

УДК 378:004.67

DOI: 10.31652/2412-1142-2024-74-49-58

Луценко Галина Василівна

докторка педагогічних наук, доцент,
професор кафедри автоматизації та комп'ютерно-інтегрованих технологій,
Черкаський національний університет імені Богдана Хмельницького,
м. Черкаси, Україна
ORCID ID: 0000-0002-9727-7836
lutsenkog@gmail.com

ВИКОРИСТАННЯ СЕРВІСІВ ВІДКРИТИХ ДАНИХ У ПІДГОТОВЦІ СТУДЕНТІВ ДО ДОСЛІДНИЦЬКОЇ ДІЯЛЬНОСТІ

Анотація. У статті висвітлюються прикладні аспекти підготовки студентів педагогічних спеціальностей до науково-дослідницької діяльності. Важливим складником підготовки студентів спеціальності 014.09 Середня освіта (Інформатика) до роботи з дослідницькими даними визначено формування навичок автономного й відповідального застосування сучасного цифрового інструментарію для їх отримання, обробки й візуалізації. У роботі окреслено проблематику, пов'язану з практичною підготовкою майбутніх учителів інформатики до застосування цифрових технологій і сервісів для роботи з освітніми даними. У статті висвітлюються шляхи впровадження проектно орієнтованого навчання, що трактується як частина загальної стратегії підготовки студентів до написання кваліфікаційної бакалаврської роботи. Описано досвід організації дослідницької діяльності студентів спеціальності 014.09 Середня освіта (Інформатика) у форматі командних проєктів з дисципліни «Аналіз і візуалізація даних». У статті розкрито поняття сервісів відкритих даних, визначено їх характеристики та особливості використання. Наведено приклади сервісів, які містять набори відкрити освітніх даних та можуть використовуватися для підготовки студентів спеціальності 014.09 Середня освіта (Інформатика) до практичної роботи з даними, отриманими в педагогічних дослідженнях. Аналізуються прикладні аспекти використання сервісів відкритих даних Kaggle та UC Irvine Machine Learning Repository для отримання й подальшої обробки даних з освітньої сфери. Визначено, що робота з сервісами відкритих даних не лише покращує аналітичні здібності студентів, але й сприяє формуванню професійної залученості, навичок критичного мислення. Стаття також акцентує увагу на використанні цифрових інструментів та технологій в освітньому процесі для підвищення рівня цифрової грамотності студентів та їх готовності до науково-дослідницької діяльності.

Ключові слова: науково-дослідницька діяльність студентів; відкриті дані; аналіз даних.

1. ВСТУП

Постановка проблеми. Важливим складником освітньо-професійних програм підготовки здобувачів освіти першого та другого рівнів галузі знань 01 Освіта/Педагогіка є науково-дослідницька діяльність. Залучення студентів різних освітніх рівнів до педагогічних досліджень, зокрема, у форматі кваліфікаційної роботи, сприяє їх професійному й особистісному розвитку та має вагомий потенціал для формування емпіричного досвіду роботи з учнями, педагогічної творчості та формування професійної рефлексії. Студенти отримують можливість вивчати актуальні проблеми педагогічної практики та шукати оригінальні підходи і рішення.

Попри те, що кожна зі спеціальностей галузі знань 01 Освіта/Педагогіка, має власну специфіку й особливі вимоги до написання кваліфікаційних робіт, розвинуте критичне мислення, аналітичні навички і вміння працювати з даними є обов'язковими якостями для здобувачів.

Наголосимо також, що отримувати й обробляти дослідницькі дані в сучасних умовах неможливо без використання цифрових технологій. Відповідно системна й послідовна підготовка здобувачів до провадження педагогічних досліджень, зокрема, у частині збору, обробки, аналізу й візуалізації дослідницьких даних, є одним із завдань системи вищої освіти, вирішення якої потребує науково і методично виважених підходів та інструментарію.

Формування навичок використання цифрових технологій та застосунків для роботи з даними досліджень може здійснювати як прямо, у рамках спеціалізованих курсів з математичної статистики й аналізу даних, так і опосередковано, під час виконання студентами командних проєктів тощо.

Ідентифікація й апробація цифрових платформ і сервісів, що можуть використовуватися в освітньому процесі для формування в здобувачів навичок пошуку, якісного й кількісного аналізу дослідницьких даних, є важливим прикладним завданням, вирішення якого дозволить підвищити практичну спрямованість роботи студентів і пов'язати її з реальними освітніми проблемами.

Аналіз останніх досліджень і публікацій. Питання, пов'язані з підготовкою здобувачів у галузі освіти до науково-дослідницької роботи, знаходяться в фокусі активної уваги вітчизняних науковців, серед яких В. Андрущенко, С. Гончаренко, І. Зязюн, В. Кремень, В. Луговий, Н. Ничкало, С. Сисоєва, Т. Кристопчук та ін.

