

УЗГОДЖЕННЯ ОСВІТНІХ ПРОГРАМ З ВИМОГАМИ РИНКУ ПРАЦІ ДЛЯ ФАХІВЦІВ У СФЕРІ ДАНИХ: КЛАСТЕРНИЙ ПІДХІД

У цій статті розглядається проблема узгодження освітніх програм закладів вищої освіти зі швидко змінюваними вимогами ринку праці в галузі науки про дані. Використовуючи відкритий набір даних щодо вакансій в ІТ-сфері, ми застосували кластеризацію методом К-середніх для визначення чотирьох окремих категорій ролей фахівців з даних: аналітик і інженер даних, інженер платформи даних, спеціаліст із науки та інженерії даних, інженер хмарних даних. Кожен кластер представляє унікальну комбінацію затребуваних навичок, що дає уявлення про поточні вимоги галузі. Нами пропонується новий підхід для зіставлення цих кластерів навичок з освітніми програмами, використовуючи частку кредитів, виділених на відповідні компетентності. Цей метод демонструється на прикладі магістерської програми з інформаційних систем і технологій ХДУ. Для перевірки узгодженості вимог до навичок у кластері спеціаліста із науки та інженерії даних і структури компетенцій академічної програми застосовується χ^2 -тест. Наші результати свідчать про те, що цей підхід може ефективно сприяти розробці освітніх програм та допомагати студентам обирати відповідні програми, які узгоджуються з їхніми кар'єрними цілями. Це дослідження сприяє подоланню розриву між академічною сферою та індустрією, потенційно підвищуючи працевлаштованість випускників у галузі науки про дані, що стрімко розвивається.

Ключові слова: *Навички фахівців з даних, аналіз ринку праці, машинне навчання без учителя, кластеризація, узгодження освітніх програм, спеціальні компетентності*

This paper consider the challenge of aligning higher education programs with rapidly evolving labour market demands in the field of data science. Using an open dataset of IT job postings, we applied K-means clustering to identify 4 distinct categories of data specialist roles: Data Analyst & Engineer, Data Platform Engineer, Data Science & Engineering Specialist, and Cloud Data Engineer. Each cluster represents a unique combination of in-demand skills, providing insights into current industry requirements. We propose a novel approach to match these skill clusters with academic programs, using the proportion of credits allocated to relevant competencies. This method is demonstrated using a case study of a Master's program in Information Systems and Technologies at KSU. The χ^2 test is employed to validate the alignment between the skill requirements of the Data Science & Engineering Specialist cluster and the academic program's competency structure. Our findings suggest that this approach can effectively guide academic

program development and help students select appropriate academic programs aligned with their career goals. This research contributes to bridging the gap between academia and industry, potentially improving graduate employability in the rapidly evolving field of data science.

Key words: *data specialist skills, labour market analysis, unsupervised learning, clustering, academic program alignment, specific subject competences.*

Динамічний характер попиту на ринку праці, особливо в ІТ галузях, вимагає безперервного розвитку навичок, щоб здобувачі могли працевлаштовуватися [1, 2]. Відкриті освітні ресурси стають цінним джерелом навчальних матеріалів, які можна використовувати для персоналізованого навчання [3].

Штучний інтелект і методи інтелектуального аналізу даних можна використовувати для аналізу вакансій і визначення вимог до навичок для конкретних професій [4-6]. Системи рекомендацій на основі ШІ можуть підбирати здобувачам відповідний контент для опанування на основі вилучених вимог до професійних навичок та індивідуальних цілей навчання.

Збільшення обсягу даних, необхідність їх аналізу та інтерпретації збільшило попит на фахівців з аналізу даних, бізнес аналітики, машинного навчання [7-11].

Метою статті є розробка підходу до узгодження спеціальних компетентностей фахівців із даних для освітніх програм з навичками, затребуваними на ринку праці, для обґрунтування вибору такої освітньої траєкторії для здобувача, яка б дозволила підготувати його або її під час навчання до отримання бажаної посади фахівця з даних на ринку праці.

Згідно з методом ліктя для визначення оптимальної кількості кластерів на основі відкритих даних, найбільш значне зниження коефіцієнта інерції, який показує різницю між кластерами, відбувається при кількості кластерів, що дорівнює 4 (Рис. 1).

