

Ю.А. Мала, Т.В. Селівьорстова, А.І. Гуда

## ВИКОРИСТАННЯ ТЕХНОЛОГІЇ ORANGE ДЛЯ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ В ОСВІТНІЙ ГАЛУЗІ

*Анотація.* Розвиток систем електронного навчання та впровадження цифрових технологій у традиційний освітній процес призводить до появи великих інформаційних об'ємів, дослідженням яких займається інтелектуальний аналіз освітніх даних (ІАОД) наукова дисципліна, пов'язана із застосуванням методів інтелектуального аналізу даних до інформації, яку продукують освітні заклади. Актуальність теми зумовлена необхідністю підвищення якості освітнього процесу за рахунок виявлення прихованих закономірностей в масиві освітніх даних. У роботі розкрито визначення інтелектуального аналізу освітніх даних, показана актуальність і необхідність подальшого розвитку даного напрямку в сучасних умовах. В якості прикладу побудована схема аналізу освітніх даних з метою демонстрації ефективного використання інструменту візуального програмування Orange для дослідження освітніх даних з застосуванням методів Data Mining, що дозволяє фахівцям в освітній галузі проводити якісний аналіз з подальшим використанням отриманих результатів при розробці стратегій для забезпечення ефективного процесу навчання та розвитку освіти.

*Ключові слова:* інтелектуальний аналіз даних, освітні дані, інструмент Orange, віджети, візуалізація, методи, кластеризація, Data Mining.

**Актуальність теми.** Зростання обсягів даних, що пов'язано з інтенсивним розвитком інформаційних технологій, створює потребу у високоефективних методах їх обробки та аналізу з метою ефективного управління різними видами організацій також з метою стратегічного планування та конкурентоспроможності їх на ринку. Такими методами обробки та аналізу даних є методи інтелектуального аналізу даних. Під інтелектуальним аналізом даних (ІАД, Data Mining), за визначенням засновника цього напрямку в сфері аналізу даних Г. Пятецький-Шапіро [1], розуміють сукупність методів виявлення в даних раніш невідомих, нетривіальних, практично корисних і доступних інтерпретації знань, що необхідні для прийняття рішень в різних галузях людської діяльності. Спектр використання методів інтелектуального аналізу даних дуже широкий, основною умовою використання є наявність великої

кількості даних. Одним із перспективних напрямків застосування методів інтелектуального аналізу даних є освітня галузь. Погіршення якості навчання дітей в Україні, що показали результати тестування PISA, пов'язано з різними факторами, а особливо протягом останніх років пов'язано зі складними умовами провадження освітньої діяльності через виникнення пандемій і військового стану в країні. Таким чином залучення нових ефективних методів інтелектуального аналізу освітніх даних для виявлення певних закономірностей і факторів, що призводять до неякісних результатів в навчання і подальша розробка планів подолання негативних факторів є актуальною задачею сьогодення.

**Аналіз останніх досліджень і публікацій.** Питанням застосування методів інтелектуального аналізу даних приділили увагу багато науковців, як і зарубіжних вчених так і вітчизняних, зокрема: Колодчак О.М., Черняк О. І., Ілляшенко К., Г. Пятецький-Шапіро [1], Усама Файяд [1], А. Барсегіяна, Ч. Танга і ін. Дослідженню у сфері інтелектуального аналізу освітніх даних присвячені роботи [2 – 8, 11]. Огляд та приклади застосування технологій візуального програмування розглянуто в працях [2, 3, 9].

**Основною метою** роботи є дослідження питання використання інструменту візуального програмування для задач інтелектуального аналізу даних та його застосування для аналізу освітніх даних, побудова схеми аналізу освітніх даних з використанням технологій Orange для демонстрації ефективності і зручності при виявленні прихованих закономірностей з метою вдосконалення освітніх процесів.

**Виклад основного матеріалу.** Інтелектуальний аналіз даних – це об'єднання математичних алгоритмів та новітніх досягнень в сфері інформаційних технологій, яке надає можливість видобутку корисної інформації з наявних даних і подальшого прийняття обґрунтованих і ефективних рішень.

Основні задачі, що вирішуються з використанням методів ІАД: класифікація – полягає в виявленні ознак, які характеризують групу до якої належить той чи інший об'єкт; кластеризація – полягає в пошуку незалежних груп (кластерів) та їх характеристик у всій множині аналізованих даних; асоціація – полягає в знаходженні частих залежностей (або асоціацій) між об'єктами чи подіями. Знайдені залежності представляються у вигляді правил і можуть бути використані як для кращого розуміння природи аналізованих даних, так і для передбачення появи подій; аналіз відхилень – полягає в виявленні найбільш

нетипових шаблонів; прогнозування – полягає в тому, що на основі особливостей історичних даних оцінюється пропущені або ж майбутні значення цільових чисельних показників.

На даний час існує достатньо велика кількість інструментів, що допомагають вирішувати задачі ІАД:

- Мови програмування: Python, R, C++.
- Програмне забезпечення: KNIME, Orange, WEKA і інші.

Вибір програмного забезпечення залежить від типу даних, кількості даних, мети використання, можливостей замовників. Вибір інструменту ІАД обмежується вартістю та функціоналом програмного продукту. Щодо використання програмного продукту в навчальному процесі головними критеріями є: відкритий вихідний код, простота в використанні, наочність. Серед наявного програмного забезпечення в даній роботі буде розглядатися платформа Orange [10].