Особливості впровадження цифрових технологій у навчання і науково-дослідницьку діяльність здобувачів галузі знань 01 Освіта/Педагогіка вивчають, зокрема, В. Биков, Р. Гуревич, В. Кухаренко, Л. Лупаренко, Н. Морзе, О. Овчарук, С. Семеріков, О. Спірін та ін [1, 2, 3]. Дослідники наголошують, що підготовка кваліфікованих дослідників, напрацювання шляхів системного вдосконалення цифрової компетентності магістрантів та докторантів у галузі освіти є актуальним завданням, що сприятиме інтенсифікації наукового пошуку.

Здатність автономно і відповідально використовувати цифрові технології на практиці для розв'язування задач у певній предметній галузі, визначено обов'язковим складником системи професійних якостей науковців [4]. Концептуально-референтна Рамка цифрової компетентності педагогічних й науково-педагогічних працівників визначає науково-дослідницьку діяльність як один зі складників сфери с2 «Професійна залученість». Опис відповідного дескриптора включає використання цифрових технологій на різних етапах наукових досліджень, зокрема для пошуку і добору даних, їх організації, якісного і кількісного аналізу та оцінювання на рівні, необхідному для інтегрування знань та розв'язування задач у широких або мультидисциплінарних контекстах. Роль роботи з даними відображається також у дескрипторі с1 к2 «Інформаційна та медіаграмотність» сфери с1 «Цифрова грамотність», що включає вміння досліджувати проблеми за допомогою сучасних цифрових засобів, працювати з великими масивами даних, робити і презентувати висновки [5].

Вивченню проблематики використання цифрового інструментарію в підготовці майбутніх магістрів до провадження науково-педагогічних досліджень присвячені праці О. Бескорсої, Л. Гаврілової, Я. Топольник та А. Яцишин [6, 7, 8, 9]. Дослідницькі уміння студентів та організаційно-педагогічні умови їх формування досліджують Н. Бондаренко, В. Грубінко, О. Дубасенюк, Р. Павлюк, О. Повідайчик, Т. Підгорна, А. Степанюк [10, 11, 12, 13]. Автори наголошують на важливості залучення студентів до науково-дослідницької діяльності, починаючи з молодших курсів. Одним зі способів підтримки студентів у написанні бакалаврських робіт є спеціалізовані курси та тренінги з академічного письма [14].

Водночас, аналіз публікацій та практичний досвід керування науково-дослідницькою роботою студентів, вказує на труднощі, пов'язані зі збором і аналізом даних [15, 16]. Для подолання таких труднощів, науковці пропонують використовувати адаптовані дослідницькі завдання, тривалість та рівень складності яких враховують поточний рівень підготовки здобувачів, розглядається як спосіб створення сприятливих умов для залучення студентів до досліджень.

Мета статті. Метою статті є аналіз можливостей використання сервісів відкритих даних у підготовці студентів педагогічних спеціальностей до дослідницької діяльності та висвітлення прикладних аспектів їх застосування на прикладі сервісів Kaggle і UM Machine Learning.

2. ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

Виклад основного матеріалу. Кваліфікаційна (бакалаврська) робота є обов'язковим компонентом освітньо-професійної програми з підготовки здобувачів першого (бакалаврського) рівня зі спеціальності 014.09 Середня освіта (Інформатика) в Черкаському національному університеті імені Богдана Хмельницького. Підготовка бакалаврської роботи виявляє вміння студента застосовувати наукові методи в освітньому процесі, використовувати різні підходи до вирішення педагогічних проблем, орієнтуватися в інформаційному просторі, а також шукати, аналізувати, візуалізувати та критично оцінювати інформацію, представляти власні наукові результати

Для формування навичок роботи з дослідницькими даними, починаючи з молодших курсів, освітньою програмою передбачено послідовне вивчення дисциплін «Теорія ймовірності та математична статистика», «Аналіз і візуалізація даних» та «Методологія педагогічних досліджень». Обраний підхід забезпечує дотримання принципів науковості, послідовності та системності та дозволяє поступово розвивати навички для роботи з освітніми даними та ефективного використання цифрових технологій для роботи з ними.

Зупинимося детальніше на структурі й наповненні перших двох дисциплін. У контексті змістового наповнення, забезпечується поступове ускладнення матеріалу та адаптація студентів до аналітичної роботи. У курсі «Теорія ймовірності та математична статистика» розглядаються базові теоретичні концепції і методи статистичного аналізу, наприклад, емпіричні розподіли та їх побудова, розрахунок мір центральної тенденції і мір мінливості, вводиться поняття статистичних гіпотез та ін.

