім балом 151. Харківський регіон можна зачислити до середнього рівня, оскільки він тільки один раз (у 2012 р.) увійшов до кластера високого рівня. Відповідно, до середнього рівня одноразово увійшла також Миколаївська область у 2011 р., яка чотири роки поспіль була в кластері низького рівня (середній бал кластера – 150 балів), разом з Херсонською областю, котра єдина з усіх регіонів була в низькому рівні усі п'ять років.

Таблиця 3

Результат кластеризації результатів ЗНО з математики за 2009 – 2013 роки

Кластер/ Рік	2009	2010	2011	2012	2013
Високий рівень	Закарпатський, Івано-Франківський, Львівський, Чернівецький, м. Київ	Закарпатський, Івано-Франківський, Львівський, Сумський, м. Київ	Івано-Франківський, Львівський, м. Київ	Закарпатський, Івано-Франківський, Львівський, Чернігівський, м. Київ	Дніпропетр., Івано-Франківський, Львівський, м. Київ

Середній рівень	АР Крим, Волинський, Дніпропетр., Донецький, Запорізький, Київський, Миколаївський, Одеський, Рівненський, Сумський, Харківський, Хмельницький, Черкаський	Вінницький, Волинський, Дніпропетр., Донецький, Запорізький, Київський, Миколаївський, Одеський, Рівненський, Херсонський, Хмельницький, Чернівецький, Чернігівський	АР Крим, Дніпропетр., Донецький, Закарпатський, Київський, Одеський, Рівненський, Сумський, Харківський, Хмельницький, Черкаський, Чернігівський	Дніпропетр., Донецький, Житомирський, Запорізький, Київський, Миколаївський, Одеський, Полтавський, Рівненський, Сумський, Харківський, Херсонський, Хмельницький, Черкаський, м. Севастополь	Донецький, Житомирський, Закарпатський, Запорізький, Київський, Кіровоград., Миколаївський, Одеський, Рівненський, Сумський, Харківський, Херсонський, Черкаський, Чернігівський
Низький рівень	Вінницький, Житомирський, Кіровоград., Луганський, Полтавський, Тернопільський, Херсонський, Чернівецький, м. Севастополь	АР Крим, Житомирський, Кіровоград., Луганський, Полтавський, Тернопільський, Харківський, Черкаський, м. Севастополь	Вінницький, Волинський, Житомирський, Запорізький, Кіровоград., Луганський, Миколаївський, Полтавський, Тернопільський, Херсонський, Чернівецький, м. Севастополь	АР Крим, Вінницький, Волинський, Кіровоград., Луганський, Тернопільський, Чернівецький	АР Крим, Вінницький, Волинський, Луганський, Полтавський, Тернопільський, Хмельницький, Чернівецький, м. Севастополь

У табл. 3 подано об'єднання регіонів України, які отримали за допомогою кластерного аналізу результатів оцінювання знань з математики.

Івано-Франківський, Львівський регіони і м. Київ становлять основу кластера високого рівня з середнім балом 154, до якого також три роки входила Закарпатська область (2009 – 2010 і 2012). Кластер середнього рівня (середній бал – 151 бал) усі 5 років містить Донецьку, Київську, Одеську та Рівненську області, а кластер низького рівня регіони Луганська та Тернополя з середнім балом – 151. Варто звернути увагу, що Запоріжжя у 2011 р., а також Хмельницьк і Черкаси у 2013 р. потрапили у кластер низького рівня, тоді як Вінницька область і м. Сімферополь, у 2010 та 2012 роках, відповідно, навпаки, потрапили в об'єднання регіонів, які належать до кластера середнього рівня.

Таблиця 4

Результат кластеризації результатів ЗНО з фізики за 2009 – 2013 роки

Кластер/ Рік	2009	2010	2011	2012	2013
Високий рівень	Запорізький, Івано-Франківський, Рівненський, Сумський, Чернівецький, Чернігівський, м. Київ, м. Севастополь	Івано-Франківський, Тернопільський, Херсонський, Черкаський, м. Київ	АР Крим, Запорізький, Івано-Франківський, Львівський, Рівненський, м. Київ	Дніпропетр., Івано-Франківський, Львівський, Херсонський, Чернігівський, м. Київ,	Івано-Франківський, Львівський, Рівненський, Тернопільський, Чернівецький, Чернігівський, м. Київ,

