

ЗАТВЕРДЖУЮ
Проректор з наукової роботи
Дніпровського національного
університету імені Олеся Гончара
Олег МАРЕНКОВ
«11» 2026 р.

ВИСНОВОК

про наукову новизну, теоретичне та практичне значення результатів дисертації Земляного Олексія Дмитровича на тему «Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних», представленій на здобуття ступеня доктора філософії зі спеціальності 121 Інженерія програмного забезпечення

ВИТЯГ

протоколу №4 засідання міжкафедрального семінару при постійнодіючому семінарі «Актуальні питання оптимізації та дискретної математики» при Науковій раді НАН України з проблеми «Кібернетика» факультету прикладної математики та інформаційних технологій Дніпровського національного університету імені Олеся Гончара від «03» червня 2026 року

ПРИСУТНІ: 38 членів наукового семінару.

ГОЛОВА НАУКОВОГО СЕМІНАРУ: член-кореспондент НАН України, д-р фіз.-мат. наук, проф. Кісельова О.М. (01.05.01 – теоретичні основи інформатики та кібернетики), декан факультету прикладної математики, професорка кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

СЕКРЕТАР ЗАСІДАННЯ: канд. фіз.-мат. наук, доц. Кузенков О.О. (01.05.02 – математичне моделювання та обчислювальні методи) доцент кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара.

ЧЛЕНИ НАУКОВОГО СЕМІНАРУ: д-р фіз.-мат. наук, проф. Гук Н. А. (01.02.04 – механіка деформівного твердого тіла), в. о. проректора з науково-педагогічної роботи, професорка кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

д-р фіз.-мат. наук, проф. Кузьменко В. І. (01.02.04 – механіка деформівного твердого тіла), професор кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

д-р фіз.-мат. наук, проф. Шевельова А.Є. (01.02.04 – механіка деформівного твердого тіла), професорка кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

д-р фіз.-мат. наук, проф. Гарт Л.Л. (01.05.01 – теоретичні основи інформатики та кібернетики), професорка кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

д-р техн. наук, проф. Байбуз О.Г. (05.22.20 – експлуатація та ремонт засобів транспорту), завідувач кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Турчина В.А. (01.05.02 – математичне моделювання та обчислювальні методи), завідувачка кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

канд. техн. наук, доц. Зайцева Т.А. (01.02.04 – механіка деформівного твердого тіла), завідувачка кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Волошко В.Л. (01.02.04 – механіка деформівного твердого тіла), доцент кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

д-р фіз.-мат. наук, проф. Білозьоров В.Є. (01.05.04 – системний аналіз і теорія оптимальних рішень), професор кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

д-р техн. наук, проф. Книш Л.І. (05.14.06 – технічна теплофізика та промислова теплоенергетика), професорка кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. техн. наук, доц. Золотько К.Є. (05.14.04 – промислова теплоенергетика), доцент кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Зайцев В.Г. (01.05.02 – математичне моделювання та обчислювальні методи), доцент кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. техн. наук Дзюба П. А. (01.02.04 – механіка деформівного твердого тіла), доцент кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Хижа О.Л. (01.01.01 – математичний аналіз), доцент кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. техн. наук, доц. Мацуга О.М. (05.13.06 – автоматизовані системи управління та прогресивні інформаційні технології), доцентка кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук Козакова Н. Л. (01.02.04 – механіка деформівного твердого тіла), доцентка кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

канд.фіз.-мат. наук Михальчук Г.Й. (01.02.04 – механіка деформівного твердого тіла), доцентка кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Тонкошкур І.С. (01.02.05 – механіка рідини, газу та плазми), доцент кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук Степанова Н.І. (01.02.04 – механіка деформівного твердого тіла), доцентка кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Сафронова І.А. (01.02.04 – механіка деформівного твердого тіла), доцентка кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

канд. техн. наук, доц. Іванченко М.Г. (05.13.06 – інформаційні технології), доцентка кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Божуха Л.М. (01.01.01 – математичний аналіз), доцентка кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

канд. техн. наук Білобородько О.І. (05.13.06 – автоматизовані системи управління та прогресивні інформаційні технології), доцентка кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

канд. техн. наук, доц. Ємел'яненко Т.Г. (05.13.06 – автоматизовані системи управління та прогресивні інформаційні технології), доцентка кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Наконечна Т.В. (01.01.01 – математичний аналіз), доцентка кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

канд. фіз.-мат. наук, доц. Трофімов О.В. (01.02.04 – механіка деформівного твердого тіла), доцент кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

доктор філософії (PhD) Антонюк В.А. (121 – інженерія програмного забезпечення), доцент кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

доктор філософії (PhD) Єгошкін Д.І., доцент кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

Полонська А.Є., асистентка кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара;

Лисиця Н.М., асистентка кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

Сірик С.Ф., асистентка кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

Красношарпа Д.В., старший викладач кафедри комп'ютерних технологій Дніпровського національного університету імені Олеся Гончара;

Лапець О.В., асистент кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

Мащенко Л.В., старший викладач кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара;

Соломатін В.А., асистент кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара.

Стружко В.Р., асистент кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара.

ЗАПРОШЕННІ ФАХІВЦІ (4 особи, з правом голосу):

д-р техн. наук, проф. Корчинський В.М. (05.01.01 – прикладна геометрія, інженерна графіка), професор кафедри телекомунікаційних систем та мереж факультету фізики, електроніки та комп'ютерних систем, Дніпровський національний університет імені Олеся Гончара;

д-р техн. наук, проф. Голуб С.В. (05.13.06 – інформаційні технології), завідувач кафедри програмного забезпечення автоматизованих систем факультету інформаційних технологій і систем, Черкаський державний технологічний університет;

канд. техн. наук, доц. Клименко С.В. (05.13.06 – інформаційні технології), завідувачка кафедри кібербезпеки та комп'ютерно-інтегрованих технологій, Дніпровський національний університет імені Олеся Гончара;

канд. техн. наук Губський А.М. (05.13.07 – автоматизація процесів керування), доцент кафедри інформатики та програмної інженерії факультету інформатики і обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

На засіданні присутні аспіранти: Форкерт П.П., Божуха Д.І.

Аспіранти участі в голосуванні не брали.

Порядок денний: розгляд та обговорення дисертаційної роботи Земляного Олексія Дмитровича на тему «Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних», поданої на здобуття ступеня доктора філософії за спеціальністю 121 Інженерія програмного забезпечення.

Тема дисертації затверджена на засіданні вченої ради Дніпровського національного університету імені Олеся Гончара (протокол № 4 від 01 грудня 2022 року) та уточнена на засіданні вченої ради факультету прикладної математики та інформаційних технологій Дніпровського національного університету імені Олеся Гончара (протокол № 6 від 23 лютого 2026 року) у формулюванні «Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних».

Науковий керівник: доктор технічних наук, професор Байбуз Олег Григорович.

Підготовка здобувача третього рівня вищої освіти здійснюється за акредитованою освітньо-науковою програмою «Інженерія програмного забезпечення» зі спеціальності 121 Інженерія програмного забезпечення (Сертифікат про акредитацію освітньої програми 13179, дійсний до 01.07.2030 р.).

СЛУХАЛИ:

Обговорення дисертації аспіранта 4 року навчання Земляного Олексія Дмитровича на тему «Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних» на здобуття ступеня доктора філософії зі спеціальності 121 Інженерія програмного забезпечення.

Перевірку на плагіат здійснювала комісія у складі: кандидат фіз.-мат. наук, доцент кафедри обчислювальної математики та математичної кібернетики Кузенков О.О., кандидат фізико-математичних наук, доцент кафедри обчислювальної математики та математичної кібернетики Козакової Н.Л., провідний інженер НДЛ ОСС Яцечко Н.Є. За результатами перевірки дисертаційної роботи на плагіат програмою «Strikeplagiarism» зроблено висновок: дисертаційна робота **Земляного О.Д.** має високий рівень унікальності (**88,93%**) і може бути допущена до захисту.