Курс «Аналіз та візуалізація даних» додає практичний компонент, пов'язаний з використанням цифрових технологій роботи з даними. Студенти знайомляться з концепціями відкритих даних і вебскрейпінгу, навчаються працювати з реальними наборами даних предметної області, висуваючи і перевіряючи власні гіпотези. Завдання курсу «Аналіз та візуалізація даних» передбачають також роботу з інструментами для створення інфографік, графіків та інтерактивних візуалізацій, що допомагає інтерпретувати результати статистичних аналізів та підвищує ефективність сприйняття інформації.

Вибрані теми з інтелектуального аналізу даних (Data Mining) завершують підготовку, знайомлячи студентів із методами пошуку прихованих закономірностей у великих наборах даних, що дозволяє глибше познайомитися з поняттям освітньої аналітики. Розвиток цифрових технологій, а відповідно спрощення і прискорення процедури збору даних, сприяють інтенсифікації використання освітніх даних для аналізу успішності учнів, оптимізації освітніх програм і окремих дисциплін та оцінки ефективності навчання і викладання, а інтелектуальний аналіз даних надає можливість прогнозувати освітні тенденції.

Зазначимо, що виконання педагогічного дослідження не завжди пов'язане з аналізом великої кількості даних, адже може реалізовуватися, наприклад, у форматі кейс-стаді. Водночас ключові поняття та підходи математичної статистики й аналізу даних є валідними у всіх випадках.

У курсі «Аналіз і візуалізація даних» для майбутніх учителів інформатики передбачено виконання командного проєкту, завданням якого є збір та обробка освітніх даних [17, 18]. В основу проєктно орієнтованого навчання покладено визначення проблеми, яка стає відправною точкою освітнього процесу. У ході роботи над її вирішенням студенти мають самостійно знаходити та обробляти інформацію, інтегруючи емпіричні та теоретичні аспекти [19].

Наукові дослідження підтверджують, що проєктно орієнтоване навчання сприяє розвитку навичок самоспрямованого навчання, взаємодії, комунікації, а також навичок організації та планування діяльності [20]. Крім того, впровадження цього методу у підготовці майбутніх учителів інформатики допомагає їм опанувати активні методи, які вони зможуть застосовувати у власній педагогічній практиці [21].

Тематика командного проекту для студентів спеціальності 014.09 Середня освіта (Інформатика) у межах дисципліни «Аналіз і візуалізація даних» пов'язана з дослідженням трендів і закономірностей освітніх даних. Підбір завдань, що передбачають обробку реальних наборів даних, сприяє глибшому ознайомленню студентів з концепціями та методами освітньої аналітики. Використання проектного підходу в цій дисципліні є частиною стратегії підготовки студентів до написання кваліфікаційної бакалаврської роботи.

Наголосимо, що дані, які використовуються, мають бути реалістичними, пов'язаними з предметною областю та містити достатню кількість різнотипних характеристик, що дозволить застосувати різні методи статистичного аналізу. Також, їх використання в освітньому процесі не має порушувати прав інтелектуальної власності.

Сервіси відкритих даних відповідають вказаним вимогами та є потужним інструментом у підготовці студентів до науково-дослідницької роботи. Відкриті дані – це дані, які вільно доступні для будь-кого для використання, поширення та обробки. Вони можуть бути використані для проведення досліджень, розробки нових продуктів і послуг, а також для прийняття рішень [22].

Перелік сервісів відкритих даних включає, зокрема [23]:

- Google Dataset Search – спеціалізований пошуковий сервіс, що містить наукові, економічні, технічні дані тощо.
- UCI Machine Learning Repository – репозиторій даних орієнтований на машинне навчання.
- European Union Open Data Portal – платформа демографічних, економічних, фінансових даних з інституцій та органів ЄС.
- UN Data – портал даних ООН, що містить набори даних щодо населення, економіки, торгівлі та енергетики тощо.
- GitHub Awesome Public Datasets – репозиторій GitHub.

3. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Розглянемо прикладні аспекти використання наборів відкритих даних сервісів UC Irvine Machine Learning Repository та Kaggle. Ключовим чинником вибору вказаних сервісів, є наявність освітніх наборів даних, обробка яких є складником педагогічних досліджень.

Сервіс UC Irvine Machine Learning Repository є одним із найбільших та найпопулярніших онлайн-ресурсів для збору, зберігання та надання доступу до наборів даних, зокрема для машинного навчання [24]. У репозиторії сервісу знаходяться понад 500 простих за структурою наборів даних з освітньої сфери, медицини, біології, транспорту тощо. Усі дані доступні для безкоштовного завантаження і використання з освітньою метою. Така діяльність сервісу сприяє розробці та тестуванню нових алгоритмів аналізу даних, навчання моделей тощо.