Orange — це фреймворк для візуалізації та аналітики даних, який спеціалізується на великій кількості методів машинного навчання та аналітичних інструментах. Його основною метою є зробити складні аналітичні методи доступними для широкого кола користувачів, включаючи тих, хто не має глибоких знань у програмуванні чи статистиці. Orange використовує мову програмування Python як один з основних елементів своєї архітектури. Python можна використовувати як складову частину своїх аналітичних проєктів в Orange, що дозволяє комбінувати візуальне програмування з написанням коду на Python для отримання рішень більш складних, персоналізованих аналітичних задач. Також Orange може бути використаний як бібліотека в середовищі Python і імпортувати функціональність Orange у код Python.

Orange складається з інтерфейсу Canvas, на який користувач поміщає віджети і створює робочий процес аналізу даних. Віджети – це будівельні блоки робочих процесів аналізу даних. Віджет виконує дії залежно від того, що ви виберете в списку віджетів у лівій частині екрана.

Функціями віджетів можуть бути наступні категорії [9, 10]:

- Data – віджети для введення / виведення даних, фільтрації, різні дії з вибірками, а також велика кількість навчальних наборів даних;
- Visualize – віджети для загальної і багатовимірної візуалізації;
- Classify – набір алгоритмів машинного навчання для класифікації;
- Regression - набір алгоритмів регресійного аналізу;

- Evaluate – кросвалідація, процедури на основі вибірки, аналіз методів прогнозування;
- Unsupervised – алгоритми кластеризації і проєкції даних.

Додаткові віджети для аналізу даних при необхідності, такі як Associate, Image analytics, Network, Text mining, Time series і інші є можливість завантажити [9, 10]. Користувач може інтерактивно досліджувати візуалізації або передавати обрану підмножину в інші віджети. В Orange процес застосування ІАД реалізується за допомогою візуального програмування, а саме віджетів. Інструмент Orange вміє розпізнавати файли як у власному так і в інших форматах даних (csv, exl). Вище перераховані переваги свідчать проте, що платформа Orange є ефективною в питаннях аналізу даних з залученням методів ІАД.

Перспективним напрямком застосування методів ІАД є освітня діяльність у зв'язку з тим, що відбувається інтенсивне накопичення великих даних в освітніх установах.

Великі дані в освіті можуть допомогти університетам, коледжам і школам трансформувати свої бізнес-моделі, академічні результати успішності студентів і ефективність викладачів. Крім того, знання, отримані з великих даних, можуть допомогти навчальним закладам покращити свої технологічні системи.

Великі дані в освіті охоплюють комплексний збір, обробку та інтерпретацію великих і різноманітних джерел даних у сфері освіти, що передбачає використання масивних наборів даних, створених навчальними закладами, студентами, викладачами та пов'язаними системами. Такі набори даних часто охоплюють широкий спектр інформації, включаючи демографічні дані студентів, записи про академічну успішність, дані відвідуваності, онлайн-взаємодії, результати оцінювання тощо. На основі цих даних можна знаходити закономірності і взаємозв'язки з подальшим прогнозуванням, що будуть залучені при побудові стратегії розвитку окремої освітньої установи, або розвитку освітніх процесів на рівні регіональному чи в рамках цілої держави.

Інтелектуальний аналіз освітніх даних (ІАОД, Educational Data Mining (EDM)) — це нова міждисциплінарна область досліджень, яка займається розробкою методів дослідження даних в освітній сфері, що допомагають виявляти раніше невідомі, практично корисні знання про освітній процес і його учасників з метою підтримки прийняття рішень [болюбаш]. ІАОД використовує статистичні методи, методи машинного навчання та методи Data Mining для

аналізу освітніх даних з метою вивчення освітніх процесів, прийняття рішень і рекомендацій в освітній галузі.

Для дослідження у сфері інтелектуального аналізу освітніх даних використовують традиційні методи Data Mining: класифікацію, кластеризацію, виявлення взаємозв'язків, моделювання, пошук асоціативних правил та послідовних шаблонів, Text Mining (інтелектуальний аналіз текстів), Visual Mining. Однак реалізація цих методів в освітній сфері для аналізу цифрових даних стосовно процесів навчання має свої особливості, обумовлені цілями аналізу та специфікою даних, які аналізуються [11].

Для наочного прикладу застосування інструменту Orange було використано набір освітніх даних «World Educational Data» з ресурсу <https://www.kaggle.com/>. Задачею даного прикладу є демонстрація застосування інструменту Orange для аналізу освітніх даних з залученням методів Data Mining. Дані завантажуються в Orange за допомогою віджету File у форматі .csv, і за допомогою віджета Data Table можна переглянути в вигляді таблиці (рис.1).