Робота виконана на **169 сторінках** і містить такі складові частини: анотація, зміст, вступ, основна частина (4 розділи), висновки, список використаних джерел, додатки.

Слово надається аспіранту Землянному О.Д. Регламент виступу – 15 хвилин.

Аспірант Земляний О.Д.:

Добрий день, шановні учасники семінару. Тема моєї дисертаційної роботи: «Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних». Науковий керівник: доктор технічних наук, професор Байбуз Олег Григорович.

Метою дослідження є розроблення та оцінка нових методів імпутування пропусків у даних, які забезпечують високу точність, адаптивність та продуктивність обробки наборів даних з урахуванням кількісних, категоріальних ознак та кореляційних зв'язків.

Об'єктом дослідження є процеси обробки, аналізу та підготовки даних в системах моніторингу.

Предметом дослідження є методи алгоритми та програмні засоби інтелектуального імпутування пропущених значень у вибірках для підвищення ефективності обробки даних в системах моніторингу.

Основні завдання дослідження:

1. Провести аналіз існуючих методів імпутування та виявити переваги та обмеження наявних підходів.

2. Розробити нові методи та алгоритми імпутування, а саме запропонувати нові підходи, орієнтовані на різні типи даних і завдань, зокрема методи на основі класифікації та

регресії, ентропійному підході та гібридні методи, що поєднують ентропійний і регресійний підходи.

3. Розробити алгоритми та програмне забезпечення методів імпутування на мові програмування Python відповідно до архітектурних принципів бібліотеки scikit-learn.

4. Оцінити ефективність нових методів за точністю імпутування та часом обробки на датасетах медичного та гідрологічного моніторингу.

5. Проаналізувати вплив нових методів імпутування на якість моделей класифікації та прогнозування для даних медичного та гідрологічного моніторингу.

Робота містить вступ, 4 розділи, висновки та список використаних джерел, що містить 121 найменування.

Проблема пропусків у даних є актуальною для багатьох прикладних галузей, оскільки вона погіршує якість і надійність аналізу, знижує точність алгоритмів машинного навчання та може призводити до хибних висновків, особливо в задачах класифікації та прогнозування. Це ускладнює створення моделей, які б адекватно описували реальні процеси, і збільшує ймовірність упереджених рішень у таких критичних сферах, як охорона здоров'я, екологічний моніторинг та інші.

У першому розділі дослідження проводиться аналіз сучасних підходів до обробки даних, зокрема проблеми пропусків, яка є однією з ключових перешкод для ефективного машинного навчання. Проведено огляд методів аналізу даних при діагностуванні ішемічної хвороби серця як прикладу задачі класифікації, огляд задачі прогнозування рівня та витрат води басейну ріки Дніпро як прикладу аналізу часових рядів, а також детальний огляд існуючих методів імпутування пропусків у даних, їх класифікацію за механізмами виникнення та стратегії заповнення.

Другий розділ присвячено алгоритмічному забезпеченню імпутування пропусків у даних. У ньому описано розроблені методи імпутування, які враховують структуру даних, патерни пропусків, кореляційні та ентропійні зв'язки між ознаками, а також алгоритми перетворення якісних ознак на кількісні зі збереженням інформації про пропуски. У розділі наведено UML-діаграми діяльності для візуалізації логіки роботи алгоритмів.

Третій розділ дисертації присвячений програмному забезпеченню методів імпутування пропусків у даних, зокрема реалізації розроблених алгоритмів на мові Python з інтеграцією в екосистему бібліотеки scikit-learn, а також оптимізації коду для покращення продуктивності обчислень.

Четвертий розділ дисертації присвячений тестуванню розроблених методів імпутування пропусків у даних. У ньому описано датасети для задач класифікації та прогнозування часових рядів, а також проведено чотири типи тестів. Перший тест порівнює точність імпутування та швидкодію при використанні даних із штучно доданими пропусками з моделюванням певних патернів. Другий і третій типи тестів аналізують вплив імпутування на якість моделей у задачах класифікації та прогнозування відповідно. Тест четвертого типу дозволяє проаналізувати вплив імпутування на зменшення умовної ентропії ознак.

У висновках представлено результати дослідження. Зазначається, що розроблені методи імпутування пропусків у даних підвищують точність імпутування та покращують якість моделей у задачах класифікації та прогнозування.

Наукова новизна одержаних результатів полягає у наступному:

1. Удосконалено метод імпутування NoNa шляхом розробки методу UnifiedClassRegrImputer, який також застосовує комбінацію класифікатора та регресора з урахуванням типу даних ознак, при цьому враховує патерни пропусків у даних та надає змінний порядок обходу ознак залежно від кількості пропусків, що дозволяє підвищити точність імпутування та покращити якість моделей у задачах класифікації та прогнозування.

2. **Вперше** розроблено метод імпутування RegrImputer, який застосовує регресор з уточненням значення у випадку категоріальних даних, що дозволяє значно підвищити швидкодію та точність імпутування, а також покращити якість моделей в задачах класифікації та прогнозування.

3. **Удосконалено** ентропійний метод EntropyImputer шляхом додавання двох модифікацій: однокрокова та ітераційна процедури зменшення умовної ентропії кожної ознаки у послідовному та паралельному виконанні, що дозволяє зменшити умовну ентропію та час виконання алгоритму.

4. **Вперше** розроблено гібридний метод імпутування HybridRegrEntropyImputer, який поєднує ентропійний та регресійний підходи в залежності від типу даних ознаки та патернів пропусків, що дозволяє підвищити точність імпутування та швидкодію у порівнянні з EntropyImputer.

5. **Вперше** розроблено метод імпутування CorrelationImputer на основі виявленого кореляційного зв'язку між ознаками, який автоматизує вибір найкращої моделі залежності серед лінійних та квазілінійних моделей, що дозволяє підвищити точність імпутування, а також точність моделей в задачі прогнозування часових рядів.

6. **Вперше** запропоновано метод імпутування пропусків у часових рядах з сезонною складовою на основі виявленого кореляційного зв'язку між ознаками (метод імпутування CorrelationImputer), але дія виконується для кожного окремого місяця року. Це дозволяє підвищити точність імпутування і якість моделей прогнозування у порівнянні з аналогом без розбиття по місяцям.

7. **Удосконалено** методи перетворення якісних ознак на кількісні шляхом збереження інформації про пропуски та можливістю виконувати зворотне перетворення:

– IgnoreNaNLabelEncoder – модифікація стандартного LabelEncoder, яка не кодує пропуски, зберігає словник та дозволяє зворотне перетворення;

– IgnoreNaNFrequentEncoder – модифікація LabelEncoder та FrequencyEncoder, яка враховує частоту значень, надаючи більшу вагу найпоширенішим або найрідшим категоріям, не кодує пропуски, зберігає словник та дозволяє зворотне перетворення.

Практичне значення роботи полягає у наступному:

1. Розроблено алгоритми та програмне забезпечення для імпутування пропусків у даних, яке реалізує запропоновані методи (UnifiedClassRegrImputer, RegrImputer, EntropyImputer, HybridRegrEntropyImputer, CorrelationImputer, IgnoreNaNLabelEncoder, IgnoreNaNFrequentEncoder). Всі алгоритми написані мовою програмування Python відповідно до архітектурних принципів бібліотеки scikit-learn, що забезпечує їхню уніфіковану інтеграцію у стандартні процеси обробки даних. Це дозволяє використовувати методи в аналітичних конвеєрах (pipeline) разом з імпутуванням, масштабуванням, класифікацією чи прогнозуванням.