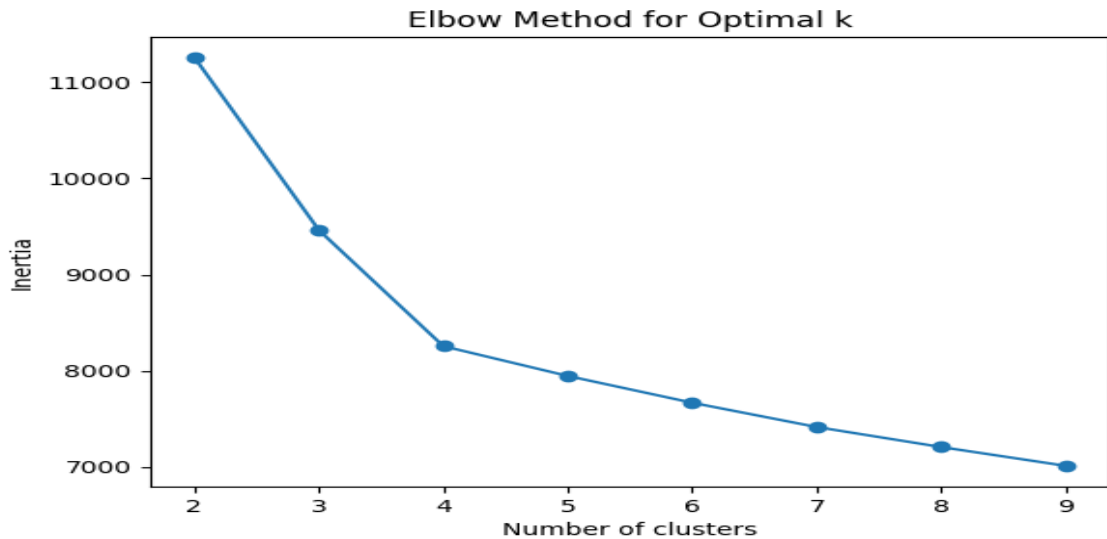


Рис. 1. Визначення оптимальної кількості кластерів за допомогою методу ліктя

Візуалізація кластерів даних за навичками проведена за допомогою методу стохастичного вкладення сусідів t-розподілу (метод t-SNE) як нелінійного методу зменшення розмірності (Рис. 2).

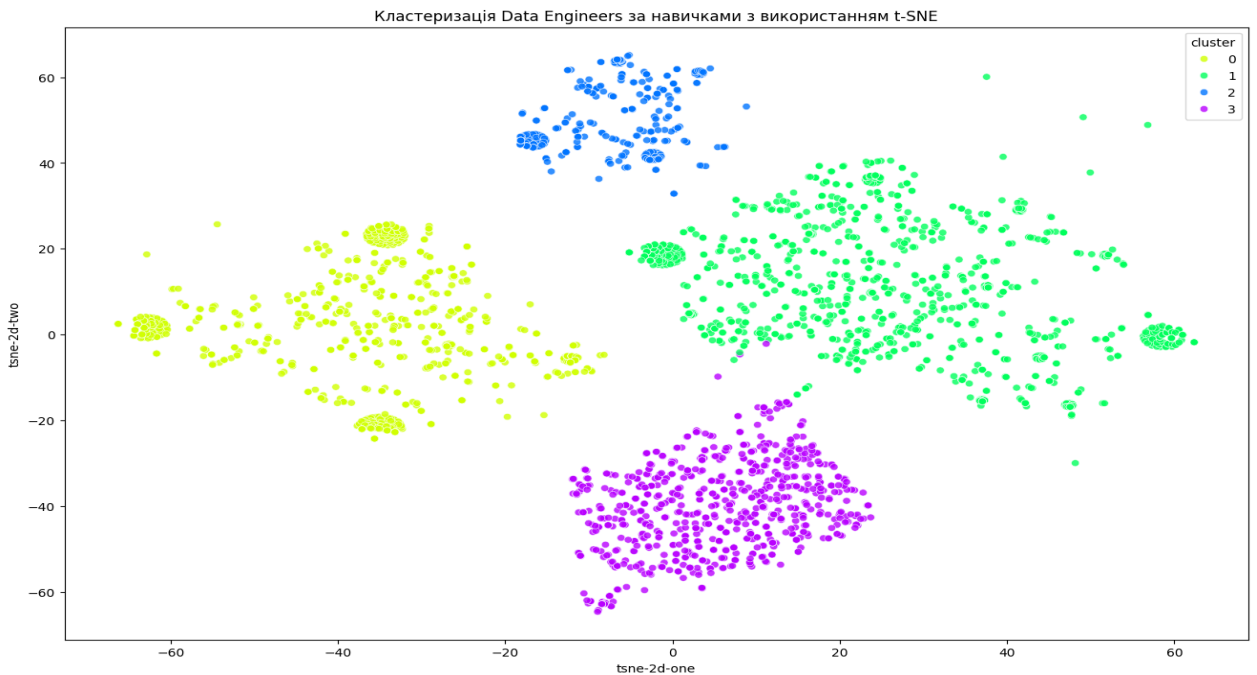


Рис. 2. Візуалізація кластерів позицій Data Specialist з використанням t-SNE

Розподіл топ навичок за кластерами (рис. 3) за допомогою методу кластеризації K-середніх дозволив виділити ключові категорії вакансій. Для кожної з них характерні свої вимоги до знань і навичок:

- Аналітики даних мають володіти навичками роботи з Python, SQL та іншими інструментами для аналізу даних.