Unemployment_Rate	Countries and areas	Latitude	Longitude	ISCED_Primary_Agt_Male	ISCED_Primary_Agt_Female	ISCED_Secondary_Agt_Male	ISCED_Secondary_Agt_Female	ISCED_Tertiary_Agt_Male	ISCED_Tertiary_Agt_Female	ISCED_Quaternary_Agt_Male	ISCED_Quaternary_Agt_Female	ISCED_Quinary_Agt_Male	ISCED_Quinary_Agt_Female	ISCED_Sixary_Agt_Male	ISCED_Sixary_Agt_Female	ISCED_Senary_Agt_Male	ISCED_Senary_Agt_Female	ISCED_Octary_Agt_Male	ISCED_Octary_Agt_Female	ISCED_Nonary_Agt_Male	ISCED_Nonary_Agt_Female	ISCED_Tenary_Agt_Male	ISCED_Tenary_Agt_Female	
28.18	South Africa	30.5595	22.9375	28	27	12	10	16	13	22	20													
23.41	Lesotho	29.61	28.2336	58	57	3	2	21	13	38	29													
20.71	Saint Lucia	13.9094	60.9789	5	0	3	0	9	10	20	21													
20.27	Namibia	22.9576	18.4904	33	30	0	0	0	0	0	0													
20.08	Gabon	0.803689	11.6094	0	0	0	0	0	0	0	0													
18.88	Saint Vincent a...	12.9843	61.2872	9	3	0	0	1	2	16	14													
18.56	Libya	26.3351	17.2283	0	0	0	0	0	0	0	0													
18.42	Bosnia and Her...	43.9159	17.6791	72	73	0	0	0	0	22	19													
18.19	Botswana	22.3285	24.6849	79	78	12	10	0	0	0	0													
17.84	Greece	39.0742	21.8243	6	5	2	1	3	5	4	6													
16.99	Armenia	40.0691	45.0382	52	50	9	9	11	9	16	4													
16.53	Sudan	12.8628	30.2176	60	60	31	35	33	35	53	50													
16.02	Tunisia	33.8869	9.5375	0	0	0	0	0	0	0	0													
14.88	Montenegro	42.7087	19.3744	23	24	3	3	8	7	12	10													
14.72	Jordan	31.3028	36.7868	55	55	18	19	29	30	48	43													
14.70	United States	37.0902	95.7129	12	8	1	1	3	1	4	3													
14.48	Georgia	42.3154	43.3569	0	0	2	2	1	2	7	4													
13.96	Spain	40.4637	3.74922	5	5	3	3	1	1	2	1													
13.78	Haiti	18.9712	72.2852	0	0	0	0	0	0	0	0													
13.48	Turkey	38.9637	35.2433	23	25	5	5	6	7	17	17													
13.27	SenSenSenSenSen...	0.18636	6.61308	49	46	6	6	12	7	19	16													
12.91	Yemen	15.5527	48.5164	96	96	10	21	23	34	46	68													
12.82	Iraq	33.2232	43.6793	0	0	0	0	0	0	0	0													
12.62	Serbia	44.0165	21.0059	9	10	2	2	2	2	14	11													
12.31	Albania	41.1533	20.1683	4	2	6	3	6	1	21	15													
12.25	Cape Verde	16.5388	23.0418	19	19	6	7	14	13	29	24													
12.22	South Sudan	6.87699	31.307	78	81	58	67	49	63	57	72													
12.08	Brazil	14.235	51.9253	0	1	1	1	2	3	16	14													
11.88	Costa Rica	9.74892	83.7534	4	4	3	3	3	1	7	5													
11.88	Guyana	4.86042	58.9302	0	0	0	0	0	0	0	0													
11.79	Algeria	28.0339	1.65963	0	0	0	0	0	0	0	0													
11.42	Zambia	13.1399	27.8493	0	0	17	13	0	0	0	0													
11.38	Iran	32.4279	53.688	50	49	0	0	3	6	25	27													
11.36	Somalia	5.15215	46.1996	0	0	0	0	0	0	0	0													
11.12	Afghanistan	33.9391	67.71	0	0	0	0	0	0	44	69													

Рисунок 1 - Приклад набору даних

В віджеті File необхідно встановити тип даних в стовпчику Type: categorical (категоріальний), numeric (числовий), text (текстовий), datetime (час і дата), також встановити роль даних в стовпчику Role: target (цільовий атрибут), features (числове значення), meta (текстові), skip (ігнорування атрибуту) (рис.2).