2. Розроблені методи дозволяють підвищити точність імпутування пропусків у даних медичного та гідрологічного моніторингу. Ефективність методів за точністю значною мірою залежить від особливостей даних, патернів пропусків та їхньої кількості. Метод RegrImputer виділяється найвищою обчислювальною ефективністю. Метод EntropyImputer забезпечує найбільш значне зменшення умовної ентропії ознак щодо цільової змінної. Метод HybridRegrEntropyImputer демонструє підвищену точність імпутування та обчислювальну ефективність порівняно з EntropyImputer. За наявності кореляційних залежностей між ознаками метод CorrelationImputer забезпечує найвищу точність імпутування. У разі наявності сезонної компоненти в часових рядах та вираженої кореляційної залежності модифікація методу CorrelationImputer з окремим підбором моделей для кожного місяця року демонструє вищу точність імпутування, однак вимагає більших обчислювальних витрат. Метод UnifiedClassRegrImputer виявляє стабільну

ефективність за точністю імпутування при різних налаштуваннях, що залежить від кількості пропусків та патернів у даних.

3. Розроблені методи дозволяють покращити якість моделей класифікації та прогнозування для даних медичного та гідрологічного моніторингу у порівнянні з базовими моделями без імпутування та найпростішим стандартним методом імпутуванням SimpleImputer.

4. Впроваджено результати дисертації в освітній процес кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара. Розроблені методи та програмне забезпечення використовуються при викладанні дисциплін «Аналіз даних на мові Python», «Оптимізація та підвищення продуктивності програмного коду», «Аналіз та візуалізація даних», «Технології пошуку структури в даних» для здобувачів першого (бакалаврського) рівня вищої освіти галузі інформаційних технологій.

5. Розроблені методи та програмне забезпечення можуть бути впроваджені в реальні системи аналізу даних, що дозволить підвищити якість моделей класифікації та прогнозування, зокрема для даних медичного та гідрологічного моніторингу.

Після закінчення доповіді до Земляного О.Д. присутніми були поставлені наступні запитання.

ЗАПИТАННЯ ТА ВІДПОВІДІ

Канд. техн. наук, доцент Клименко Світлана Володимирівна (05.13.06 – Інформаційні технології), завідувач кафедри кібербезпеки та комп'ютерно-інтегрованих технологій Дніпровського національного університету імені Олеся Гончара

1. Чи можна використовувати розроблені методи в інших сферах, або тільки для медичних даних та геологічних?

2. Який розмір вибірок та скільки в них пропусків?

3. Чому вибрана мова програмування Python?

Земляний О.Д.

1. Я розглянув дві задачі інтелектуального аналізу даних: класифікація та прогнозування, і вибрав дві різні сфери застосування саме для того, щоб продемонструвати універсальність моїх методів для таких задач.

2. В датасеті UCI Heart Disease Data 920 рядків, 13 ознак та один цільовий стовпець. Лише 33% записів є повними. Датасет Framingham Heart Study складається з 4240 рядків, 14 ознак і цільового стовпця. 86% екземплярів є повними. Датасет hydro_monitoring містить дані щоденного моніторингу за 24 роки. У стовпці рівня води відсутні 6,3% значень, у стовпці витрат води – 33,3%.

3. Python є популярною мовою розробки програмного забезпечення для обробки даних. Написано багато бібліотек для обробки даних, а писати таку роботу на C++ – це буде означати одноразову розробку з невиправдано великими затратами часу, оскільки багато корисних функцій довелося б писати самому з нуля. Навіть бібліотеки з візуалізацією – це була б складна задача для C++, а на Python можна брати відомі бібліотеки. І якби ми створили бібліотеку імп'ютерів на C++, то її можна було б одноразово використати для поповнення, а інші дії з обробки даних все одно проводити на Python. Також наші імп'ютери (UnifiedClassRegImputer, RegImputer) використовують у собі моделі регресії та класифікатори, які вже розроблені на Python. Ми можемо їх не писати з нуля.

Канд. техн. наук, доцент Іванченко Марина Геннадіївна (05.13.06 – Інформаційні технології), доцент кафедри інженерії програмного забезпечення та інформаційних технологій

Чи досліджували Ви, хто ще до Вас займався цією тематикою? Може якісь школи чи вчені? Що було зроблено до Вас?

Земляний О.Д.

Roderick Little, Donald Rubin – найбільш відомі спеціалісти зі статистичного аналізу даних з пропусками. Вони написали книгу *Statistical Analysis with Missing Data*, яка регулярно оновлюється. Саме вони запропонували класифікацію типів пропусків на MAR, MCAR, MNAR. Але зараз є достатньо багато публікацій з цього питання, бо воно актуальне. Розділ 1 дисертації містить аналіз цих публікацій.

Канд. фіз.-мат. наук, Михальчук Ганна Йосипівна (01.02.04 – механіка деформівного твердого тіла), доцент кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара

Під час доповіді Ви казали, що для оптимізації коду Ви використовували векторизацію та паралельні обчислення. Ви можете трохи детальніше пояснити, як саме це було реалізовано і чи досягли Ви якогось ефекту?

Земляний О.Д.

Для оптимізації коду ми використовували такі підходи:

1. Векторизація обчислень – це заміна циклів на векторні операції, які написані на нативному коді (C++ або Rust). Це дозволило значно прискорити обчислення. Ми спеціально провели дослідження, коли порівнювали швидкість виконання деяких кусків власного коду з використанням циклів та версій з векторизацією.

2. Багатопоточність – використання паралельних обчислень для виконання незалежних задач. Але в Python існує механізм блокування інтерпретатора GIL, який означає що, в один момент тільки один потік може виконуватись інтерпретатором. Тобто, на всі потоки в нас тільки один інтерпретатор, і потоки конкурують за нього. Внаслідок цього паралельність тут лише умовна. Тому ми показали, що багатопоточність не дає виграшу в часі виконання, незважаючи на те, що задача дійсно розділяється на підзадачі. На відміну від багатопоточності, багатопроцесність дозволяє покращити швидкодію, оскільки кожний процес має свій власний інтерпретатор, а тому процеси не конкурують за нього. Ми також показали, що виграш буде помітним, якщо кількість процесів не перебільшує кількості фізичних ядер процесора.

Д-р техн. наук, професор Голуб Сергій Васильович (05.13.06 – Інформаційні технології), завідувач кафедри програмного забезпечення автоматизованих систем факультету інформаційних технологій і систем Черкаського державного технологічного університету

Ви стверджуєте, що розроблений метод дозволяє підвищити точність імпутування, а також покращити якість моделей у задачах класифікації та прогнозування. А яким чином Ви розраховували точність імпутування та як перевіряли, що у Вас покращується якість моделей?

Земляний О.Д.

Ми проводили декілька типів тестів. В контексті цього питання за це відповідали тести типу 1, 2, 3 дослідження.

Тест типу 1 був присвячений аналізу точності імпутування та швидкодії методів. На датасетах без пропусків штучно створювалися пропуски (10–40% за MCAR) з додатковими патернами (стовпці, рядки, один пропуск). Після імпутації порівнювалися RMSE імпутованих даних з оригінальними та час виконання методів. Порівняння представлених методів проводилось з методами бібліотеки scikit-learn. Найкращі результати за точністю та швидкістю показали UnifiedClassRegrImputer, RegrImputer, HybridRegrEntropy, а на датасеті hydro_monitoring – CorrelationImputer. Результати цього тесту можна побачити на слайдах 29 та 30.

Тести типу 2 та 3 відповідали за аналіз впливу імпутування на якість моделей класифікації та прогнозування відповідно. Для задачі класифікації покращення якості моделей перевірялося шляхом порівняння метрик accuracy і recall. Порівнювали моделі, отримані після навчання тільки на повному наборі даних, на наборах даних з імпутацією пропусків стандартними методами, та нашими методами. Для класифікації досліджували модель Random Forest, тут accuracy зростала з 83% до 91% на UCI Heart Disease Data, а recall – з 78% до 90%. Результат цього тесту представлено на слайді 32. Для задачі прогнозування використовували метрику RMSE відхилення прогнозу на тестовій частині часового ряду від оригінальних даних. Порівнювали моделі, навчені на даних без імпутування, на даних з імпутуванням стандартними методами та на даних з імпутацією нашими методами. Досліджували прогнозування методами XGBoost, Prophet, ARIMA. Імпутація представленими в дисертації методами CorrelationImputer, UnifiedClassRegrImputer, RegrImputer давала найкращі результати, тоді як стандартний SimpleImputer міг погіршувати якість. Результат цього тесту представлено на слайді 33.