- Інженери платформи даних повинні вміти працювати з хмарними платформами (AWS, Azure) та знати принципи побудови архітектури систем обробки даних.

- Спеціалісти з науки про дані мають глибокі знання з машинного навчання, статистики та програмування.

- Інженери хмарних даних займаються підтримкою та оптимізацією хмарних інфраструктур для зберігання та аналізу великих обсягів інформації.

Узгодження цих кластерів з освітніми програмами дозволить забезпечити відповідність знань та компетентностей випускників реальним потребам ринку праці.



Рис. 3. Розподіл навичок Data Specialist по професійних кластерах

Для студентів, які планують кар'єру у сфері науки про дані, важливо обирати такі освітні програми, які відповідають їхнім кар'єрним цілям. Програми мають включати курси з програмування (Python, R), роботи з базами даних (SQL), машинного навчання та хмарних технологій.

На прикладі навчальних дисциплін освітньої програми магістерського рівня Інформаційні системи та технології в Херсонському державному університеті (табл. 1) визначена відповідність між вимогами ринку праці і часткою компетентностей навчальних дисциплін (рис. 4).

Таблиця 1.

**Освітні компоненти програми магістратури з ІСТ
(ОС – обов'язкова компонента, ЕС – вибіркова компонента)**

№	Дисципліни освітньої програми	ECTS
ОС 3	Формальні методи інженерії програмного забезпечення	3
ОС 4	Управління технологіями розробки фінансових інструментів	3
ОС 5	Цифрові валюти та блокчейн-технології	3

OC 6	Моделі прогнозування часових рядів для бізнес-аналітики	6,5
OC 7	Моделювання та проєктування інформаційних систем	3
EC	EC 7, EC 8, EC 9, EC 10	3*4
OC 4	Управління технологіями розробки фінансових інструментів	3

№	Cluster 2	Number	Skill share, SS	ECTS	Competences share, CS	Average, Av	χ^2_{fact}	Courses
1	Python	1409	24.3%	5	13.5%	18.9	1.54	OC 3, EC 7.1, EC 8.4
2	SQL	1067	18.4%	4	10.8%	14.6	0.99	OC 4, EC 8.3
3	Data Engineering	416	7.2%	6	16.2%	11.7	1.74	OC 3, EC 7.1, EC 8.1, EC 9.2
4	Data Analysis	384	6.6%	5.5	14.9%	10.7	1.59	OC 6, EC 7.4, EC 9.1
5	Machine Learning	385	5.8%	5.5	14.9%	10.3	1.99	OC 6, EC 7.4, EC 10.1
6	ETL	290	5.0%	4.5	12.2%	8.6	1.49	OC 6, EC 9.1
7	Data Modelling	278	4.8%	4.5	12.2%	8.5	1.60	OC 7, OC 5
8	Java	272	4.7%	0	0.0%	2.4	2.35	
9	Data Warehousing	254	4.4%	1	2.7%	3.6	0.20	EC 8.1
10	Databricks	208	3.6%	0	0.0%	1.8	1.80	
11	Snowflake	193	3.3%	0	0.0%	1.7	1.65	
12	Azure	137	2.4%	0	0.0%	1.2	1.20	
13	Airflow	115	2.0%	0	0.0%	1.0	1.00	
14	Git	110	1.9%	1	2.7%	2.3	0.07	OC 9
15	Kafka	112	1.9%	0	0.0%	1.0	0.95	
16	DevOps	104	1.8%	0	0.0%	0.9	0.90	
17	Kubernetes	61	1.1%	0	0.0%	0.6	0.55	
18	Docker	57	1.0%	0	0.0%	0.5	0.50	

Рис. 4. Розрахунок часток навичок і компетентностей

Отримані результати демонструють, що розподіл спеціальних компетентностей освітньої програми з інформаційних систем та технологій відповідає вимогам кластера спеціалістів з Data Science та Engineering, і цю освітню програму можна обрати для розвитку відповідних навичок, щоб отримати бажану посаду на ринку праці.

Висновки. Оптимальна кількість кластерів була визначена на основі відкритих даних про посади спеціалістів з даних, методу кластеризації та методу ліктя: аналітик даних та інженер, інженер платформ даних, спеціаліст з Data Science та Engineering, інженер з хмарних даних. Кожен з цих кластерів має унікальну комбінацію від 17 до 20 навичок для спеціалістів з даних.