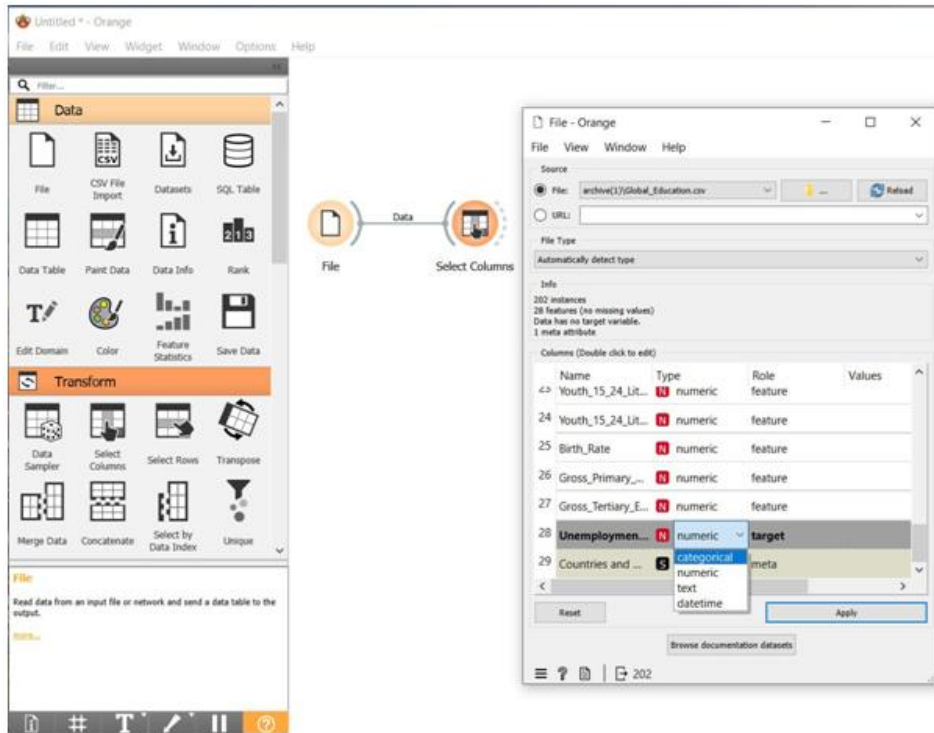


Рисунок 2 - Вибір типу даних і роль атрибуту

В випадку, коли при дослідженні не всі потрібні атрибути в таблиці, або необхідно видалити нульові значення в рядочках, то необхідно застосувати віджет Select Columns, Select Rows або Rank і з'єднавши з віджетом Data Table можна переглянути відредаговану таблицю (рис. 3).

The image shows the 'Data Table' widget configuration window. The 'Info' section indicates 79 instances (no missing data), 1 feature, and 1 meta attribute. The 'Variables' section has checkboxes for 'Show variable labels (if present)', 'Visualize numeric values', and 'Color by instance classes'. The 'Selection' section has a checkbox for 'Select full rows'. The main table displays the following data:

	Unemployment_Rate	Countries and areas	1_15_24_Literacy_Rate
13	3.68	Central African ...	48
65	12.24	South Sudan	48
49	0.47	Niger	51
40	7.22	Mali	58
11	6.26	Burkina Faso	62
19	3.32	Ivory Coast	64
8	2.23	Benin	70
26	4.30	Guinea	70
63	4.43	Sierra Leone	71
69	16.53	Sudan	73
1	11.12	Afghanistan	74
16	4.34	Comoros	78
38	1.76	Madagascar	82
50	8.10	Nigeria	82
72	4.55	East Timor	82
58	1.03	Rwanda	84

Рисунок 3 - Відредагована таблиця даних з видаленими нульовими значеннями та видаленими зайвими стовпчиками

### «Системні технології» 3 (152) 2024 «System technologies»

Для проведення первинного розвідувального статистичного аналізу була складена схема аналізу з використанням віджетів Scatter Plot та Distributions з розділу Visualize з метою виявлення наявних закономірностей, припущень відносно даних (рис. 4.).

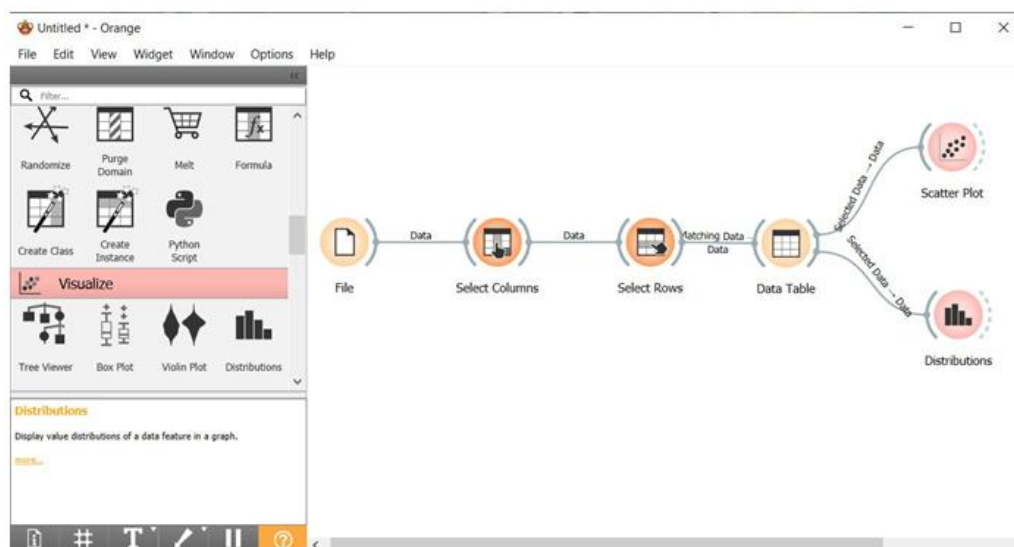


Рисунок 4 - Схема аналізу даних в Orange

Результати представлені розвідувального статистичного аналізу на рис. 5, рис. 6.

