

Рішення
разової спеціалізованої вченої ради
про присудження ступеня доктора філософії

Здобувач ступеня доктора філософії Кириченко Євген Олександрович, 1998 року народження, громадянин України, освіта вища: закінчив у 2021 році Львівський національний університет імені Івана Франка за спеціальністю «Комп'ютерні науки», працює асистентом кафедри математичних проблем управління і кібернетики у навчально-науковому інституті фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича з 2025 р., аспірант Чернівецького національного університету імені Юрія Федьковича, виконав акредитовану освітньо-наукову програму Інженерія програмного забезпечення (сертифікат №16846).

Разова спеціалізована вчена рада, утворена наказом Чернівецького національного університету імені Юрія Федьковича, Міністерства освіти і науки України, м.Чернівці від «31» березня 2026 року № 112, у складі:

Голови разової
спеціалізованої вченої ради –

Юрія Ушенка, доктора фізико-математичних наук (01.04.05 – Оптика, лазерна фізика), професора, завідувача кафедри комп'ютерних наук навчально-наукового інституту фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича Ушенко Юрій Олександрович.

Рецензентів –

Дмитра Угриня, доктора технічних наук (05.13.06 – «Інформаційні технології»), професора, професора кафедри комп'ютерних наук навчально-наукового інституту фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича,

Сергія Баловсяка, доктора технічних наук (05.13.05 – «Комп'ютерні системи та компоненти»), доцента, доцента кафедри комп'ютерних систем та мереж Чернівецького національного університету імені Юрія Федьковича,

Офіційних опонентів –

Ярослава Виклюка, доктора технічних наук (01.05.02 – «Математичне моделювання та обчислювальні методи»), професора, професора кафедри систем штучного інтелекту Національного

університету «Львівська політехніка»,
Олени Криворучко, доктора технічних наук (05.13.22 – «Управління проектами і програмами»), професора, професора кафедри комп'ютерних систем, мереж та кібербезпеки факультету інформаційних технологій Національного університету біоресурсів і природокористування України),

на засіданні «22» травня 2026 року прийняла рішення про присудження ступеня доктора філософії з галузі знань 12 Інформаційні технології Євгену Кириченку на підставі публічного захисту дисертації «Оптимізація структури гетерогенних даних в Big Data» за спеціальністю 121 Інженерія програмного забезпечення.

Дисертацію виконано у Чернівецькому національному університеті імені Юрія Федьковича, Міністерство освіти і науки України, м. Чернівці.

Науковий керівник Малик Ігор Володимирович, доктор фізико-математичних наук, професор, завідувач кафедри математичних проблем управління і кібернетики навчально-наукового інституту фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича.

Дисертацію подано у вигляді спеціально підготовленого рукопису українською мовою. Дисертація Кириченка Євгена Олександровича на тему «Оптимізація структури гетерогенних даних в Big Data», що подана на здобуття ступеня доктора філософії з галузі знань 12 Інформаційні технології за спеціальністю 121 Інженерія програмного забезпечення є самостійним, завершеним дослідженням, має наукову новизну, практичну значущість і відповідає вимогам Наказу Міністерства освіти і науки України № 40 від 12.01.2017 р. «Про затвердження вимог до оформлення дисертації» (зі змінами, внесеними згідно з Наказом МОН України від 31.05.2019 р. №579) та «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженого Постановою Кабінету Міністрів України від 12 січня 2022 р. № 44 (зі змінами, внесеними згідно з Постановою Кабінету Міністрів України № 341 від 21.03.2022 р., №502 від 19.05.2023 р., №507 від 03.03.2024 р.).