Загалом, кластеризація дозволяє визначити найбільш затребувані групи спеціалістів з даних з найбільш затребуваними навичками. На основі

порівняння часток навичок у профілі спеціалістів кожного кластера та часток професійних компетентностей в освітніх програмах пропонується підхід до визначення подібностей між вимогами ринку праці та освітньою програмою для підготовки спеціалістів з даних. Представлено приклад застосування цього підходу до кластера спеціалістів з Data Science та Engineering та програми магістратури з інформаційних систем і технологій Херсонського державного університету. Використовуючи статистичний критерій, підтверджується подібність між структурою кластера за навичками та структурою освітньої програми за професійними компетенціями.

ЛІТЕРАТУРА:

1. Tavakoli, M., Kismihok, G., Mol, S.T. «Система персоналізованих рекомендацій на основі інформації ринку праці для безперервного навчання з використанням відкритих освітніх ресурсів (OER)». Праці Міжнародної конференції з підтримки освіти за допомогою комп'ютерів (CSEDU). SciTePress, 2020. DOI: 10.48550/arXiv.2005.07465.
2. Wang, F., Jiang, Z., Li, X., Li, G. «Когнітивні фактори передачі емпіричних інженерних знань: поведінкове та fNIRS-дослідження». *Advanced Engineering Informatics* 2021; 47: 101207. DOI: 10.1016/j.aei.2020.101207.
3. Colombo, E., Mercurio, F., Mezzanzanica, M. «Застосування інструментів машинного навчання для аналізу вакансій в Інтернеті щодо ринку праці та навичок. В «Термінатор чи Джетсони?». Економічні та політичні наслідки штучного інтелекту 2018.
4. Li, X., Jiang, Z., Guan, Y., Li, G., Wang, F. «Сприяння передачі емпіричних інженерних знань в умовах зміни технологічної парадигми: експериментальне дослідження на етапі концептуального проектування». *Advanced Engineering Informatics* 2019; 41: 100927. DOI: 10.1016/j.aei.2019.100927.
5. Molavi, M., Tavakoli, M., Kismihok, G. «Виділення тем з відкритих освітніх ресурсів». Праці Європейської конференції з технологій, що підсилюють освіту. Springer, 2020, сс. 455-460. DOI: 10.1007/978-3-030-57717-9_44.
6. Kobets, V., Osypova, N. V. «Identification of factors for providing the higher education quality assurance for students». *International Journal for Quality Research* 2023; 17(1): 195–208. DOI: 10.24874/ijqr17.01-12.
7. Kobets, V., Yatsenko, V., Buiak, L. «Bridging Business Analysts Competence Gaps: Labor Market Needs Versus Education Standards». *Proceedings of the Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2020. CCIS, 2021. vol. 1308, pp. 22-45. Springer, Cham, DOI: 10.1007/978-3-030-77592-6_2.*
8. Kobets, V.; Tsiuriuta, N.; Lytvynenko, V.; Mykhaylova, V. «Web-service management system for job search using competence-based approach». *Proceedings of 16th International Conference on Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2020. CEUR-WS, 2020. vol.2732, pp. 290-302. Aachen University. – Available at: <https://ceur-ws.org/Vol-2732/20200290.pdf>*
9. Kravtsov, H., Kobets, V. «Evolutionary Revision Model for Improvement of Computer Science Curriculum». *Proceedings of the Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2018. CCIS, 2019. vol 1007, pp. 127-147. Springer, Cham. DOI: 10.1007/978-3-030-13929-2_7.*
10. Kobets, V.; Yatsenko, V.; Buiak, L. «Identifying the gaps in the preparing of a business analyst between the requirements of the labor market and the standards of study programs:

- Case of Ukraine”. Proceedings of 16th International Conference on Information and Communication Technologies in Education, Research, and Industrial Applications. ICTERI 2020, CEUR-WS, 2020. vol.2732, pp. 499-514, Aachen University. - Available at: <https://ceur-ws.org/Vol-2732/20200499.pdf>
11. Kobets, V., Tsiuriuta, N., Lytvynenko, V., Novikov, M., Chizhik, S. “Recruitment Web-Service Management System Using Competence-Based Approach for Manufacturing Enterprises”. Proceedings of Advances in Design, Simulation and Manufacturing II. DSMIE 2019. Lecture Notes in Mechanical Engineering, 2020, pp. 138–148. Springer, Cham. DOI: 10.1007/978-3-030-22365-6_14.

Науковий керівник доктор економічних наук, професор Кобець В.М.