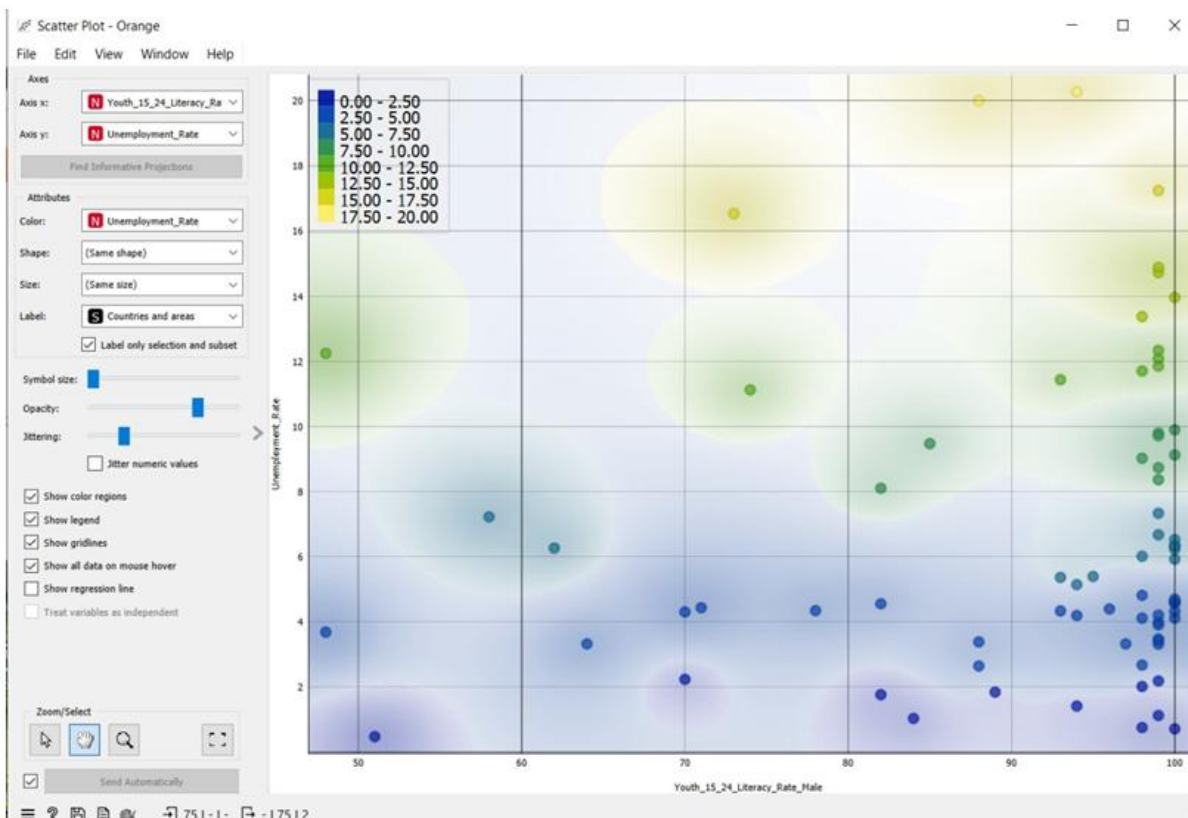


Рисунок 5 - Результат застосування Scatter Plot

Віджет Scatter Plot дає можливість візуалізувати дані у вигляді точкових графіків з попереднім аналізом та інтелектуальними поліпшеннями візуалізації даних. Основна мета цього віджету створити візуальну форму для кращого та ефективнішого розуміння закономірностей, прихованих у даних.

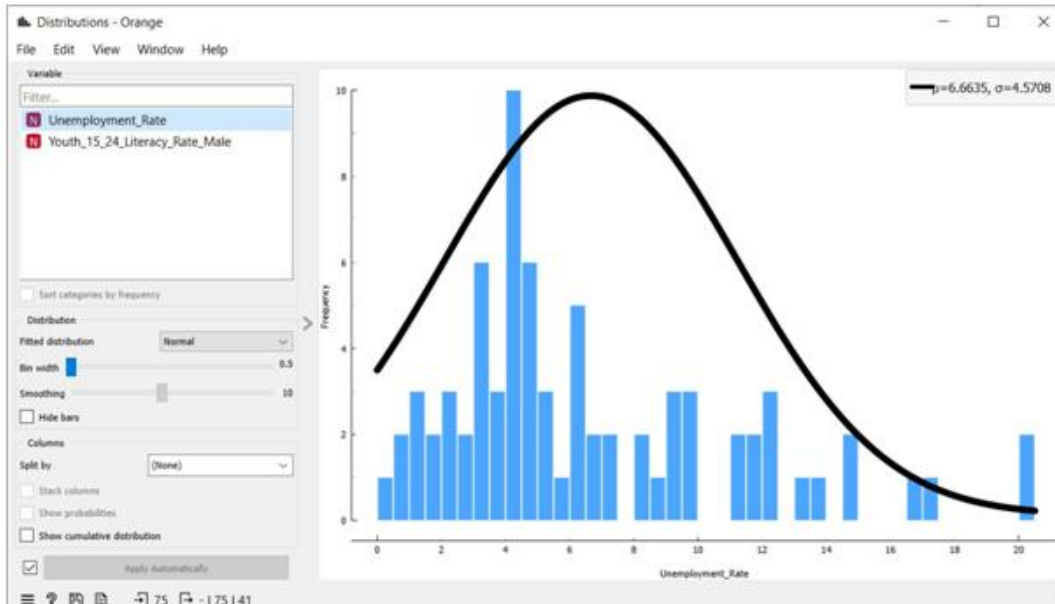


Рисунок 6 - Результат застосування віджету Distributions

Як видно з рис. 5 по осі x показано атрибут Youth\_15\_24\_Literacy\_Rate\_Male, що показує рівень грамотності серед молодих чоловіків віком 15-24 роки, по осі y - Unemployment\_Rate, що показує рівень безробіття у відповідних країнах, кольором виділено рівень безробіття у відповідних країнах. Рис. 5 демонструє цілком очевидний результат – рівень безробіття у відповідних країнах тим нижче, чим нижче рівень грамотності серед молодих чоловіків віком 15-24 роки.

Віджет Distributions дозволяє будувати діаграми, що демонструють динаміку змін в даному прикладі рівень безробіття у відповідних країнах.

На рис. 7 представлена як додаткова гілка в попередній схемі аналізу даних в Orange представляє собою процес застосування методу кластеризації k-Means.