Здобувач має 13 наукових публікацій за темою дисертації, з них 4 праці у наукових фахових виданнях України, одна з яких проіндексована у наукометричній базі Scopus, 1 публікація у закордонному періодичному науковому виданні, проіндексованому у базі даних Scopus (Q3); 6 праць, що засвідчують апробацію матеріалів дисертації та 1 праця, яка додатково відображає наукові результати дисертації:

1. Кириченко О., Кириченко Є. Асимптотичні властивості випадкових матриць = Asymptotic Properties of Randommatrices. Вісник Київського

національного університету імені Тараса Шевченка. Фізико-математичні науки. 2024. 78(1). С. 41–44. (Scopus)

2. Malyk I. V., Kyrychenko Y., Gorbatenko M., Lukashiv T. Data Optimization through Compression Methods Using Information Technology. International Journal of Information Technology and Computer Science. 2025. Vol.17, No.5. P. 84-99. (Scopus, Q3)

3. Кириченко Є. О., Кириченко О. Л. Побудова моделей нейронних мереж у хмарному середовищі засобами AWS EMR. Вчені записки Таврійського національного університету імені В.І. Вернадського. Серія : Технічні науки. 2024. 35(74), №6. С. 75-79.

4. Kyrychenko Y., Malyk I. Information Technology for Data Compression and Transformation by Means of Apache Airflow and Amazon EMR. Security of Infocommunication Systems and Internet of Things. 2025. Vol. 3, no. 1. P. 01004.

5. Кириченко Є., Малик, І. Один результат кластеризації на основі випадкових матриць. Information Technology: Computer Science, Software Engineering and Cyber Security. 2025. №3. С. 257-270.

У дискусії взяли участь голова, рецензенти, офіційні опоненти та висловили зауваження:

Голова спеціалізованої вченої ради Ушенко Юрій Олександрович, доктор фізико-математичних наук, професор, завідувач кафедри комп'ютерних наук навчально-наукового інституту фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича. Зауваження до дисертаційного дослідження відсутні.

Опонент Вихлюк Ярослав Ігорович, доктор технічних наук, професор, професор кафедри систем штучного інтелекту Національного університету «Львівська політехніка». Зауваження:

1. Поряд із добре розробленим математичним апаратом, математична модель містить ряд невідомих параметрів (ваг w) та неоднозначність вибору метрик при використанні DISS. Слід більш детально звернути увагу саме на підбір даних параметрів.

2. У роботі зазначається про широкий спектр проведених експериментальних досліджень, але основна увага приділена опису експериментів над фінансовими даними.

3. Теоретичні результати зосереджені на розробці методів кластеризації графів за допомогою спектральних методів для випадку важких хвостів елементів матриці суміжності графу для $\alpha \in (2,3]$, тобто за умови існування скінченного математичного сподівання. Оскільки відповідна стохастична матриця будується на основі нормування рядків, то варта провести моделювання відповідних графів та дослідити валідність одержаних оцінок для $\alpha \in (1,2]$.

Опонент Криворучко Олена Володимирівна, докторка технічних наук, професорка, професор кафедри комп'ютерних систем, мереж та кібербезпеки факультету інформаційних технологій Національного університету біоресурсів і природокористування України. Зауваження:

1. При викладенні матеріалу у розділі 3 пункт 3.2 (стор. 94-95) бажано було б приділити більше уваги опису нефункціональних вимог до розробленої інформаційної системи, зокрема питанням безпеки, надійності та зручності використання (usability).

2. Обґрунтування хмарної реалізації передбачає, серед іншого, і зниження витрат на зберігання та обробку даних, однак у роботі відсутня розгорнута економічна модель, яка б формально пов'язувала обсяг CDR-представлень, кількість запусків конвеєра та режими роботи EMR з відповідними показниками сукупних експлуатаційних витрат. Доповнення експериментальної частини таким аналізом (наприклад, порівнянням вартості на 1 ГБ вхідних даних у режимах EMR на EC2 та EMR Serverless) надало б отриманим інженерним рішенням додаткового економічного обґрунтування.

3. Архітектура (підрозділ 3.3) жорстко прив'язана до екосистеми AWS: S3, Glue, Athena, EMR, EMR Serverless, Airflow на MWSA. Така vendor-lock-in архітектура суперечить заявленім у вступі принципам переносимості та інтероперабельності інформаційної системи. Бажано було сформулювати архітектуру у термінах абстракцій (object storage, metadata catalog, distributed compute engine, workflow orchestrator) і лише потім зіставити з конкретними сервісами AWS, з паралельними аналогами в Azure (Blob Storage, Synapse, Data Factory) та GCP (Cloud Storage, Dataproc, Composer). У такій формі робота мала б суттєво вищу практичну цінність та довговічність.

