Облікова картка дисертації

I. Загальні відомості

Image not found or type unknow

Державний обліковий номер: 0826U003636

Особливі позначки: відкрита

Дата реєстрації: 15-07-2026

Статус: Наказ про видачу диплома

Реквізити наказу МОН / наказу закладу: Наказ по ДНУ № 1034-с від 22.09.2026 р.

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Земляний Олексій Дмитрович
2. Oleksii D. Zemlianyi

Кваліфікація:

Ідентифікатор ORCID ID: 0009-0001-6157-8725

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 121

Назва наукової спеціальності: Інженерія програмного забезпечення

Галузь / галузі знань:

Освітньо-наукова програма зі спеціальності: Інженерія програмного забезпечення

Дата захисту: 27-08-2026

Спеціальність за освітою: Інженерія програмного забезпечення

Місце роботи здобувача: Організація відсутня

Код за ЄДРПОУ: 00000000

Місцезнаходження: ------, Київ, 00000, Україна

Форма власності: Змішана

Сфера управління:

Ідентифікатор ROR:

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): PhD 16125

Повне найменування юридичної особи: Дніпровський національний університет імені Олеся Гончара

Код за ЄДРПОУ: 02066747

Місцезнаходження: проспект Науки, Дніпро, Дніпровський р-н., 49045, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

IV. Відомості про підприємство, установу, організацію, в якій було виконано дисертацію

Повне найменування юридичної особи: Дніпровський національний університет імені Олеся Гончара

Код за ЄДРПОУ: 02066747

Місцезнаходження: проспект Науки, Дніпро, Дніпровський р-н., 49045, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

V. Відомості про дисертацію

Мова дисертації: Українська

Коди тематичних рубрик: 27.43.51, 50.41.25, 20.54.07

Тема дисертації:

1. Розроблення методів та програмного забезпечення інтелектуального імпутування пропусків даних

2. Development of methods and software for intelligent missing data imputation

Реферат/анотація:

1. Проблема пропусків у даних є актуальною для багатьох прикладних галузей, оскільки вона суттєво погіршує якість 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 ризик упереджених рішень у критичних сферах, таких як охорона здоров'я та екологічний моніторинг. Дисертація присвячена розробленню методів, алгоритмів та програмних засобів інтелектуального імпутування пропусків, які забезпечують точність і ефективність у задачах моніторингу. Автором запропоновано інноваційні підходи, що поєднують класифікацію, регресію, ентропійний аналіз і кореляційні залежності, з урахуванням структури даних та патернів пропусків. Методи реалізовано на мові Python з інтеграцією в екосистему бібліотеки scikit-learn, що забезпечує їх уніфіковане застосування в аналітичних конвеєрах (pipeline) разом з імпутуванням, масштабуванням, класифікацією чи

прогнозуванням. Наукова новизна полягає в удосконаленні та розробці нових методів імпутування: • UnifiedClassRegrImputer – удосконалення методу NoNa шляхом комбінації класифікатора та регресора з урахуванням типу даних і патернів пропусків; • RegrImputer – метод на основі регресора з уточненням значень для категоріальних даних, який значно підвищує швидкодію та точність; • EntropyImputer – удосконалений ентропійний метод з однокроковою та ітераційною процедурами зменшення умовної ентропії; • HybridRegrEntropyImputer – гібридний метод, що поєднує ентропійний та регресійний підходи; • CorrelationImputer – метод на основі кореляційного зв'язку між ознаками, який автоматизує вибір найкращої моделі залежності; • Метод імпутування для часових рядів з сезонною складовою – модифікація CorrelationImputer, яка діє для кожного місяця окремо; • IgnoreNANLabelEncoder та IgnoreNANFrequentEncoder – удосконалені методи перетворення якісних ознак на кількісні зі збереженням інформації про пропуски та можливістю зворотного перетворення. Практичне значення роботи полягає в розробці програмного забезпечення, яке реалізує запропоновані методи та дозволяє підвищити точність імпутування пропусків у даних медичного та гідрологічного моніторингу. Ефективність методів залежить від особливостей даних, патернів пропусків та їхньої кількості. Метод RegrImputer виділяється найвищою обчислювальною ефективністю, EntropyImputer забезпечує найбільше зменшення умовної ентропії, а HybridRegrEntropyImputer демонструє підвищену точність імпутування. За наявності кореляційних залежностей CorrelationImputer забезпечує найвищу точність, а його модифікація для сезонних даних підвищує якість прогнозування, хоча й вимагає більших обчислювальних витрат. Розроблені методи покращують якість моделей класифікації та прогнозування порівняно з базовими моделями без імпутування та найпростішим стандартним методом SimpleImputer. Результати дисертації впроваджено в освітній процес кафедри інженерії програмного забезпечення та інформаційних технологій Дніпровського національного університету імені Олеся Гончара, де вони використовуються при викладанні дисциплін «Аналіз даних на мові Python», «Оптимізація та підвищення продуктивності програмного коду», «Аналіз та візуалізація даних», «Технології пошуку структури в даних».