Середній рівень	АР Крим, Дніпропетр., Донецький, Житомирський, Київський, Львівський, Харківський, Хмельницький	Волинський, Дніпропетр., Закарпатський, Запорізький, Київський, Львівський, Миколаївський, Одеський, Харківський, Хмельницький, Чернігівський	Дніпропетр., Донецький, Закарпатський, Київський, Миколаївський, Одеський, Сумський, Тернопільський, Харківський, Хмельницький, Черкаський, Чернівецький, м. Севастополь	АР Крим, Вінницький, Київський, Кіровоград., Миколаївський, Одеський, Сумський, Тернопільський, Харківський, Хмельницький, Черкаський, м. Севастополь	АР Крим, Вінницький, Волинський, Дніпропетр., Житомирський, Запорізький, Київський, Миколаївський, Одеський, Сумський, Херсонський, Хмельницький, Черкаський,
Низький рівень	Вінницький, Волинський, Закарпатський, Кіровоград., Луганський, Миколаївський, Одеський, Полтавський, Тернопільський, Херсонський, Черкаський	АР Крим, Вінницький, Донецький, Житомирський, Кіровоград., Луганський, Полтавський, Рівненський, Сумський, Чернівецький, м. Севастополь	Вінницький, Волинський, Житомирський, Кіровоград., Луганський, Полтавський, Херсонський, Чернігівський	Волинський, Донецький, Житомирський, Закарпатський, Запорізький, Луганський, Полтавський, Рівненський, Чернівецький	Донецький, Закарпатський, Кіровоград., Луганський, Полтавський, Харківський, м. Севастополь

Фізика не належить до обов'язкових предметів, однак варто зауважити, що цей предмет обов'язковий для вступу в ВНЗ, для подальшого вивчення фізико-математичних наук. У 2009 р. кількість учасників ЗНО, які брали участь в оцінюванні своїх знань з фізики, становила 32 161 особи, а у 2013 р. – 61 560 осіб. За п'ять років кількість учасників, які перевіряли свої знання з фізики, збільшилася практично вдвічі. Проаналізуємо оцінки учасників тестування з фізики за ці п'ять років.

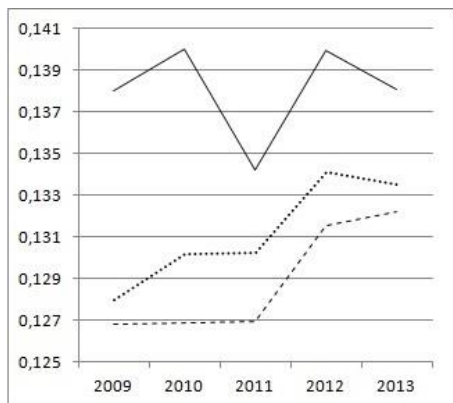
Як видно з табл. 4, до кластера високого рівня (середня оцінка 154 бали) увійшли Івано-Франківський регіон і м. Київ; Львівський регіон входив до цього кластера у 2011-2013 рр., а Рівненський – у 2009, 2011 і 2013 рр. Кластер середнього рівня має середній бал – 152, і протягом усього досліджуваного періоду до нього входила лише Хмельницька область. Як таких, що постійно перебувають у середньому рівні, варто виділити Дніпропетровський регіон (увійшов до кластера високого рівня у 2012 р.), Миколаївську, Одеську і Харківську області (перші дві входили до кластера низького рівня у 2009 р., а третя – у 2013 р.). Низький рівень з середнім балом 151 усі п'ять років у своїх об'єднаннях містить лише два регіони: Луганський і Полтавський, Кіровоградська область тільки один раз у 2011 р. входила до кластера середнього рівня.

Кожен кластер має власні середній бал і дисперсію. Дослідимо як змінювалися оцінки середнього рівня знань учасників зовнішнього незалежного тестування. Розглянемо коефіцієнти варіації для кожного кластера і коефіцієнт варіації як відношення квадратного кореня дисперсії до середнього балу України, які зображені на лінійному графіку (див. рис.).

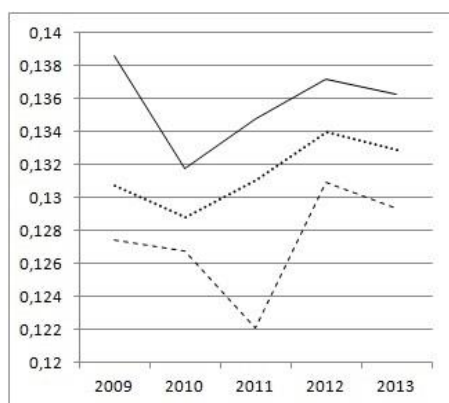
Спостерігаємо у порівняльних графіках математики та фізики (рис. 6, 2) практично паралельні лінії різних кластерів. Зокрема, у 2011 р. кластер низького рівня оцінок з математики мав найнижчий коефіцієнт варіації, і до нього увійшло 12 регіонів, що є рекордом за всі п'ять років; а от у 2012 р. спостерігалось різке



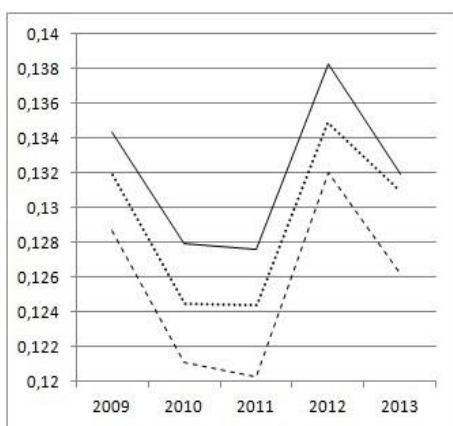
підвищення рівня оцінок, у кластер середнього рівня об'єднано 15 регіонів, що є найбільшою кількістю за всі п'ять років, і всього 7 регіонів у кластері низького рівня.