Канд. техн. наук, доцент Мацуга Ольга Миколаївна (05.13.06 – Автоматизовані системи управління та прогресивні інформаційні технології), доцент кафедри інженерії програмного забезпечення та інформаційних технологій

1. Мені здалося, що ідея методу, що враховує кореляційний зв'язок, не нова. В чому новизна Вашого методу? А якщо між ознаками взагалі не буде в наборі даних якоюсь кореляції, як цей метод застосовувати?

2. Чи є якісь обмеження взагалі у тих методів, які Ви запропонували? Тобто є такі ситуації, коли Ви навпаки не радили би застосовувати Ваші методи?

3. У Вас в одному з наборів даних дуже велика кількість пропусків. Чи можна використовувати такий датасет для Вашого дослідження?

4. Коли Ви працюєте з моделями класифікації, Ви аналізуєте метрики accuracy та recall? Наскільки я знаю, recall завжди аналізується в сукупності з precision. Чому Ви розглядаєте лише recall, а precision ігноруєте?

Земляний О.Д.

1. Щодо ідеї використання кореляційного зв'язку для імпутування, то вона не нова. Я автоматизував вибір найкращої моделі залежності серед лінійних та квазілінійних моделей. Це дозволяє підвищити точність імпутування, а також точність моделей в задачі прогнозування часових рядів. Всередині методу обчислюються коефіцієнти Пірсона та Спірмена. Якщо вони незначимі, то буде використано SimpleImputer.

2. Оскільки наші моделі потребують навчання, то на малих вибірках вони не будуть працювати.

3. Датасет UCI Heart Disease отриманий об'єднанням схожих датасетів із чотирьох кардіологічних центрів Угорщини, Швейцарії та США. В цьому датасеті 920 рядків, 299 екземплярів (33%). Існує дослідження, яке показує, що має сенс об'єднувати схожі датасети (ми це наводили у статті з аналізом методів інтелектуального аналізу даних при прогнозуванні ішемічної хвороби). Це дозволяє покращити моделі класифікації, оскільки залучається більше схожих даних. Тобто збільшується навчальна вибірка.

4. Метрика Accuracy показує відсоток правильних передбачень від усіх випадків. А щодо вибору між recall та precision, поки що ми виходили з того, що нам достатньо визначити, яку частку реальних позитивних випадків (хворих) модель знайшла. Це показує метрика recall. І не звертаємо увагу на помилкову тривогу, яку описує метрика precision. В майбутньому варто розглянути і цей показник. Дякую за запитання та зауваження.

Канд. техн. наук Білобородько Оксана Іванівна (05.13.06 – Автоматизовані системи управління та прогресивні інформаційні технології), доцент кафедри інженерії програмного забезпечення та інформаційних технологій

1. В якому вигляді створено Ваше програмне забезпечення? Як воно реалізовано, де воно існує, як до нього звернутись? Це якийсь локальний репозиторій або використовує якісь хмарні технології. Це бібліотека або це якийсь єдиний програмний продукт? Чи використовували Ви якісь інші бібліотеки в розробці?

2. А які складності були з інтеграцією?

Земляний О.Д.

1. Я розробив бібліотеку, яка зберігається у віддаленому репозиторії на GitLab. При розробці використовувалась бібліотека scikit-learn. Мої методи розроблялись таким чином, щоб їх можна було інтегрувати в архітектуру цієї бібліотеки.

2. Реалізація методів імпутування в архітектурі scikit-learn була виконана шляхом створення класів, які відповідають інтерфейсу цієї бібліотеки. Ми створили класи імпутерів для кожного методу (наприклад, UnifiedClassRegrImputer, EntropyImputer, CorrelationImputer) як нащадки BaseEstimator і TransformerMixin з пакету sklearn.base з методами fit() та transform(), і перевизначили методи fit, transform та fit_transform згідно з алгоритмом конкретного імпутера. Це дозволило використовувати наші методи в pipeline scikit-learn. Основні складнощі були пов'язані з оптимізацією коду для покращення швидкодії. Ми використовували техніку векторизації обчислень та паралельні обчислення (у багатопроцесному виконанні).

Голова семінару, член-кореспондент НАН України, доктор фізико-математичних наук, проф. Кісельова О.М., декан факультету прикладної математики та інформаційних технологій, професор кафедри обчислювальної математики та математичної кібернетики Дніпровського національного університету імені Олеся Гончара:

Запитань більше немає. Переходимо до обговорення дисертаційної роботи. Слово має науковий керівник.

ВИСТУП НАУКОВОГО КЕРІВНИКА:

Дисертаційна робота Земляного Олексія Дмитровича присвячена дослідженню, розробці та впровадженню сучасних методів та програмного забезпечення для імпутування пропусків даних. Робота спрямована на вдосконалення існуючих підходів до обробки даних та розробку нових технологічних рішень для забезпечення більш точного відновлення пропущених значень у даних, що є важливим для підвищення якості аналізу та прийняття рішень у таких галузях, як медицина, екологія, та інші.

Дисертаційні дослідження частково відносяться до тем науково-дослідної роботи № ФПМ-2-22 «Розроблення програмного забезпечення аналізу та кластеризації часових рядів» 2022-2024 рр. номер держреєстрації 0122U001465 та № 58 – ФПМ-2-25 «Розроблення інформаційної технології обробки статистичних даних» 2025-27 рр. номер держреєстрації 0125U002280, що виконуються на кафедрі інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара відповідно до тематичних планів науково-дослідних робіт. Дослідження було частково профінансоване Міністерством закордонних справ Чеської Республіки в межах проекту співробітництва між Дніпровським національним університетом імені Олеся Гончара та Карловим університетом (Чеська Республіка).

Результати дисертаційної роботи є новими, отримані Землянним Олексієм Дмитровичем самостійно. Виконані наукові дослідження вносять вагомий внесок у розвиток сучасних методів обробки даних, зокрема у сфері імпутування пропусків, що є критично важливим для підвищення якості аналізу та прийняття рішень у таких галузях, як медицина, екологічний моніторинг та економіка. Розроблені методи та алгоритми, засновані на поєднанні класифікаційних, регресійних, ентропійних та кореляційних підходів, дозволяють підвищити точність заповнення пропусків, враховуючи складні патерни даних

та залежності між ознаками. Крім того, реалізація цих методів у вигляді програмного забезпечення, інтегрованого в архітектуру scikit-learn, забезпечує уніфіковане застосування розроблених рішень у сучасних аналітичних системах.

Вважаю, що дисертаційна робота Земляного Олексія Дмитровича є завершеним науковим дослідженням, виконана на високому науковому рівні, містить нові теоретично обґрунтовані результати. Розроблені методи та алгоритми пройшли апробацію на реальних даних, зокрема в медичних та гідрологічних задачах, де продемонстрували високу ефективність та точність.

Хочу також зазначити, що мені було приємно працювати з Олексієм Дмитровичем. Він демонстрував глибоке розуміння предметної області, творчий підхід до вирішення наукових завдань та здатність до самостійної роботи. Сподіваюся на Вашу підтримку його дисертаційної роботи, яку я, як науковий керівник, рекомендую до захисту на здобуття ступеня доктора філософії зі спеціальності 121 Інженерія програмного забезпечення.

Дякую за увагу.

В ОБГОВОРЕННІ ДИСЕРТАЦІЇ ЗЕМЛЯНОГО О.Д. ВЗЯЛИ УЧАСТЬ:

Канд. техн. наук, доцент Мацуга Ольга Миколаївна – доцент кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара:

Дисертаційна робота та доповідь Земляного О.Д. справили добре враження. Робота, на мій погляд, дуже хороша, якісна і відповідає спеціальності, і має всі елементи і наукової новизни, і практичної цінності.