Репозиторій пропонує детальні описи кожного набору даних, що включають інформацію про джерело даних, структуру, опис змінних, а також метадані, що сприяє кращому розумінню їх специфіки та контексту. Багато наборів даних супроводжується також описовими статтями, публікаціями або інструкціями.

Реєстрація в сервісі UC Machine Learning Repository дуже проста та може здійснюватися з використанням корпоративного Google-акаунту студентів.

Наведений в сервісі перелік предметних областей включає соціальні науки, інженерію, фізику та хімію, біологію, бізнес та низку інших.

Для нашого прикладу обрано мультиваріативний набір даних «Students Performance» [25]. Наголосимо, що сервіс підтримує політику доброчесного використання даних, тож обов'язково слід посылатися на них. Опис посылання автоматично формується при натисканні кнопки Cite, кожен з наборів має оригінальний doi.



Рис. 1. Головне вікно сервісу UC Machine Learning Repository

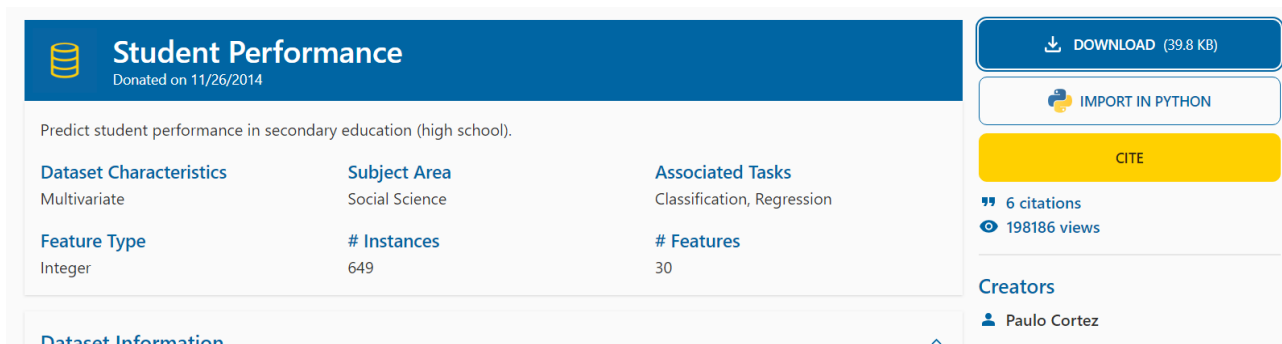


Рис. 2. Вікно роботи з набором даних «Student Performance»

Файл даних у форматі csv завантажується як архів і може бути відкритий за допомогою будь-якого табличного процесора тощо (рис. 3). У випадку нашого курсу, студенти послуговуються Google Таблицями, з доступом через корпоративний акаунт. UC Machine Learning Repository підтримує імпорт даних для обробки в Python (опція «Import in Python»).

Завантажений набір даних містить 30 характеристик, серед яких вік респондентів, стать, освіта й зайнятість батьків, кількість годин витрачених на самостійну роботу, кількість невдало опанованих курсів раніше, оцінки за три періоди та ін. Наявність широкого спектру характеристик є винятково важливим, адже дозволяє використовувати різні формати роботи студентів. Наприклад, для студенти можуть перевіряти гіпотезу про рівність параметрів, використовуючи параметричні або непараметричні критерії.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	school	sex	age	address	famsize	Patatus	Medu	Fedu	Mjob	Fjob	reason	guardian	traveltime	studytime	failures	schoolsup	famsup
2	GP	F	18	U	GT3	A		4	4 at_home	teacher	course	mother	2	2	0	yes	no
3	GP	F	17	U	GT3	T		1	1 at_home	other	course	father	1	2	0	no	yes
4	GP	F	15	U	LE3	T		1	1 at_home	other	other	mother	1	2	0	yes	no
5	GP	F	15	U	GT3	T		4	2 health	services	home	mother	1	3	0	no	yes
6	GP	F	16	U	GT3	T		3	3 other	other	home	father	1	2	0	no	yes
7	GP	M	16	U	LE3	T		4	3 services	other	reputation	mother	1	2	0	no	yes
8	GP	F	16	U	LE3	T		2	2 other	other	home	mother	1	2	0	no	no
9	GP	F	17	U	GT3	A		4	4 other	teacher	home	mother	2	2	0	yes	yes
10	GP	M	15	U	LE3	A		3	2 services	other	home	mother	1	2	0	no	yes
11	GP	M	15	U	GT3	T		3	4 other	other	home	mother	1	2	0	no	yes
12	GP	F	15	U	GT3	T		4	4 teacher	health	reputation	mother	1	2	0	no	yes
13	GP	F	15	U	GT3	T		2	1 services	other	reputation	father	3	3	0	no	yes
14	GP	M	15	U	LE3	T		4	4 health	services	course	father	1	1	0	no	yes
15	GP	M	15	U	GT3	T		4	3 teacher	other	course	mother	2	2	0	no	yes
16	GP	M	15	U	GT3	A		2	2 other	other	home	other	1	3	0	no	yes