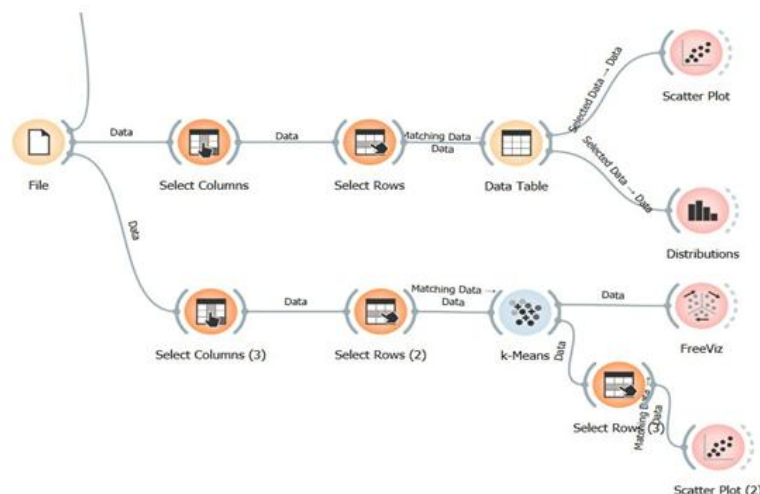


Рисунок 7 - Схема застосування методу кластеризації k-Means

Віджет k-Means з розділу Unsupervised застосовує алгоритм кластеризації k-Means до даних і виводить новий набір даних, до якого мітка кластера додається як метаатрибут. При застосуванні віджету k-Means необхідно застосувати один із віджетів з розділу Visualize для візуалізації результатів кластеризації. В нашому випадку для кластеризації були обрані такі атрибути: Unemployment\_Rate - рівень безробіття у відповідних країнах, Primary\_End\_Proficiency\_Math - знання математики наприкінці початкової освіти, Gross\_Tertiary\_Education\_Enrollment - загальна кількість зареєстрованих у вищій освіті. Для візуалізації результатів методу кластеризації були обрані два віджета Scatter Plot і FreeViz. Результати представлені на рис. 8. та рис. 9.

На рис. 8 представлено 3 кластера, кількість яких задано заздалегідь в віджеті k-Means, що виділені різним кольором.

Розмір точок показує рівень безробіття в країнах, і можливо зробити висновок, що низький рівень безробіття пов'язаний з низьким знань математики наприкінці початкової середньої освіти, що підтверджується і на рис. 9 з застосуванням віджету Scatter Plot.

Треба відмітити, що серед віджетів є віджет Python Script в закладці Transform, що дозволяє розширити функціональні можливості за допомогою сценаріїв Python у вхідних даних, якщо відповідна функціональність не реалізована в існуючому віджеті. Така можливість надає інструменту візуального програмування Orange гнучкості в застосуванні і вирішенні різноманітних задач інтелектуального аналізу даних. Взагалі технології Orange є вивченим інструментом для проведення аналізу даних в освітній сфері з використанням методів Data Mining.