Я підтримую цю роботу і пропоную винести позитивне рішення нашого семінару та рекомендувати її до захисту на разовій спеціалізованій вченій раді.

Канд. техн. наук, доцент Клименко Світлана Володимирівна – завідувач кафедри кібербезпеки та комп'ютерно-інтегрованих технологій Дніпровського національного університету імені Олеся Гончара:

Тема роботи Земляного О.Д. є актуальною. Всі формальності з публікаціями виконані, здобувач набув необхідних професійних компетенцій для захисту роботи і отримання ступеня доктора філософії. Етапи дисертаційної роботи Земляного О.Д. висвітлені в його доповіді достатньо повно та ґрунтовно.

Вважаю, що робота Земляного О.Д. повністю відповідає вимогам до дисертацій на здобуття ступеня доктора філософії і може бути рекомендована до захисту на разовій спеціалізованій вченій раді за спеціальністю 121 Інженерія програмного забезпечення.

Д-р техн. наук, професор Голуб Сергій Васильович – завідувач кафедри програмного забезпечення автоматизованих систем факультету інформаційних технологій і систем Черкаського державного технологічного університету:

Я згодний з іншими виступами, вважаю, що робота є завершеною і може бути рекомендована до захисту на разовій спеціалізованій вченій раді за спеціальністю 121 Інженерія програмного забезпечення.

ВИСНОВОК

Актуальність теми дисертації. Широке впровадження цифрових технологій у різних галузях сучасного суспільства супроводжується експоненційним зростанням обсягів даних, які стають основою для автоматизованого аналізу та прийняття рішень у медицині, економіці, екології та інших сферах. Однак під час збору та обробки таких даних виникають значні проблеми, пов'язані з неповнотою інформації. Пропуски у даних можуть з'являтися через технічні збої, помилки введення, об'єднання різнорідних інформаційних джерел або інші фактори. Наслідком цього є зниження якості аналізу та підвищення ризику помилкових висновків, що, у свою чергу, негативно впливає на ефективність управлінських рішень.

Незважаючи на існування численних методів імпутації пропусків у даних, більшість традиційних підходів не враховують складних закономірностей та залежностей, притаманних сучасним наборам даних. Це обмежує їхню застосовність у реальних умовах, де дані часто мають неоднорідну структуру та містять значні обсяги пропусків.

Актуальність цього дослідження полягає у розробці сучасних методів імпутування, які здатні забезпечити високу точність та надійність заповнення пропусків за рахунок врахування структури даних, патернів пропусків, а також кореляційних зв'язків між ознаками. Запропоновані рішення спрямовані на підвищення якості аналізу та зниження ризику упередженості, що є критично важливим для ефективного використання даних у сучасних автоматизованих системах прийняття рішень.

Затвердження теми та плану дисертації. Тема дисертації затверджена на засіданні вченої ради Дніпровського національного університету імені Олеся Гончара (протокол № 4 від 01 грудня 2022 року) та уточнена на засіданні вченої ради факультету прикладної математики та інформаційних технологій Дніпровського національного університету імені Олеся Гончара (протокол № 6 від 23 лютого 2026 року).

Науковим керівником дисертаційної роботи призначено доктора технічних наук, професора Байбуза Олега Григоровича.

Підготовка здобувача третього рівня вищої освіти здійснювалась за акредитованою освітньо-науковою програмою «Інженерія програмного забезпечення» зі спеціальності 121 Інженерія програмного забезпечення (сертифікат про акредитацію освітньої програми № 13179, дійсний до 01.07.2030 р.).

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконувалась у відповідності з індивідуальним планом підготовки аспіранта кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара.

Дослідження здійснювалось в рамках науково-дослідної роботи № ФПМ-2-22 «Розроблення програмного забезпечення аналізу та кластеризації часових рядів» 2022-2024 рр. номер держреєстрації 0122U001465 та № 58 – ФПМ-2-25 «Розроблення інформаційної технології обробки статистичних даних» 2025-27 рр. номер держреєстрації 0125U002280.

Дослідження було частково профінансоване Міністерством закордонних справ Чеської Республіки в межах проєкту № 24-PKVV-UM-011 «Посилення стандартів викладання, досліджень та міжнародного співробітництва в Дніпровському національному університеті імені Олеся Гончара (ДНУ)», реалізованого Карловим університетом і ДНУ: Project “Strengthening the Standards of Teaching, Research and International Cooperation at Oles Honchar Dnipro National University (DNU)” granted by the Ministry of Foreign Affairs of the Czech Republic 2024, Registration Number 24-PKVV-UM-011, Activity 2.1.4: Providing mini-grants for research and publication activities of DNU researchers, project name – Imputation of gaps in hydrological monitoring data.

Публікації та особистий внесок здобувача. За темою дисертації опубліковано 11 робіт: 4 статті у наукових фахових виданнях України категорії Б, 2 статті у інших збірниках наукових праць та 5 тез доповідей у збірниках матеріалів наукових конференцій. Аналіз літературних джерел, розроблення алгоритмів та програмного забезпечення імпутування пропусків у даних, порівняльний аналіз методів, оптимізація коду на мові програмування Python, обробка отриманих результатів виконані автором самостійно. Постановка мети і завдань дослідження, аналіз і узагальнення отриманих результатів проводились спільно з науковим керівником д.т.н., проф. О. Г. Байбузом. Публікації Земляного О.Д. відповідають вимогам пп. 8, 9 «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).

Обґрунтованість наукових положень, висновків і рекомендацій. Обґрунтованість наукових положень, висновків та рекомендацій дисертації забезпечується використанням апробованих математичних моделей, коректністю постановок технічних завдань, а також застосуванням теоретично обґрунтованих технологій і методів, які не суперечать відомим науковим даним. Отримані результати узгоджуються з раніше опублікованими дослідженнями інших авторів, що підтверджує їхню достовірність та несуперечливість.

Наукова новизна одержаних результатів полягає у наступному:

1. **Удосконалено** метод імпутування NoNa шляхом розробки методу UnifiedClassRegrImputer, який також застосовує комбінацію класифікатора та регресора з урахуванням типу даних ознак, при цьому враховує патерни пропусків у даних та надає змінний порядок обходу ознак залежно від кількості пропусків, що дозволяє підвищити точність імпутування та покращити якість моделей у задачах класифікації та прогнозування.

2. **Вперше** розроблено метод імпутування RegrImputer, який застосовує регресор з уточненням значення у випадку категоріальних даних, що дозволяє значно підвищити швидкодію та точність імпутування, а також покращити якість моделей в задачах класифікації та прогнозування.

3. **Удосконалено** ентропійний метод EntropyImputer шляхом додавання двох модифікацій: однокрокова та ітераційна процедури зменшення умовної ентропії кожної ознаки у послідовному та паралельному виконанні, що дозволяє зменшити умовну ентропію та час виконання алгоритму.

4. **Вперше** розроблено гібридний метод імпутування HybridRegrEntropyImputer, який поєднує ентропійний та регресійний підходи в залежності від типу даних ознаки та патернів пропусків, що дозволяє підвищити точність імпутування та швидкодію у порівнянні з EntropyImputer.

5. **Вперше** розроблено метод імпутування CorrelationImputer на основі виявленого кореляційного зв'язку між ознаками, який автоматизує вибір найкращої моделі залежності серед лінійних та квазілінійних моделей, що дозволяє підвищити точність імпутування, а також точність моделей в задачі прогнозування часових рядів.

6. **Вперше** запропоновано метод імпутування пропусків у часових рядах з сезонною складовою на основі виявленого кореляційного зв'язку між ознаками (метод імпутування CorrelationImputer), але дія виконується для кожного окремого місяця року. Це дозволяє підвищити точність імпутування і якість моделей прогнозування у порівнянні з аналогом без розбиття по місяцям.