4. У таблиці 4.5 наведено зменшення розміру даних «більш ніж на 99,99% без втрати статистично значущої інформації». Твердження потребує строгого доведення: яка саме статистично значуща інформація зберігається? Запропонований CDR за побудовою втрачає індивідуальні значення спостережень, всю автокореляційну структуру часових рядів (для фінансових даних – критично важливу), будь-які залежності, що не вкладаються в перші чотири моменти. Адекватним було б продемонструвати збереження якогось конкретного функціоналу на оригінальних та CDR-представлених даних.

5. Дисертаційна робота добре стилістично і грамотно оформлена, хоча і містить незначну кількість технічних неточностей.

Рецензент Угрин Дмитро Ілліч, доктор технічних наук, професор, професор кафедри комп'ютерних наук навчально-наукового інституту фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича. Зауваження:

1. У розділі 1, присвяченому аналізу спектральних методів кластеризації, наведено ґрунтовний теоретичний огляд сучасних підходів, однак відсутня чітка систематизація обмежень існуючих методів у контексті роботи з гетерогенними табличними даними. Зокрема, доцільно було б узагальнити недоліки розглянутих підходів у вигляді порівняльної таблиці або структурованого висновку, що дозволило б чіткіше обґрунтувати необхідність запропонованого автором методу.

2. У підрозділі 2.3 описано процедуру стиснення числових змінних через вектори моментів до четвертого порядку, проте не наведено кількісного порівняння інформаційних втрат у залежності від порядку моменту, що

ускладнює вибір оптимальної глибини зведення для прикладних задач з високою варіабельністю розподілів.

3. У підрозділі 2.4.1 під час обґрунтування ваг у DISS-метриці подано загальні рекомендації щодо їх встановлення експертом або навчання, проте відсутній систематичний аналіз чутливості метрики до вибору ваг для різних класів корпусів даних, що могло б посилити узагальненість висновків.

4. Розділ 3 містить опис архітектури системи на базі AWS (S3, EMR Serverless, Glue, Athena, Airflow), однак характеризується переважно якісним описом переваг кожного компонента. Не наведено кількісних показників: залежності часу обробки від обсягу даних, залежності продуктивності від кількості еxecutor'ів у Spark-кластері, обсягу використаної оперативної пам'яті, часу cold-start ініціалізації EMR-кластера. Доцільно було б провести експеримент з нарощуванням обсягу даних принаймні до десятків ГБ та виміряти показники прискорення обчислень.

5. Експериментальні дослідження у розділі 4 виконано на фінансових часових рядах із двох відкритих наборів даних, що доцільно було розширити за рахунок корпусів з інших предметних областей (наприклад, біомедичних або сенсорних даних IoT) для підтвердження універсальності запропонованого підходу.

6. У роботі недостатньо уваги приділено аналізу обчислювальної складності запропонованих алгоритмів, зокрема оцінці їх масштабованості при збільшенні кількості ознак та об'єктів у наборі даних. Включення асимптотичних оцінок складності та експериментального підтвердження ефективності алгоритмів при великих обсягах даних дозволило б підсилити практичну значимість дослідження.

7. У дисертаційній роботі спостерігається певна нерівномірність між теоретичною та експериментальною частинами: при достатньо глибокому математичному обґрунтуванні моделей експериментальна валідація могла б бути більш розширеною та різноплановою. Зокрема, додаткове тестування на різномірних типах даних та детальніший аналіз отриманих результатів підвищили б рівень доказовості зроблених висновків.

Рецензент Баловсяк Сергій Васильович, доктор технічних наук, доцент, доцент кафедри комп'ютерних систем та мереж Чернівецького національного університету імені Юрія Федьковича. Зауваження:

1. У розділі 2 відсутні схеми для розроблених алгоритмів.

2. У роботі не приведено оцінку часових витрат на обробку даних в розробленій інформаційній технології, проте, такі часові оцінки були б корисними в процесі практичного застосування розроблених програмних засобів.