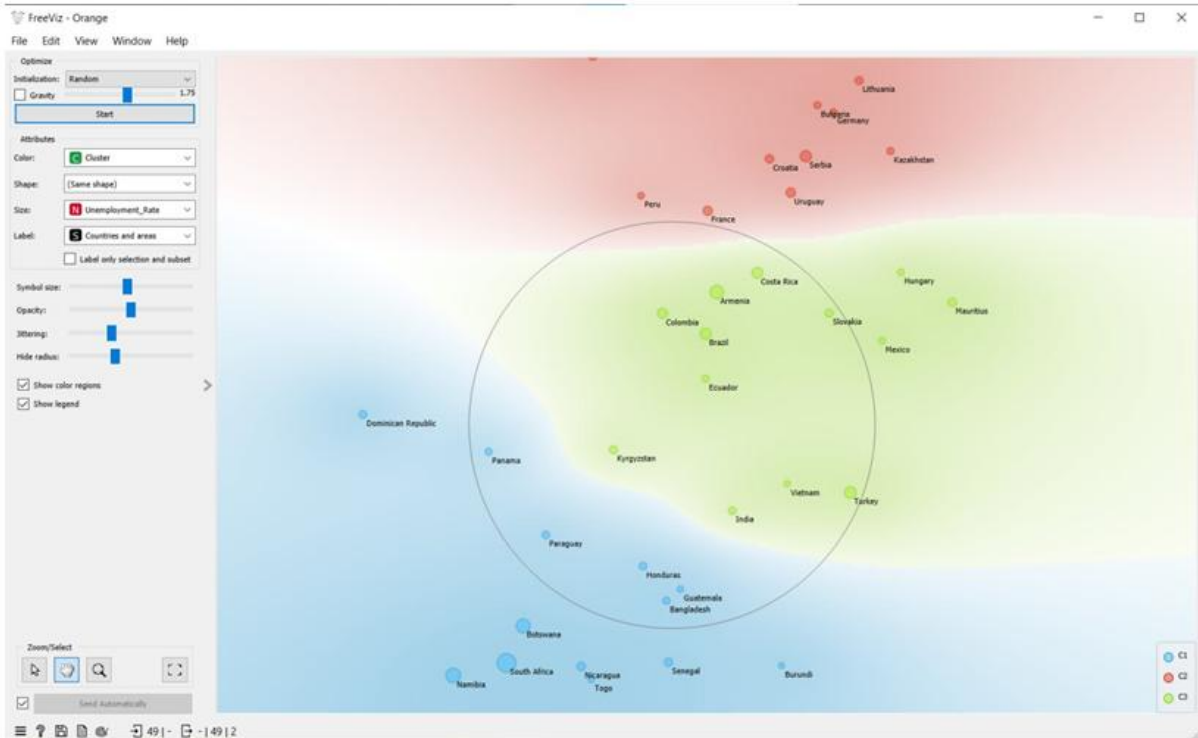


Рисунок 8 - Результат застосування віджета k-Means і FreeViz

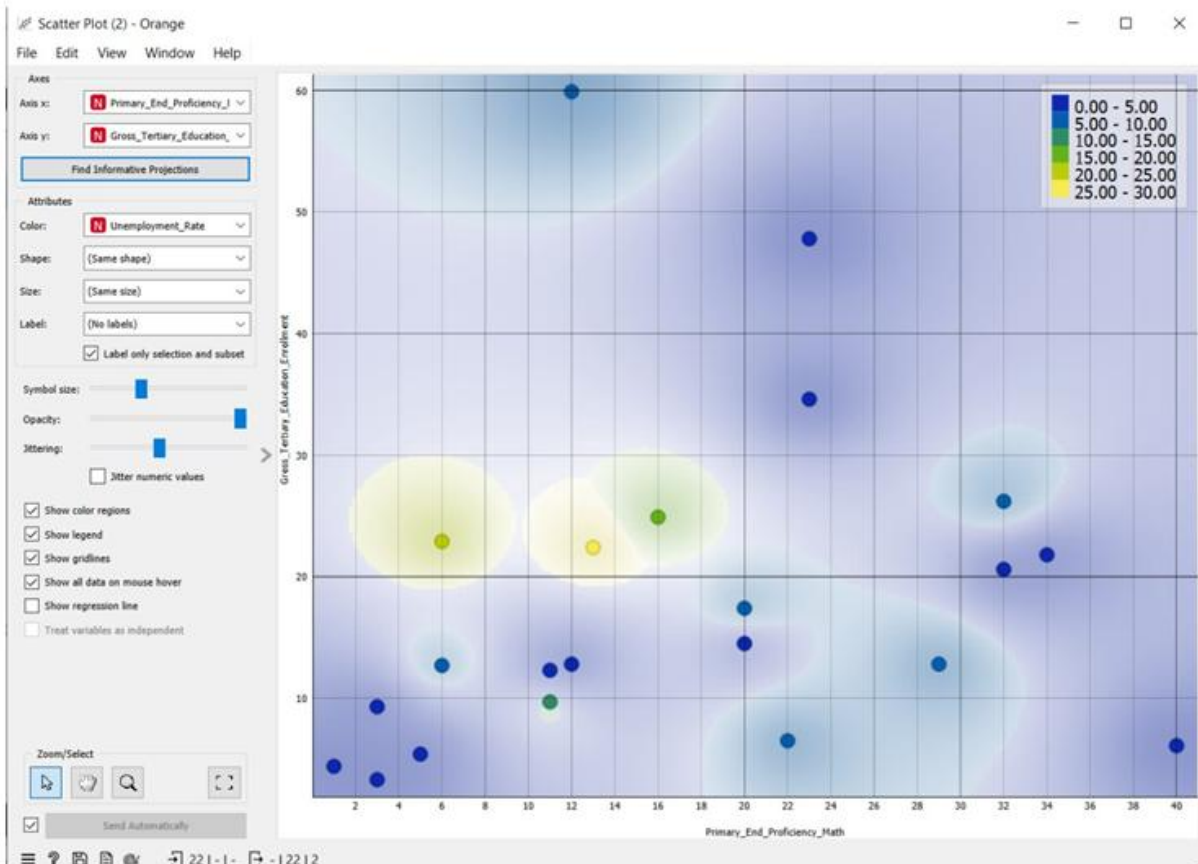


Рисунок 9 - Результат застосування віджета k-Means і Scatter Plot

**Висновки.** В роботі на основі проведеного дослідження та побудови наочного прикладу схеми аналізу освітніх даних було показано, що використання методів Data Mining і застосування в якості інструменту технології візуального програмування Orange дозволяє фахівцям в освітній галузі проводити якісний аналіз з подальшим використанням висновків при розробці стратегій для забезпечення ефективного процесу навчання.

#### **ЛІТЕРАТУРА**

1. Piatetsky-Shapiro G., Frawley W. Knowledge Discovery in Databases. - Publisher: AAAI Press, 1991. 540 pp.
2. Abdelmagid A. S., Qahmash A. I. M.. Utilizing the Educational Data Mining Techniques "Orange Technology" for Detecting Patterns and Predicting Academic Performance of University Students // Information Sciences Letters. An International Journal, 2023. Vol. 12. No. 3. P. 1415-1431.
3. Hernández-Leal E., Darío Duque-Méndez N. C. Cechinel Unveiling educational patterns at a regional level in Colombia: data from elementary and public high school institutions // Heliyon, 2021. Vol. 7. P. 3-17.
4. M. Goyal, R. Vohra. Applications of Data Mining in Higher Education // International Journal of Computer Science Issues, 2012. Vol. 9. Issue 2. No 1. P. 113 – 120.
5. Вергун В. Р. Характеристика методів розв'язання задачі класифікації в інтелектуальному аналізі даних навчальних програм // Науковий вісник НЛТУ України, 2019. Т. 29. № 6. С. 136-139.
6. Romero C., Ventura S. Educational Data Mining: A Review of the State of the Art // Transactions on Systems Man and Cybernetics Part C (Applications and Reviews), 2010. 40(6). P. 601 – 618.
7. Romero C., Ventura S. Data mining in education // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2013. Vol. 3(1). P. 12-27.
8. Chao Zhang, Linling He. Data Mining Technology in Teaching Evaluation of Colleges and Universities // SHS Web of Conferences, 2024. Vol. 187. 4 pp.
9. Пронін С. В., Сотников А. Д. Використання платформи orange для аналізу даних // Вісник ХНАДУ, 2022. Вип. 99. С. 131 – 137.
10. Data Mining Fruitful and Fun Open source machine learning and data visualization. <https://docs.biolab.si/orange/2/widgets/rst/index.html#data>
11. Болюбаш Н. М. Задачі та методи інтелектуального аналізу освітніх даних // Наукові праці. Педагогіка, 2019. Випуск 311. Том 323. С. 26 - 29.