Рис. 3. Набір даних, відкритий в Google Таблиці

Сервіс Kaggle фактично є глобальною онлайн-спільнотою фахівців у сфері обробки даних [26]. На платформі опубліковано багато безкоштовних навчальних матеріалів, підтримуються можливості для спілкування користувачів. Також Kaggle підтримує різноманітні формати публікації наборів даних, зокрема csv, JSON, SQLite та низку інших [27].

Як і у випадку UC Machine Learning Repository студенти можуть безкоштовно долучитися до спільноти Kaggle, зареєструвавшись з використанням університетського акаунту Google. На рис. 4. наведено приклад набору даних «Students Performance Factors», що

містить більше 6600 записів для 20 змінних. Серед факторів, що потенційно можуть впливати на успішність студентів, є, зокрема, кількість годин, витрачених на виконання домашніх завдань, відсоток відвіданих занять, середній бал класу тощо. Спектр статистичних методів, які можуть використовуватися, є надзвичайно широким, включаючи t-test, кореляційний та регресійний аналіз тощо.

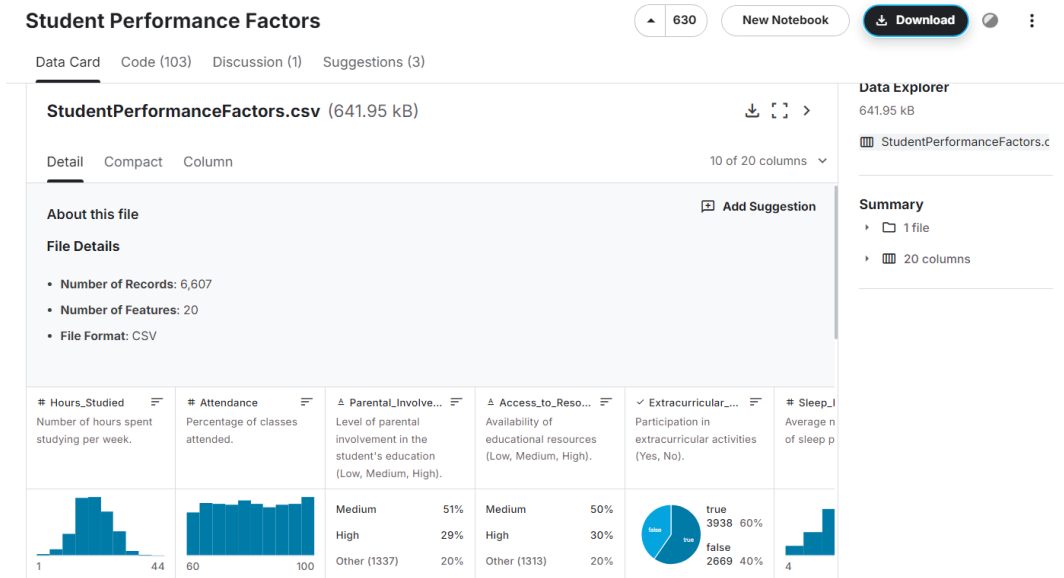


Рис. 4. Вікно роботи з набором даних «Student Performance Factors»

Прикладом дослідження, що може реалізовуватися як частина командного проєкту, є перевірка гіпотези щодо різниці успішності для студентів різної статі. Попередньо, команди мають підготувати отримані дані, очистивши та сформувавши відповідні вибірки у новому аркуші (рис. 5).

D	E	F	G	H	I
Male	Exam	Female	Exam		
Male		67	Female	61	Середня успішність чоловіків
Male		74	Female	70	Середня успішність жінок
Male		71	Female	64	Дисперсія успішності чоловіків
Male		71	Female	65	Дисперсія успішності жінок
Male		67	Female	67	
Male		66	Female	69	
Male		69	Female	66	TTEST
Male		72	Female	63	0,9009672009
Male		68	Female	65	
Male		71	Female	66	

Рис. 5. Приклад розрахунків для перевірки гіпотези про відмінність середньої успішності

Наступним кроком є формулювання нульової гіпотези H_0 , відповідно до якої середні бали успішності чоловіків і жінок не відрізняються та альтернативної – про їх відмінність. Подальша робота з даними включає вибір критерію і рівня значущості для його перевірки. Для наведеного випадку використовувався двосторонній t-тест для вибірок із різними дисперсіями (так званий? t-тест Велча). Він є модифікацією стандартного t-тесту і підходить для ситуацій, коли варіативність даних у групах суттєво різниться, забезпечуючи точніші результати. Отримане розрахункове значення критерію дозволяє стверджувати, що різниця середніх балів успішності між двома групами (чоловіків і жінок) є статистично незначущою.