7. **Удосконалено** методи перетворення якісних ознак на кількісні шляхом збереження інформації про пропуски та можливістю виконувати зворотне перетворення:

– IgnoreNaNLabelEncoder – модифікація стандартного LabelEncoder, яка не кодує пропуски, зберігає словник та дозволяє зворотне перетворення;

– IgnoreNaNFrequentEncoder – модифікація LabelEncoder та FrequencyEncoder, яка враховує частоту значень, надаючи більшу вагу найпоширенішим або найрідшим категоріям, не кодує пропуски, зберігає словник та дозволяє зворотне перетворення.

Практичне значення роботи полягає у наступному:

1. Розроблено алгоритми та програмне забезпечення для імпутування пропусків у даних, яке реалізує запропоновані методи (UnifiedClassRegrImputer, RegrImputer, EntropyImputer, HybridRegrEntropyImputer, CorrelationImputer, IgnoreNaNLabelEncoder, IgnoreNaNFrequentEncoder). Всі алгоритми написані мовою програмування Python відповідно до архітектурних принципів бібліотеки scikit-learn, що забезпечує їхню уніфіковану інтеграцію у стандартні процеси обробки даних. Це дозволяє використовувати методи в аналітичних конвеєрах (pipeline) разом з імпутуванням, масштабуванням, класифікацією чи прогнозуванням.

2. Розроблені методи дозволяють підвищити точність імпутування пропусків у даних медичного та гідрологічного моніторингу. Ефективність методів за точністю значною мірою залежить від особливостей даних, патернів пропусків та їхньої кількості. Метод RegrImputer виділяється найвищою обчислювальною ефективністю. Метод EntropyImputer забезпечує найбільш значне зменшення умовної ентропії ознак щодо цільової змінної. Метод HybridRegrEntropyImputer демонструє підвищену точність імпутування та обчислювальну ефективність порівняно з EntropyImputer. За наявності кореляційних залежностей між ознаками метод CorrelationImputer забезпечує найвищу точність імпутування. У разі наявності сезонної компоненти в часових рядах та вираженої кореляційної залежності модифікація методу CorrelationImputer з окремим підбором моделей для кожного місяця року демонструє вищу точність імпутування, однак вимагає більших обчислювальних витрат. Метод UnifiedClassRegrImputer виявляє стабільну ефективність за точністю імпутування при різних налаштуваннях, що залежить від кількості пропусків та патернів у даних.

3. Розроблені методи дозволяють покращити якість моделей класифікації та прогнозування для даних медичного та гідрологічного моніторингу у порівнянні з базовими моделями без імпутування та найпростішим стандартним методом імпутування SimpleImputer.

4. Впроваджено результати дисертації в освітній процес кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара. Розроблені методи та програмне забезпечення використовуються при викладанні дисциплін «Аналіз даних на мові Python», «Оптимізація та підвищення продуктивності програмного коду», «Аналіз та візуалізація даних», «Технології пошуку структури в даних» для здобувачів першого (бакалаврського) рівня вищої освіти галузі інформаційних технологій.

5. Розроблені методи та програмне забезпечення можуть бути впроваджені в реальні системи аналізу даних, що дозволить підвищити якість моделей класифікації та прогнозування, зокрема для даних медичного та гідрологічного моніторингу.

Список опублікованих праць за темою дисертації

Статті у наукових фахових виданнях України:

1. Земляний О.Д., Байбуз О.Г. Огляд методів інтелектуального аналізу даних та методів машинного навчання при прогнозуванні ішемічної хвороби серця. *Актуальні проблеми автоматизації та інформаційних технологій*, т. 27, Дніпро, 2023, с. 109–129. DOI: <http://dx.doi.org/10.15421/432311> (особистий внесок Земляного О.Д.: проведення аналізу літературних джерел щодо застосування методів машинного навчання для діагностики ішемічної хвороби серця, систематизація існуючих підходів та алгоритмів, а також підготовка порівняльної характеристики наборів даних, аналіз результатів; Байбуза О.Г.: постановка завдання дослідження, узагальнення отриманих результатів).

2. Земляний О.Д., Байбуз О.Г. Методи імпутування пропусків у даних про ішемічну хворобу серця. *Системні технології. Регіональний міжвузівський збірник наукових праць*, вип. 2(151), Дніпро, 2024, с. 33–49. DOI: <https://doi.org/10.34185/1562-9945-2-151-2024-04> (особистий внесок Земляного О.Д.: розроблення та реалізація алгоритмів імпутування пропусків у медичних даних на основі класифікації та регресії для роботи з категоріальними та кількісними ознаками, проведення експериментального тестування запропонованих методів на двох популярних датасетах, оцінка їхньої точності та швидкодії, аналіз впливу імпутування на якість моделей класифікації, програмування алгоритмів, обробка даних, візуалізація та аналіз результатів; Байбуза О.Г.: постановка мети дослідження, контроль та узагальнення отриманих результатів).

3. Земляний О.Д., Байбуз О.Г. Алгоритми імпутування пропусків у даних на основі ентропії. *Системні технології. Регіональний міжвузівський збірник наукових праць*, вип. 6(155), Дніпро, 2024, с. 133–149. DOI: <https://doi.org/10.34185/1562-9945-6-155-2024-12> (особистий внесок Земляного О.Д.: розроблення алгоритму *EntropyImputer* на основі ентропійного підходу до імпутування пропусків, реалізація методу мовою Python, оптимізація коду для підвищення продуктивності, проведення аналізу ефективності запропонованого методу; Байбуза О.Г.: постановка мети та завдання дослідження, контроль та узагальнення отриманих результатів).

4. Земляний О.Д., Байбуз О.Г. Імпутування пропусків у даних гідрологічного моніторингу. *Актуальні проблеми автоматизації та інформаційних технологій*, т. 28, Дніпро, 2024, с. 147–160. DOI: <http://dx.doi.org/10.15421/432413> (особистий внесок Земляного О.Д.: проведення статистичного аналізу та дослідження особливостей даних гідрологічного моніторингу басейну ріки Дніпро, розроблення методів імпутування, адаптованих для часових рядів, реалізація алгоритмів мовою Python, аналіз можливостей щодо оптимізації обчислень, проведення порівняльного аналізу ефективності розроблених методів порівняно з традиційними методами, аналіз результатів; Байбуза О.Г.: постановка задачі, аналіз результатів).

Наукові праці, які засвідчують апробацію матеріалів дисертації:

5. Oleksii Zemlianyi, Oleh Baibuz. Development of an intelligent monitoring system for making decisions about the state of the cardiovascular system. *Матеріали Міжнародної науково-практичної інтернет-конференції «Тенденції та перспективи розвитку науки і освіти в умовах глобалізації»*, вип. 95, Переяслав, 2023, с. 62–63. URL: <https://0a30397da1.clvaw-cdnwnd.com/12ac69b5c0bec343f11779551473023e/200000540-7809b7809d/%D0%97%D0%B1%D1%96%D1%80%D0%BD%D0%B8%D0%BA%2095-5.pdf?ph=0a30397da1> (особистий внесок Земляного О.Д.: розроблення алгоритмів послідовного аналізу для прийняття рішень щодо категорії захворювання серцево-судинної системи на основі методів перевірки гіпотез про параметри нормального розподілу, реалізація програмного забезпечення для аналізу даних добового моніторингу артеріального тиску, проведення порівняльного аналізу класичних та послідовних методів; Байбуза О.Г.: постановка завдання, узагальнення отриманих результатів).

6. Земляний О.Д., Байбуз О.Г. Аналіз існуючих методів інтелектуального аналізу даних при прогнозуванні ішемічної хвороби серця. *Тези доповідей XXI Міжнародної науково-практичної конференції «Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2023)»*, Дніпро, 22-24 листопада 2023 р., Дніпро, 2023, с. 133–134. URL: <http://mpzis.dnu.dp.ua/wp-content/uploads/2023/11/mpzis-2023.pdf> (особистий внесок Земляного О.Д.: проведення огляду літератури щодо застосування методів машинного навчання для діагностики ішемічної хвороби серця, систематизація та порівняльний аналіз алгоритмів, аналіз програмних засобів для обробки даних, підготовка порівняльної характеристики наборів даних, аналіз результатів; Байбуза О.Г.: постановка мети і узагальнення отриманих результатів).