2. The problem of missing data is relevant for many applied fields, as it significantly degrades the quality of analysis, reduces the accuracy of machine learning algorithms, and can lead to incorrect conclusions, especially in classification and prediction tasks. This complicates the creation of models that would adequately describe real-world processes and increases the risk of biased decisions in critical areas such as healthcare and environmental monitoring. The dissertation is dedicated to the development of methods, algorithms, and software tools for intelligent imputation of missing values, ensuring accuracy and efficiency in monitoring tasks. The author proposes innovative approaches that combine classification, regression, entropy analysis, and correlation dependencies, taking into account the structure of data and missing data patterns. The methods are implemented in Python with integration into the scikit-learn library ecosystem, ensuring their unified application in analytical pipelines (pipeline) alongside imputation, scaling, classification, or forecasting. The scientific novelty lies in the improvement and development of new imputation methods: • UnifiedClassRegrImputer – an enhancement of the NoNa method by combining a classifier and a regressor, considering data types and missing data patterns; • RegrImputer – a method based on a regressor with value refinement for categorical data, significantly improving speed and accuracy; • EntropyImputer – an improved entropy-based method with single-step and iterative procedures for reducing conditional entropy; • HybridRegrEntropyImputer – a hybrid method combining entropy and regression approaches; • CorrelationImputer – a method based on the correlation between features, which automates the selection of the best dependency model; • Seasonal time series imputation method – a modification of CorrelationImputer that operates separately for each month; • IgnoreNANLabelEncoder and IgnoreNANFrequentEncoder – improved methods for transforming qualitative features into quantitative ones while preserving information about missing values and allowing reverse transformation. The practical significance of the work lies in the development of software that implements the proposed methods and allows improving the accuracy of imputation of missing data in medical and hydrological monitoring. The effectiveness of the methods depends on the characteristics of the data, missing data patterns, and their quantity. The RegrImputer method stands out with the highest computational efficiency, EntropyImputer provides the most significant reduction in

conditional entropy, and HybridRegrEntropyImputer demonstrates improved imputation accuracy. In the presence of correlation dependencies, CorrelationImputer ensures the highest accuracy, and its modification for seasonal data improves prediction quality, although it requires greater computational costs. The developed methods improve the quality of classification and prediction models compared to baseline models without imputation and the simplest standard imputation method, SimpleImputer. The results of the dissertation have been implemented in the educational process of the Department of Software Engineering and Information Technology at Oles Honchar Dnipro National University, where they are used in teaching the disciplines «Data Analysis in Python», «Optimization and Improvement of the Productivity of the Software Code», «Data Analysis and Visualization», and «Technologies for Finding Structure in Data».

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки: Інформаційні та комунікаційні технології

Стратегічний пріоритетний напрям інноваційної діяльності: Розвиток сучасних інформаційних, комунікаційних технологій, робототехніки

Підсумки дослідження: Нове вирішення актуального наукового завдання

Публікації:

- 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, 2023, с. 109–129. DOI: http://dx.doi.org/10.15421/432311 URL: https://actualproblems.dp.ua/index.php/APAIT/article/view/236
- Земляний О. Д., Байбуз О. Г. Методи імпутування пропусків у даних про ішемічну хворобу серця. Системні технології. Регіональний міжвузівський збірник наукових праць, т. 2, № 151, 2024, с.33–49. DOI: https://doi.org/10.34185/1562-9945-2-151-2024-04 URL: https://journals.nmetau.edu.ua/index.php/st/article/view/1708/1008
- Земляний О. Д., Байбуз О. Г. Алгоритми імпутування пропусків у даних на основі ентропії. Системні технології. Регіональний міжвузівський збірник наукових праць, т. 6, № 155, 2024, с.133–149. DOI: https://doi.org/10.34185/1562-9945-6-155-2024-12 URL: https://journals.nmetau.edu.ua/index.php/st/article/view/1920/1194
- Земляний О. Д., Байбуз О. Г. Імпутування пропусків у даних гідрологічного моніторингу. Актуальні проблеми автоматизації та інформаційних технологій, т. 28, 2024, с. 147–160. DOI: http://dx.doi.org/10.15421/432413 URL: https://actualproblems.dp.ua/index.php/APAIT/article/view/258
- 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%B A%2095-5.pdf?ph=0a30397da1
- 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
- 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