3. У дисертації проведено аналіз фінансових даних, проте, для демонстрації можливостей розробленої інформаційної технології було б доцільним детальніше висвітлити результати обробки гетерогенних даних інших типів.

4. У таблиці 4.5 для поля «Space Saved» було б доцільно використати більшу кількість цифр після коми з метою кращого розрізнення результатів обробки даних.

На дисертацію Кириченка Є.О. надійшло три звернення. Вони позитивні та підкреслюють актуальність і значущість проведеного дослідження, хоча містять окремі зауваження та рекомендації щодо перспектив подальших досліджень:

Перше звернення:

Петрик Михайло Романович – доктор фізико-математичних наук, професор, завідувач кафедри програмної інженерії Тернопільського національного технічного університету імені Івана Пулюя. У своєму зверненні Петрик М. Р. зазначив, що на особливу увагу заслуговує сформульований дисертантом критерій визначення оптимальної кількості кластерів, заснований на підрахунку власних значень стохастичної матриці, що перевищують заданий поріг. Такий підхід дозволяє коректно працювати з матрицями суміжності, елементи яких можуть мати розподіли з важкими хвостами та не мати скінченних моментів другого порядку, що є типовим для реальних Big Data систем, де класичні припущення центральної граничної теореми втрачають силу. Поєднання цього критерію з типо-орієнтованою моделлю компактного представлення даних та зваженою метрикою структурної подібності формує логічно замкнений математичний апарат – від класифікації змінних до спектральної оцінки кластерної структури корпусу.

Практична реалізація цього апарату у вигляді масштабованої хмарної інформаційної технології на базі Apache Spark та сервісів AWS свідчить про готовність запропонованих методів до застосування у реальних конвеєрах обробки даних.

Як окремі зауваження можна відзначити досить стислий опис проведених експериментів, а також деякі описки та стилістичні неточності, що не впливають на наукову цінність роботи.

Друге звернення:

Цюцюра Світлана Володимирівна – доктор технічних наук, професор, професор кафедри інженерії програмного забезпечення та кібербезпеки факультету інформаційних технологій Державного торговельно-економічного університету. У зверненні зазначається цінність роботи з позицій інженерії інформаційних систем, яка полягає у цілісному охопленні життєвого циклу технології – від формальної моделі до робочого програмного продукту. Окремо відзначено практичну придатність запропонованого конвеєра. Як побажання до подальшого розвитку роботи слід зауважити, що питанням супроводу запропонованої інформаційної системи – моніторингу конвеєрів, логуванню помилок та збору експлуатаційних метрик на рівні Apache Airflow і Apache Spark – у дисертації приділено менше уваги, ніж проектуванню обчислювального ядра. Опрацювання цього аспекту посилює б експлуатаційну придатність технології у виробничих умовах.

Третє звернення:

Самойленко Ігор Валерійович – доктор фізико-математичних наук, професор, професор кафедри дослідження операцій Київського національного університету імені Тараса Шевченка. Автор зазначає, що актуальним є створення теоретичних і прикладних засад для комплексного аналізу гетерогенних даних, включаючи методи класифікації змінних, формування компактних дескрипторів (CDR), побудову метрик подібності (DISS) та реалізацію відповідної інформаційної системи у масштабованих середовищах. У зверненні відзначається практична цінність дисертації та повнота апробації результатів. Зауваження до роботи відсутні.

Результати відкритого голосування:

«За» – 5 (п'ять) членів ради,

«Проти» – 0 (нуль) членів ради.

На підставі результатів відкритого голосування разова спеціалізована вчена рада присуджує Євгену Кириченку ступінь доктора філософії з галузі знань 12 Інформаційні технології за спеціальністю 121 Інженерія програмного забезпечення.

Відеозапис трансляції захисту дисертації додається.

Голова спеціалізованої вченої ради

PhD 12668,

доктор фізико-математичних наук,
професор, завідувач кафедри
комп'ютерних наук навчально-
наукового інституту фізико-технічних
та комп'ютерних наук Чернівецького
національного університету імені Юрія
Федьковича

Юрій УШЕНКО