7. Земляний О.Д., Байбуз О.Г. Розроблення методів для імпутування пропусків у даних в архітектурі Scikit-learn Python. *Тези доповідей XXII Міжнародної науково-практичної конференції «Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2024)»*, Дніпро, 20-22 листопада 2024 р., Дніпро, 2024, с. 141–143. URL: <http://mpzis.dnu.dp.ua/wp-content/uploads/2024/11/%D0%9C%D0%9F%D0%97%D0%86%D0%A1-2024-1.pdf> (особистий внесок Земляного О.Д.: розроблення та реалізація класів для імпутування пропусків згідно архітектурних принципів бібліотеки scikit-learn, забезпечення їхньої

сумісності з pipeline, тестування на реальних наборах даних, оцінка точності та швидкодії методів, програмування алгоритмів та обробка даних; Байбуза О.Г.: постановка мети і завдання дослідження, контроль результатів).

8. Земляний О.Д., Байбуз О.Г. Розроблення гібридного методу для імпутування пропусків у даних на основі регресії та ентропійного підходів. *Тези доповідей XXVII Міжнародної конференції «Автоматика 2024»*, Дніпро, 20-22 листопада 2024 р., Дніпро, 2024, с.114. URL: <http://mpzis.dnu.dp.ua/wp-content/uploads/2025/11/%D0%90%D0%B2%D1%82%D0%BE%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0-2024-%D1%82%D0%B5%D0%B7%D0%B8-%D0%B4%D0%BE%D0%BF%D0%BE%D0%B2%D1%96%D0%B4%D0%B5%D0%B9.pdf> (особистий внесок Земляного О.Д.: розроблення гібридного методу HybridRegrEntropyImputer, який поєднує регресійний та ентропійний підходи для імпутування кількісних і якісних ознак, реалізація класу мовою Python згідно архітектурних принципів бібліотеки scikit-learn, проведення порівняльного аналізу ефективності методу з іншими розробленими імпл'ютерами, оцінка якості імпутування, швидкодії та впливу на зменшення умовної ентропії ознак, аналіз результатів; Байбуза О.Г.: постановка задачі дослідження, аналіз і узагальнення результатів).

9. Земляний О.Д., Байбуз О.Г. Підвищення якості моделей класифікації та прогнозування шляхом імпутації пропусків. *Тези доповідей XXIII Міжнародної науково-практичної конференції «Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2025)»*, Дніпро, 19-21 листопада 2025 р., Дніпро: ДНУ, 2025, с.152–154. URL: <http://mpzis.dnu.dp.ua/wp-content/uploads/2025/11/%D0%9C%D0%9F%D0%97%D0%86%D0%A1-2025.pdf> (особистий внесок Земляного О.Д.: розроблення та реалізація методів імпутування пропусків у даних, аналіз ефективності розроблених методів, інтеграція методів в екосистему бібліотеки scikit-learn, проведення апробації на реальних наборах даних, аналіз впливу імпутування на якість моделей класифікації та прогнозування, візуалізація та аналіз результатів; Байбуза О.Г.: постановка мети і завдання дослідження, аналіз і узагальнення отриманих результатів).

Наукові праці в інших виданнях, які додатково відображають зміст дисертації:

10. Земляний О.Д., Байбуз О.Г. Порівняння багатопроекторної та багатопоточної реалізацій ентропійного підходу для імпутування пропусків у даних на мові програмування Python. *Виклики та проблеми сучасної науки*, т. 2, Дніпро, 2024, с. 300–304. DOI: <https://doi.org/10.15421/cims.2> URL: <https://cims.fti.dp.ua/j/article/view/131/159> (особистий внесок Земляного О.Д.: дослідження підходів щодо оптимізації обчислень при реалізації ентропійного підходу для імпутування пропусків у даних на мові Python, проведення програмного експерименту на реальних наборах даних, дослідження впливу GIL на багатопоточну обробку в Python, аналіз результатів; Байбуза О.Г.: постановка задачі, аналіз результатів).

11. Земляний О.Д., Байбуз О.Г. Векторизація обчислень для оптимізації коду на мові програмування Python. *Виклики та проблеми сучасної науки*, т. 3, Дніпро, 2024, с. 144–149. DOI: <https://doi.org/10.15421/cims.3> URL: <https://cims.fti.dp.ua/j/article/view/215/208>. PURL: <https://purl.org/cims/2403.017> (особистий внесок Земляного О.Д.: дослідження застосування векторизації для підвищення продуктивності та читабельності Python-коду, проведення експериментів на реальних наборах даних, аналіз результатів; Байбуза О.Г.: постановка мети дослідження, аналіз результатів).

На підставі заслуховування та обговорення доповіді Земляного О.Д. про основні положення дисертаційної роботи, питань та відповідей на них

УХВАЛИЛИ:

1. Вважати, що за актуальністю, ступенем новизни, обґрунтованості, наукової та практичної цінності здобутих результатів дисертація Земляного Олексія Дмитровича на тему «Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних» відповідає вимогам Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії від 12 січня 2022 р. № 44 (зі змінами).

2. Рекомендувати дисертаційну роботу Земляного Олексія Дмитровича на тему «Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних» до захисту в разовій спеціалізованій вченій раді на здобуття ступеня доктора філософії за спеціальністю 121 Інженерія програмного забезпечення.

3. Клопотати перед вченою радою університету розглянути питання про створення спеціалізованої вченої ради для проведення разового захисту дисертації на здобуття ступеня доктора філософії за спеціальністю 121 Інженерія програмного забезпечення Земляного Олексія Дмитровича у такому складі:

№ з/п	Прізвище, ім'я, по батькові	Місце основної роботи, підпорядкування, посада	Науковий ступінь, шифр, назва спеціальності, за якою захищена дисертація, рік присудження	Вчене звання (за спеціальністю, кафедрою), рік присвоєння	Наукові публікації, опубліковані за останні п'ять років, за науковим напрямом, за яким підготовлено дисертацію здобувача, з яких не менше однієї публікації у виданнях, проіндексованих у базах даних Scopus та/або Web of Science Core Collection
1	2	3	4	5	6
1.	Корчинський Володимир Михайлович (голова)	Дніпровський національний університет імені Олеся Гончара Міністерства освіти і науки України, професор кафедри телекомунікаційних систем та мереж факультету фізики, електроніки та комп'ютерних систем	доктор технічних наук, 05.01.01 – Прикладна геометрія, інженерна графіка, 2000 рік	професор кафедри електронних засобів телекомунікацій, 2003 рік	1. Корчинський В.М., Козарь І.О. Застосування компресії сигналів для управління швидкістю передачі даних у телекомунікаційних каналах з адитивним шумом. <i>Вчені записки Таврійського національного університету імені В.І.Вернадського. Серія: Технічні науки.</i> 2023. Т. 34(73). № 4. С. 1-4. DOI: https://doi.org/10.32782/2663-5941/2023.4/01 (фахове видання категорії Б) 2. Корчинський В.М. Оптимізація пропускної здатності інформаційних каналів передачі растрових зображень. <i>Вчені записки Таврійського національного університету імені В.І.Вернадського. Серія: Технічні науки.</i> 2024. Т. 35(74). № 4. С. 33-37. DOI: https://doi.org/10.32782/2663-5941/2024.4/06 (фахове видання, категорія Б, Index Copernicus) 3. Корчинський В.М., Тимченко О.С. Оптимізаційний метод суміщення багатоканальних інформаційних сигналів. <i>Вчені записки Таврійського національного університету імені В.І.Вернадського. Серія: Технічні науки.</i>