Значний обсяг вибірок та велика кількість характеристик, що можуть бути проаналізовані, дозволяють урізноманітнити проєктні завдання студентів. Для студентів

спеціальності 014.09 Середня освіта (Інформатика) тривалість проєкту складала 6 тижнів. Команди, що формувалися з 3-4 осіб, з урахуванням індивідуальних сильних сторін кожного з учасників, мали ознайомитися з роботою онлайн-сервісів відкритих даних, а також отримати, очистити й упорядкувати один з наборів. Процедура аналізу включала побудову емпіричних розподілів для ключових показників досліджуваних даних, визначення мір центральної тенденції та мір мінливості. Окрім цього, студенти мали ідентифікувати характеристики, що потенційно можуть корелювати та розрахувати коефіцієнт кореляції. Проєкт включав також перевірку гіпотез і елементи регресійного аналізу. Обов'язковим складником проєкту було визначено візуалізацію даних з використанням оптимальної техніки візуалізації.

4. ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Аналіз наборів даних щодо успішності учнів чи студентів є важливим складником педагогічних досліджень. Працюючи з реалістичними масивами освітніх даних, майбутні учителі інформатики отримують можливість не лише відпрацьовувати навички застосування на практиці апарату математичної статистики й відповідних цифрових інструментів, а й вдосконалюють дослідницькі навички. У ході реалізації командних проєктів, студенти навчаються формулювати нульову та альтернативну гіпотези досліджень, обирати валідні критерії перевірки гіпотез, здійснювати розрахунок мір центральної тенденції і мір мінливості тощо. Наголосимо, що визначені в командних проєктах завдання відповідають принципам реального кваліфікаційного дослідження, що буде реалізовуватися здобувачами освіти протягом фінального року навчання.

Важливо, що обізнаність з предметною областю допомагає також майбутнім учителям інформатики не лише усвідомлено аналізувати отримані результати, а й пропонувати рекомендації щодо покращення успішності (наприклад, проведення додаткових занять для студентів з групи ризику, збільшення годин для вивчення окремих тем тощо).

Перспективи подальших досліджень пов'язані з використанням відкритих наборів даних у навчанні інтелектуального аналізу даних. Їх застосування сприятиме розвитку навичок побудови прогностичних моделей за допомогою методів машинного навчання, зокрема таких як класифікація та регресія. У перспективі, пропонований підхід надасть студентам можливість працювати з реальними, масштабованими наборами інформації, що значно покращить їх навички в обробці великих даних і дозволить розробляти адаптивні системи для підвищення ефективності освітнього процесу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Биков, В., Спирін, О., Білошчинський, А., & Кучанський, О. (2020). Відкриті цифрові системи в оцінюванні результатів науково-педагогічних досліджень. *Information Technologies and Learning Tools*, 75(1), 294-305. doi:10.33407/itlt.v75i1.3589
2. Кухаренко, В. М. (2015). Хмарні технології у наукових дослідженнях. *New computer technology*, 13, 146-157.
3. Шишкіна, М. П. (2018). Використання хмарних технологій для підтримки освітніх досліджень у просторі відкритої науки. *New computer technology*, 16, 105-115.
4. Спирін, О. М., & Одуд, О. А. (2016). Модель формування інформаційно-комунікаційної компетентності доктора філософії на основі використання хмарних сервісів Google Scholar. *Інформаційні технології і засоби навчання*, 56(6), 204-218.
5. МЦТУ. (2021). Концептуально-референтна Рамка цифрової компетентності педагогічних й науково-педагогічних працівників (проєкт). Отримано з https://osvita.diia.gov.ua/uploads/0/2629-frame_pedagogical.pdf
6. Яцишин, А. В. (2020). Цифрові відкриті системи у підготовці аспірантів і докторантів: монографія. Київ: ЦК "Компринт".
7. Топольник, Я. В. (2018). Система інформаційно-комунікаційної підтримки наукових досліджень майбутніх магістрів і докторів філософії в галузі знань „Освіта”: монографія. Слов'янськ: Вид-во Б.І. Маторіна.