## REFERENCES

1. Piatetsky-Shapiro G., Frawley W. Knowledge Discovery in Databases. - Publisher: AAAI Press, 1991. 540 pp.
2. Abdelmagid A.S., Qahmash A.I. M.. Utilizing the Educational Data Mining Techniques "Orange Technology" for Detecting Patterns and Predicting Academic Performance of University Students // Information Sciences Letters. An International Journal, 2023. Vol. 12. No. 3. P. 1415-1431.
3. Hernández-Leal E., Darío Duque-Méndez N. C. Cechinel Unveiling educational patterns at a regional level in Colombia: data from elementary and public high school institutions // Heliyon, 2021. Vol. 7. P. 3-17.
4. M. Goyal, R. Vohra. Applications of Data Mining in Higher Education // International Journal of Computer Science Issues, 2012. Vol. 9. Issue 2. No 1, P. 113 – 120.
5. Verhun V.R. Kharakterystyka metodiv rozviazannia zadachi klasyfikatsii v intelektualnomu analizi danykh navchalnykh proham // Naukovyi visnyk NLTU Ukrainy, 2019. Tom. 29. № 6. C. 136-139.
6. Romero C., Ventura S. Educational Data Mining: A Review of the State of the Art // Transactions on Systems Man and Cybernetics Part C (Applications and Reviews), 2010.40(6). P. 601 – 618.
7. Romero C., Ventura S. Data mining in education // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2013. Vol. 3(1). P. 12-27.
8. Chao Zhang, Linling He. Data Mining Technology in Teaching Evaluation of Colleges and Universities // SHS Web of Conferences, 2024. Vol. 187. 4 pp.
9. Pronin S. V., Sotnykov A. D. Vykorystannia platformy orange dlia analizu danykh // Visnyk KhNADU, 2022. Vyp. 99. C. 131 – 137.
10. Data Mining Fruitful and Fun Open source machine learning and data visualization. <https://docs.biolab.si/orange/2/widgets/rst/index.html#data>
11. Boliubash N. M. Zadachi ta metody intelektualnoho analizu osvitynykh danykh. Naukovi pratsi. Pedahohika, 2019. Vypusk 311. Tom 323. C. 26 - 29 .

Received 18.04.2024.

Accepted 19.04.2024.

### ***Using orange technology for data mining in the education sector***

*The growth of data volumes due to the intensive development of information technology creates a need for highly efficient methods of data processing and analysis for the effective management of various types of organizations, as well as for strategic planning and market competitiveness. Such methods of data processing and analysis are methods of data mining. One of the most promising areas of application of data mining methods is in the education sector. The development of e-learning systems and the introduction of*

*digital technologies into the traditional educational process leads to the emergence of large amounts of information, which are studied by the intellectual analysis of educational data (IAED), a scientific discipline related to the application of data mining methods to information produced by educational institutions. The relevance of the topic is due to the need to improve the quality of the educational process by identifying hidden patterns in the array of educational data, and the need for further development of this area in modern conditions is shown.*

*The paper shows that among the available software, Orange technologies are effective and easy to use, the main purpose of which is to make complex analytical methods accessible to a wide range of users, including those who do not have deep knowledge of programming or statistics.*

*As an example, a scheme for analyzing educational data is built to demonstrate the effective use of the Orange visual programming tool for researching educational data using data mining methods, which allows specialists in the field of education to conduct a qualitative analysis and then use the results obtained in developing strategies to ensure an effective learning process and educational development.*

**Мала Юлія Анатоліївна** - кандидат технічних наук, доцент, доцент кафедри комп'ютерних наук та інженерії програмного забезпечення, факультет інноваційних технологій, Університет митної справи та фінансів.

**Селівьорстова Тетяна Віталіївна** – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій і систем, факультет прикладних комп'ютерних технологій, Український державний університет науки і технологій.

**Гуда Антон Ігорович** – доктор технічних наук, доцент, професор кафедри Інформаційних технологій і систем, Український державний університет науки і технологій.

**Mala Yuliia** - PhD, Associate Professor, Associate Professor at the Department of Computer Science and Software Engineering, the University of Customs and Finance.

**Selivorstova Tatyana** – candidate of technical sciences, associate professor, associate professor department of Information Technologies and Systems, Faculty of Applied Computer Technologies at Ukrainian State University of Science and Technologies.

**Guda Anton** – Doctor of Technical Sciences, Associate Professor, Professor of the Department of Information Technologies and Systems, Ukrainian State University of Science and Technology.