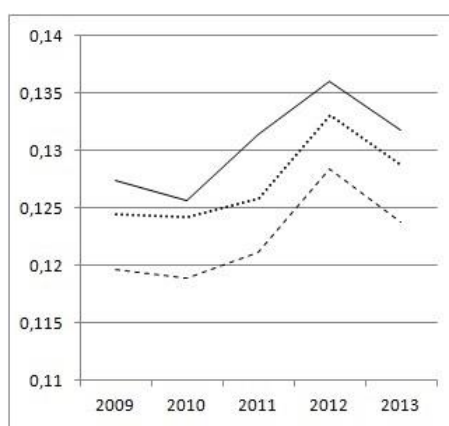
a



b



v



z

— Високий рівень    ..... Середній рівень    - - - - - Низький рівень

Порівняльний графік коефіцієнтів варіації кластерів: a – українська мова і література; b – історія України; v – математика; z – фізика

Зображені лінії графіків української мови і літератури (рис. a) та історії України (рис. б), навпаки, не є паралельними. Спостерігаємо різке зниження коефіцієнта варіації високого рівня у 2011 р. з української мови, і коефіцієнта низького рівня з історії України у тому ж році. Це означає, що середній бал і дисперсія цих кластерів значно знизилася порівняно з 2010 і 2012 роками, відповідно це також означає, що регіони, які входять до цих кластерів, мають показники різкого зниження середнього рівня оцінок, тоді як інші кластери доводять, навпаки,

підвищення середнього значення балу та дисперсії. У 2013 р. майже у всіх кластерів з чотирьох предметів відбувалось зниження рівня знань (виняток становить кластер низького рівня української мови та літератури).

### 3. ВИСНОВКИ

Такий розподіл дає змогу проаналізувати середній рівень знань України на підставі регіонів, що допоможе провести педагогічні заходи щодо підвищення ефективності технології здобуття знань з української мови, історії України, математики та фізики в середніх школах, наприклад, у регіонах, які найчастіше потрапляють до кластера низького рівня.

Результати проведеного дослідження свідчать про те, що лідерами серед регіонів України за результатами тестів є: з української мови і літератури – Закарпатський та Львівський регіони; з історії України – Закарпатський, Івано-Франківський і Львівський регіони; з математики – м. Київ, Івано-Франківська і Львівська області; з фізики – м. Київ та Івано-Франківська область. Аналіз кластера низького рівня допоміг виявити групу областей, які мають найслабші знання з п'яти предметів і потребують першочергового втручання Міністерства освіти і науки для поліпшення ситуації. До цього кластера входять: з української мови та літератури – Кіровоградська та Херсонська області; з історії України та хімії – Херсонська область; з математики – Луганський і Тернопільський регіони; з фізики – Луганський і Полтавський регіони.

Цей аналіз виявив не рейтинг конкретного регіону і його середніх загальноосвітніх навчальних закладів, а здатність учня самостійно опанувати навчальний матеріал поза межами свого закладу, що залежить не лише від організації роботи навчального закладу, а й від педагогічної майстерності вчителів і від активного застосування інформаційних технологій.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Українська Вікіпедія про Зовнішнє незалежне оцінювання. [Електронний ресурс]. – Режим доступу: [http://uk.wikipedia.org/wiki/Зовнішнє\\_незалежне\\_оцінювання](http://uk.wikipedia.org/wiki/Зовнішнє_незалежне_оцінювання)
2. Мандель И.Д. Кластерный анализ / И.Д. Мандель. – М.: Финансы и статистика, 1988. – 176 с.
3. StatSoft. Электронный учебник по статистике. [Електронний ресурс]. – Режим доступу: <http://www.statsoft.ru/home/textbook/default.htm>
4. Бахрушин В.С. Методи аналізу даних: навч. посібник для студентів / В.С. Бахрушин. – Запоріжжя: КПУ, 2011. – 268 с.
5. Офіційний сайт Українського центру оцінювання якості освіти [Електронний ресурс]. – Режим доступу: <http://testportal.gov.ua/>

*Стаття: надійшла до редколегії 20.05.2015  
доопрацьована 10.06.2015  
прийнята до друку 17.06.2015*

**CLUSTER ANALYSIS WHILE ASSESSING THE LEVEL OF KNOWLEDGE  
ACCORDING TO THE EXTERNAL INDEPENDENT TESTING  
IN THE REGIONS OF UKRAINE**

**N. Zadorozhnia, Y. Eleyko, Yu. Shcherbina**

*Ivan Franko National University of Lviv,  
Universytetska Str., 1, Lviv, 79000*

In this article, we study the possibility of using cluster analysis for the assessment of Ukrainian educational level. Clustering regions of Ukraine by the results of independent external testing of four subjects during 2009-2013 is considered. Analyze the group of homogeneous regions and through comparative graph of the coefficients of variation to compare the results of clustering.

*Key words:* cluster analysis, cluster, method of  $k$ -means, independent external testing (IET), the coefficient of variation.