8. Топольник, Я. В. (2019). Педагогічні умови ефективності інформаційно-комунікаційної підтримки наукових досліджень майбутніх магістрів і докторів філософії в галузі освіти. *Молодий вчений*, 9(73), 104-108. doi:10.32839/2304-5809/2019-9-73-23
9. Бескорса, О., & Гаврілова, Л. (2023). Цифрові технології в науково-педагогічних дослідженнях майбутніх магістрів початкової освіти. *Професіоналізм педагога: теоретичні й методичні аспекти*, 1, 4-57. doi:10.31865/2414-9292.1.2022.275154
10. Грубінко, В., & Степанюк, А. (2016). Система формування дослідницьких умінь майбутніх учителів природничих дисциплін. *Педагогічні науки: теорія, історія, інноваційні технології*, 2(56), 227-235.
11. Повідайчик, О. С. (2017). Вітчизняний досвід науково-дослідницької підготовки студентів у вищих навчальних закладах. *Науковий вісник Ужгородського університету. Серія: «Педагогіка. Соціальна робота»*, 2(41), 198-201.
12. Повідайчик, О. С., Герич, М. С., & Повідайчик, М. М. (2020). Залучення студентів до науково-дослідницької діяльності з перших курсів як педагогічна умова формування конкурентоздатних фахівців. *Інноваційна педагогіка*, 2(20), 97-100. doi:10.32843/2663-6085-2020-20-2-20
13. Луценко, Г. В., & Луценко, Г. В. (2010). Науково-дослідницька діяльність студентів у творчих колективах як методологічна основа підвищення якості підготовки фахівців. *Вісник Черкаського університету. Серія "Педагогічні науки"*, 181(3), 137-140.
14. Heimdal, A., Svennevig, P. R., & Norheim, T. G. (2021). Students' preparedness towards writing a bachelor thesis: a survey-based investigation. *International Conference on Engineering and Product Design Education*, 9-10 September 2021. Herning, Denmark.
15. van der Marel, I., Munneke, L., & de Bruijn, E. (2022). Supervising graduation projects in higher professional education—A literature review. *Educational Research Review*, 100462.
16. Bayona-Oré, S., & Bazan, C. (2020). Why Students Find It Difficult to Finish their Theses? *Journal of Turkish Science Education*, 17(4), 591-602.
17. Луценко, Г. В. (2024). Розвиток інформаційної грамотності майбутніх учителів інформатики в курсі "Аналіз даних". у О. В. Овчарук (Ред.), *Цифрова компетентність вчителя нової української школи: інновації в умовах змін : монографія* (сс. 143-147). Київ: ІЦО НАПН України.
18. Луценко, Г. В. (2023). Цифрові аспекти командних проєктів для майбутніх учителів інформатики. V Міжнародна науково-методична конференція «Розвиток інтелектуальних умінь і творчих здібностей учнів та студентів у процесі навчання дисциплін природничо-математичного циклу – ІТМ*плюс-2023». Суми, Україна.
19. Mohedo, M. T., & Bújez, A. V. (2014). Project based teaching as a didactic strategy for the learning and development of Basic competences in future teachers. *Procedia-Social and Behavioral Sciences*, 141, 232-236. doi:10.1016/j.sbspro.2014.05.040
20. Sindre, G., Giannakos, M., Krogstie, B. R., Munkvold, R., & Aalberg, T. (2018). Project-Based Learning in IT Education: Definitions and Qualities. *Uniped*, 41(2), 147-63. doi:10.18261/ISSN1893-8981-2018-02-06
21. Romero-García, C., Buzón-García, O., & de Paz-Lugo, P. (2020). Improving Future Teachers' Digital Competence Using Active Methodologies. *Sustainability*, 12. doi:10.3390/su12187798
22. Кубай, Д., & Газін, А. (2016). Відкритий посібник з відкритих даних. Київ. Отримано з <https://socialdata.org.ua/manual/>
23. Pickell, D. (2024). 50 Best Open Data Sources Ready to be Used Right Now. Retrieved from Learn Hub: <https://learn.g2.com/open-data-sources>
24. (2024). Отримано з Офіційний сайт UC Irvine Machine Learning Repository : <https://archive.ics.uci.edu/>
25. Cortez, P. (2008). Student Performance [Dataset]. UCI Machine Learning Repository. doi:10.24432/C5TG7T
26. (2024). Отримано з Офіційний сайт Kaggle: <https://www.kaggle.com/>
27. Луценко, Г. В., Подолян, О. М., & Меркулова, Д. О. (2024). Використання сервісів відкритих даних для організації командних проєктів студентів інформатичних та інженерних спеціальностей. Збірник матеріалів Міжнародної науково-практичної конференції «Цифрова трансформація освіти: теоретико-методичні засади», 28 жовтня 2024 р. Київ.