					2025. Т. 36(75). № 4. С. 69-75. DOI: https://doi.org/10.32782/2663-5941/2025.4.1/09 (фахове видання, категорія Б, Index Copernicus)
2.	Губський Андрій Миколайович (опонент)	Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського» Міністерства освіти і науки України, доцент кафедри інформатики та програмної інженерії факультету інформатики і обчислювальної техніки	кандидат технічних наук, 05.13.07 – Автоматизація процесів керування, 2017 рік		<p>1. Марченко О.І., Корнага Я.І., Губський А.М. Динамічний механізм стійкості для масштабованих інформаційних інфраструктур. <i>Вісник Херсонського національного технічного університету</i>, 2025. Т. 2. № 1 (92) С.60-65. DOI: https://doi.org/10.35546/kntu2078-4481.2025.1.2.8 (фахове видання, категорія «Б»)</p> <p>2. Солдатова М.О., Губський А.М., Максимюк А.В., М'яч Д.О. Розподіл ролей користувачів системи рейтингування науково-педагогічних працівників. <i>Телекомунікаційні та інформаційні технології</i>, 2024. С. 106-113. DOI: https://doi.org/10.31673/2412-4338.2024.019901. URL: https://tit.dut.edu.ua/index.php/telecommunication/article/download/2518/2399/ (фахове видання, категорія «Б»)</p> <p>3. Дзівідзінська М., Фіногенов О., Губський А. Інтеграція систем з використанням каталогу довідників. <i>Міжвідомчий науково-технічний збірник «Адаптивні системи автоматичного управління»</i>. 2024. Т. 1. № 44. С. 62–70. DOI: DOI: https://doi.org/10.20535/1560-8956.44.2024.302418 (фахове видання, категорія «Б»)</p> <p>4. Губський А.М. Метод імпутування пропусків у бінарних даних. Актуальні проблеми автоматизації та інформаційних технологій. 2025. Т. 29. С. 179-187. DOI: http://dx.doi.org/10.15421/432516 (фахове видання, категорія «Б»)</p>
3.	Голуб Сергій Васильович (опонент)	Черкаський державний технологічний університет Міністерства освіти і науки України, завідувач кафедри програмного забезпечення автоматизованих систем факультету інформаційних технологій і систем	доктор технічних наук 05.13.06 – Інформаційні технології, 2009 рік	професор кафедри програмного забезпечення автоматизованих систем, 2011 рік	<p>1. Голуб С. В., Толбатов Д.В. Удосконалення методу синтезу багатошарових моделей моніторингового програмного агента. <i>Актуальні проблеми автоматизації та інформаційних технологій</i>. 2023. Т. 27. С. 53-60. DOI: https://doi.org/10.15421/432306. URL: https://actualproblems.dp.ua/index.php/APAIT/article/view/231 (фахове видання, категорія «Б»)</p> <p>2. Голуб С.В., Толбатов Д.В. Машинне навчання багатошарових прогностичних моделей біржових показників. <i>Математичні машини і системи</i>. 2024. № 3-4. С. 100-108. DOI: https://doi.org/10.34121/1028-9763-2024-3-4-100-108. URL: https://www.i-j</p>

					<p>mms.de/index.php/mms/uk/article/view/2024-3-4-a7 (фахове видання, категорія «Б»)</p> <p>3. Голуб С.В., Остапюк В.В. Машинне навчання багат шарових моделей моніторингового програмного агента. <i>Математичні машини і системи</i>. 2025. № 2. С. 76-95. DOI: https://doi.org/10.34121/1028-9763-2025-2-76-95. URL: https://www.t-mms.de/index.php/mms/uk/article/view/2025-2-a10 (фахове видання, категорія «Б»)</p> <p>4. Oleksandr Derevyanchuk, Zhengbing Hu, Serhiy Balovsyak, Serhii Holub, Hanna Kravchenko, Iryna Sapsai, "Complex of Specialized Methods of Educational Data Mining for the Training of Vocational Education Teachers", <i>International Journal of Modern Education and Computer Science (IJMECS)</i>, Vol.17, No.1, pp. 28-46, 2025. DOI: https://doi.org/10.5815/ijmeecs.2025.01.03 (Scopus). URL: https://www.meecspress.org/ijmeecs/ijmeecs-v17-n1/v17n1-3.html</p>
4.	Мацуга Ольга Миколаївна (рецензент)	Дніпровський національний університет імені Олеся Гончара Міністерства освіти і науки України, доцент кафедри інженерії програмного забезпечення та інформаційних технологій	кандидат технічних наук, 05.13.06 – Автоматизовані системи управління та прогресивні інформаційні технології, 2008 рік	доцент кафедри математичного забезпечення електронних обчислювальних машин, 2011 рік	<p>1. Мацуга О.М., Фунтиков М.К. Технологія групування ознак в наборах даних на основі ієрархічної кластеризації. <i>Актуальні проблеми автоматизації та інформаційних технологій</i>: зб. наук. пр. Дніпро, 2024. Т. 28. С. 193-204. DOI: http://dx.doi.org/10.15421/432418 (фахове видання, категорія «Б»)</p> <p>2. Мацуга О.М., Шеремет В.С. Оцінювання оптимальної кількості кластерів для методу k-середніх на основі кусково-лінійної регресії з одним вузлом. <i>Актуальні проблеми автоматизації та інформаційних технологій</i>: зб. наук. пр. Дніпро, 2023. Т. 27. С. 17-24. DOI: http://dx.doi.org/10.15421/432302 (фахове видання, категорія «Б»)</p> <p>3. Мацуга О.М., Приходько М.О., Дроздов В.О. Технологія побудови моделі для прогнозування максимальної корегованої гостроти зору. <i>Актуальні проблеми автоматизації та інформаційних технологій</i>: зб. наук. пр. Дніпро, 2022. Т. 26. С. 72-79. DOI: http://dx.doi.org/10.15421/432209 (фахове видання, категорія «Б»)</p>
5.	Клименко Світлана Володимирівна	Дніпровський національний університет імені Олеся Гончара	кандидат технічних наук, 05.13.06 –	доцент кафедри радіоелектронної автоматики,	<p>1. Малайчук В.П., Клименко С.В., Астахов Д.С. Комп'ютерна обробка вимірювань в задачах спостереження за станом технічних об'єктів. <i>Вісник</i></p>

	рівна (рецензент)	Міністерства освіти і науки України, завідувачка кафедри кібербезпеки та комп'ютерно-інтегрованих технологій	Інформаційн і технології, 2011 рік	2013 рік	<p><i>Дніпровського університету. Серія: Ракетно-космічна техніка. 2022 (опубліковано 2023-02-23). Т. 30. № 4. С. 99-106. DOI: https://doi.org/10.15421/452213. URL: https://rocketspace.dp.ua/index.php/rst/article/view/147 (фахове видання, категорія «Б»)</i></p> <p>2. Димченко А., Клименко С. Алгоритми обчислення кутового відхилення системи обробки оптичних даних. <i>Вісник Дніпровського університету. Серія: Ракетно-космічна техніка. 2025. Т. 34. № 1. С. 70-77. DOI: https://doi.org/10.15421/452507. URL: https://rocketspace.dp.ua/rst/article/view/293 (фахове видання, категорія «Б»)</i></p> <p>3. Федоренко О.Д., Клим В.Ю., Клименко С.В. Непараметрична статистика випадкових величин з невідомими функціями розподілу ймовірності. <i>Системні технології. 2025. Т. 5. № 160. С. 101-111. DOI: https://doi.org/10.34185/1562-9945-5-160-2025-11. URL: https://journals.nmetau.edu.ua/index.php/st/article/view/2229 (фахове видання, категорія «Б»)</i></p>
--	----------------------	--------------------------------------------------------------------------------------------------------------	------------------------------------	----------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Усі кандидатури членів ради відповідають вимогам п.п. 14, 15 «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 (із змінами).

Результати голосування:

«За» – 42 особи,
«Проти» – немає,
«Утримались» – немає.

**Голова
наукового семінару**

Олена КІСЕЛЬОВА

Секретар

Олександр КУЗЕНКОВ