- 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/wpцcontent/uploads/2025/11/%D0%90%D0%B2%D1%82%D0%BE%D0%BC%D0%B15 0%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%D 0%B9.pdf
- 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/wpцcontent/uploads/2025/11/%D0%9C%D0%9F%D0%97%D0%86%D0%A1-2025.pdf
- Земляний О.Д., Байбуз О.Г. Порівняння багатопроцесорної та багатопоточної реалізацій ентропійного підходу для імпутування пропусків у 16 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
- 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

Наукова (науково-технічна) продукція: технології; методи, теорії, гіпотези; програмні продукти, програмно-технологічна документація

Соціально-економічна спрямованість: поліпшення стану навколишнього середовища; поліпшення якості життя та здоров'я населення, ефективності діагностики та лікування хворих

Охоронні документи на ОПІВ:

Впровадження результатів дисертації: Впроваджено

Зв'язок з науковими темами: 0122U001465, 0125U002280

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Байбуз Олег Григорович

2. Oleh H. Baibuz

Кваліфікація: д. т. н., професор, 05.22.20

Ідентифікатор ORCID ID: 0000-0001-7489-6952

Додаткова інформація: https://www.webofscience.com/wos/author/record/U-4279-2017; https://scholar.google.com/citations?user=lZTaGQsAAAAJ&hl=en

Повне найменування юридичної особи: Дніпровський національний університет імені Олеся Гончара

Код за ЄДРПОУ: 02066747

Місцезнаходження: проспект Науки, Дніпро, Дніпровський р-н., 49045, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Голуб Сергій Васильович

2. Serhii V. Holub

Кваліфікація: д. т. н., професор, 05.13.06

Ідентифікатор ORCID ID: 0000-0002-5523-6120

Додаткова інформація:
http://www.scopus.com/inward/authorDetails.url?authorID=57216031565&partnerID=MN8TOARS;
https://scholar.google.com/citations?hl=en&user=mK5GjwsAAAAJ;
https://scholar.google.com.ua/citations?hl=uk&user=18OIQFIAAAAJ

Повне найменування юридичної особи: Черкаський державний технологічний університет

Код за ЄДРПОУ: 05390336

Місцезнаходження: бульвар Шевченка, Черкаси, Черкаський р-н., 18006, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Губський Андрій Миколайович

2. Andrii M. Hubskyi

Кваліфікація: к. т. н., 05.13.07

Ідентифікатор ORCID ID: 0000-0002-5361-5485

Додаткова інформація:
http://www.scopus.com/inward/authorDetails.url?authorID=57553120300&partnerID=MN8TOARS;
https://publons.com/researcher/D-5369-2016; https://www.webofscience.com/wos/author/record/D-5369-2016; https://scholar.google.com/citations?hl=en&user=CeY9WxwAAAAJ

Повне найменування юридичної особи: Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Берестейський, Київ, 03056, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Мацуга Ольга Миколаївна

2. Olha M. Matsuha

Кваліфікація: к. т. н., доц., 05.13.06

Ідентифікатор ORCID ID: 0000-0001-6444-8566

Додаткова інформація: https://www.webofscience.com/wos/author/record/D-1507-2016; http://www.scopus.com/inward/authorDetails.url?authorID=57215127608&partnerID=MN8TOARS; https://scholar.google.com/citations?hl=en&user=-yl23-AAAAAJ

Повне найменування юридичної особи: Дніпровський національний університет імені Олеся Гончара

Код за ЄДРПОУ: 02066747

Місцезнаходження: проспект Науки, Дніпро, Дніпровський р-н., 49045, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

Власне Прізвище Ім'я По-батькові:

1. Клименко Світлана Володимирівна

2. Svitlana V. Klymenko

Кваліфікація: к. т. н., доц., 05.13.06

Ідентифікатор ORCID ID: 0000-0003-2005-9993

Додаткова інформація: https://www.webofscience.com/wos/author/record/AIC-8105-2022; http://www.scopus.com/inward/authorDetails.url?authorID=35775451500&partnerID=MN8TOARS; https://scholar.google.com/citations?user=_TFGEEMAAAAJ&hl=en

Повне найменування юридичної особи: Дніпровський національний університет імені Олеся Гончара

Код за ЄДРПОУ: 02066747

Місцезнаходження: проспект Науки, Дніпро, Дніпровський р-н., 49045, Україна

Форма власності: Державна

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR:

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради	Корчинський Володимир Михайлович
Власне Прізвище Ім'я По-батькові головуючого на засіданні	Корчинський Володимир Михайлович
Відповідальний за підготовку облікових документів	Антоненко Світлана Валентинівна
Реєстратор	Юрченко Тетяна Анатоліївна
Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності	Юрченко Тетяна Анатоліївна