USING OPEN DATA SERVICES IN PREPARING STUDENTS FOR RESEARCH ACTIVITIES

Lutsenko Galyna Vasylivna

Doctor of Science, Associated Professor

Professor of the Department of Automation and Computer-Integrated Technologies

Bohdan Khmelnytsky National University of Cherkasy, Cherkasy, Ukraine

ORCID ID: 0000-0002-9727-7836

LutsenkoG@gmail.com

Abstract. The article highlights the practical aspects of preparing students in pedagogical specializations for scientific research. A key component of preparing students majoring in Secondary Education (Informatics) (specialty 014.09) for working with research data is identified as the development of skills for the autonomous and responsible use of modern digital tools for data collection, processing, and visualization. The work outlines issues related to the practical training of future informatics teachers in applying digital technologies and services for working with educational data. The article discusses pathways for implementing project-oriented learning as part of the broader strategy of preparing students for writing their bachelor's thesis. It describes the experience of organizing research activities for students specializing in Secondary Education (Informatics) (specialty 014.09) in the format of team projects within the course "Data Analysis and Visualization." The article explains the concept of open data services, identifies their characteristics, and examines their usage. Examples of services that include open educational data sets and can be used to prepare students in the Secondary Education (Informatics) major for practical work with data obtained in educational research are provided. The applied aspects of using open data services such as Kaggle and the UC Irvine Machine Learning Repository for obtaining and processing educational data are analysed. It is determined that working with open data services not only enhances students' analytical skills but also fosters professional engagement and critical thinking skills. The article also emphasizes the use of digital tools and technologies in the educational process to improve students' digital literacy and their readiness for research activities.

Keywords: student research activity; open data; data analysis.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Bykov, V., Spirin, O., Biloshynskiy, A., & Kuchanskyi, O. (2020). Open digital systems in the assessment of scientific and pedagogical research results. *Information Technologies and Learning Tools*, 75(1), 294-305. doi:10.33407/itlt.v75i1.3589
2. Kukharenyko, V. M. (2015). Cloud technologies in scientific research. *New computer technology*, 13, 146-157.
3. Shyshkina, M. P. (2018). Use of cloud technologies to support educational research in the open science space. *New computer technology*, 16, 105-115.
4. Spirin, O. M., & Odud, O. A. (2016). A model for developing information and communication competence for PhD candidates through Google Scholar cloud services. *Information Technologies and Learning Tools*, 56(6), 204-218.
5. MTsTU. (2021). Conceptual Reference Framework for Digital Competence of Pedagogical and Scientific-Pedagogical Staff (project). Retrieved from https://osvita.diiia.gov.ua/uploads/0/2629-frame_pedagogical.pdf
6. Yatsyshyn, A. V. (2020). Digital open systems in the training of postgraduate and doctoral students: monograph. Kyiv: CK "Komprint."
7. Topolnyk, Ya. V. (2018). The information and communication support system for scientific research of future master's and PhD students in the field of "Education": monograph. Slovyansk: B.I. Matorin Publishing.
8. Topolnyk, Ya. V. (2019). Pedagogical conditions for the effectiveness of information and communication support in the scientific research of future master's and PhD students in education. *Young Scientist*, 9(73), 104-108. doi:10.32839/2304-5809/2019-9-73-23
9. Beskorsa, O., & Havrylova, L. (2023). Digital technologies in scientific and pedagogical research for future primary education master's students. *Teacher Professionalism: Theoretical and Methodological Aspects*, 1, 4-57. doi:10.31865/2414-9292.1.2022.275154
10. Hrubinko, V., & Stepanyuk, A. (2016). The system for developing research skills of future teachers of natural sciences. *Pedagogical Sciences: Theory, History, Innovative Technologies*, 2(56), 227-235.
11. Povodaychyk, O. S. (2017). Domestic experience in scientific-research training of students in higher education institutions. *Scientific Bulletin of Uzhhorod University. Series: "Pedagogy. Social Work"*, 2(41), 198-201.
12. Povodaychyk, O. S., Herych, M. S., & Povodaychyk, M. M. (2020). Involving students in research activities from early courses as a pedagogical condition for forming competitive specialists. *Innovative Pedagogy*, 2(20), 97-100. doi:10.32843/2663-6085-2020-20-2-20
13. Lutsenko, H. V., & Lutsenko, H. V. (2010). Student research activity in creative teams as a methodological foundation for improving the quality of specialist training. *Bulletin of Cherkasy University. Series "Pedagogical Sciences"*, 181(3), 137-140.
14. Heimdal, A., Svennevig, P. R., & Norheim, T. G. (2021). Students' preparadness towards writing a bachelor thesis: a survey-based investigation. *International Conference on Engineering and Product Design Education*, 9-10 September 2021. Herning, Denmark.
15. van der Marel, I., Munneke, L., & de Bruijn, E. (2022). Supervising graduation projects in higher professional education—A literature review. *Educational Research Review*, 100462.
16. Bayona-Oré, S., & Bazan, C. (2020). Why Students Find It Difficult to Finish their Theses? *Journal of Turkish Science Education*, 17(4), 591-602.