Облікова картка дисертації

I. Загальні відомості

Image not found or type unknow

Державний обліковий номер: 0820U100074

Особливі позначки: відкрита

Дата реєстрації: 01-07-2020

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Лавренюк Микола Сергійович

2. Lavreniuk Mykola

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: доктор філософії

Аспірантура/Докторантура: так

Шифр наукової спеціальності: 113

Назва наукової спеціальності: Математика та статистика. Прикладна математика

Галузь / галузі знань:

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 25-06-2020

Спеціальність за освітою: Прикладна математика

Місце роботи здобувача:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): ДФ 26.002.001

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

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Перемоги, 37, м. Київ, Київська обл., 03056, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

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

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

Код за ЄДРПОУ: 02070921

Місцезнаходження: проспект Перемоги, 37, м. Київ, Київська обл., 03056, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

V. Відомості про дисертацію

Мова дисертації:

Коди тематичних рубрик: 28.23.15

Тема дисертації:

1. Моделі та методи глибинного навчання для задач геопросторового аналізу

2. Deep learning models and methods for geospatial analysis tasks

Реферат/анотація:

1. Розвиток науки та її застосування в багатьох практичних галузях базуються на використанні штучного інтелекту. Основними його напрямками є інженерія даних, комп'ютерний аналіз, нечітка логіка, машинне сприйняття, видобуток знань та інші. Методологічну основу методів штучного інтелекту складають математичні методи та інтелектуальні моделі, які знаходяться в постійному розвитку. Особливо активно розвиваються математичні моделі та методи, які базуються на великій кількості даних та інкорпорують їх за допомогою методів глибинного навчання. Серед найважливіших практичних застосувань, в яких активно використовуються методи штучного інтелекту, основані на глибинному навчанні з використанням великих об'ємів даних, є дослідження сейсмічної активності Землі, екологічний моніторинг навколишнього середовища, виявлення аномальних даних в кібернетичній безпеці, аналіз геопросторових даних та інші. Одним з важливих аспектів, що зумовили можливість такого розвитку математичних методів глибинного

навчання є поява великих об'ємів доступних даних та обчислюваних потужностей. В зв'язку з цим останні роки швидкими темпами розвивається сфера аналізу багатовимірних зашумлених даних великого об'єму. Разом з тим для задачі аналізу багатовимірних зашумлених даних великого об'єму використання методів глибинного навчання в тому вигляді, в якому вони були розроблені для інших загальновідомих задач, неможливе. Наявні публікації та експерименти з використання глибинного навчання в цій сфері проводились лише для невеликого об'єму даних з рівномірним розподілом у вхідному просторі, що є лише частковим випадком, який неможливо поширити на дані великого об'єму з не рівномірним розподілом у вхідному просторі. Не зважаючи на велику схожість математичної постановки задачі аналізу багатовимірних зашумлених даних та постановок задач для традиційного комп'ютерного аналізу та машинного сприйняття, існують і принципові відмінності між ними. Найбільш значні та успішні результати в області штучного інтелекту з глибинним навчанням були отримані в роботах закордонних авторів Куніхіко Фукусіми, Яна Лекуна, Йошуа Бенджио та Джеффри Хинтона, а також в роботах українських авторів О.Г. Івахненка, М.З. Згуровського, І.В. Сергієнка, Н.Д. Панкратової, О.А. Павлова, Н.М. Куссуль та інших. Разом з тим прогрес потужності обчислювальної техніки та поява у вільному доступі великої кількості багатовимірних даних робить актуальним задачу розробки математичних методів та моделей штучного інтелекту з глибинним навчанням для аналізу багатовимірних зашумлених даних великого об'єму в задачах геопросторового аналізу та екологічного моніторингу. Метою дисертаційного дослідження є розробка та вдосконалення математичних методів глибинного навчання, які базуються на згорткових глибинних нейронних мережах та відрізняються ініціалізацією початкових ваг мереж при використанні немаркованих даних на основі розрідженого кодування, що веде до підвищення точності задач геопросторового аналізу. В дисертації вперше отримані такі нові наукові результати: 1. Вперше розроблено математичний метод уніфікації подання багатовимірних зашумлених геопросторових даних, оснований на розрідженому кодуванні вхідних немаркованих даних, який надає можливість побудови єдиної моделі класифікації для великих об'ємів вхідних даних, що дозволяє отримати вищу точність класифікації. 2. Удосконалено метод глибинного навчання на основі згорткових нейронних мереж, що на відміну від існуючих, ініціалізують початкові ваги не випадковими значеннями, а навчаються виокремлювати ознаки з великих обсягів немаркованих часових рядів багатовимірних даних та забезпечують суттєве підвищення загальної точності задач класифікації. 3. Вперше розроблено метод фільтрації отриманих карт класифікацій геопросторових даних для збільшення їх точності, на основі об'єктного підходу, на відміну від загальноприйнятих методів, що базуються на принципі ковзного вікна, який дозволяє зберегти форму об'єктів на карті. 4. Отримали подальший розвиток запропоновані методи класифікації багатовимірних зашумлених геопросторових даних шляхом реалізації у вигляді потоку виконання з використанням хмарної платформи Amazon, що дозволило зменшити час обробки інформації за рахунок ефективного доступу до даних та розпаралелювання.

2. The development of science and its application in many practical fields are based on the artificial intelligence usage. Data engineering, computer analysis, fuzzy logic, machine perception, data mining and other are its main areas of interest. The methodological basis of the artificial intelligence methods are mathematical methods and intellectual models that are in constant development. Mathematical models and methods that are based on a large amount of data and incorporate them through deep learning methods are particularly active develop. Among the most important practical applications that use artificial intelligence based on deep learning using large amounts of data are studies of Earth's seismic activity, environmental monitoring, anomaly detection in cybersecurity, geospatial data analysis, and more. One of the important aspects that has led to the mathematical deep learning methods development is the emergence of large volumes of available data and the computational power. Due to this, the field of big volume of multidimensional noise data analysis has been rapidly evolving in recent years. However, usage of the deep learning methods as they were designed for other well-known tasks is impossible for the task of big volume of multidimensional noise data analysis. Available publications and experiments on the deep learning usage in this field were conducted only for a small amount of data with uniform distribution in the input space, which is only a partial case that cannot be extended to the large volume of data with uneven distribution in the input space. Despite the great similarity between the mathematical formulation of the multidimensional noisy

data analysis task and the formulation for traditional computer analysis and machine perception tasks, there are fundamental differences between them. The most significant and successful results in the field of artificial intelligence with deep learning were obtained in the works of foreign authors Kunihiko Fukushima, Yann LeCun, Yoshua Bengio and Geoffrey Hinton, as well as in the works of Ukrainian authors O.G. Ivakhnenko, M.S. Zgurovsky, I.V. Sergienko, N.D. Pankratova, O.A. Pavlov, N.M. Kussul and others. However, the progress in computational resources and the emergence of the free accessed large volumes of multidimensional data make it actual to develop mathematical methods and models of artificial intelligence with deep learning to analyze large volumes of multidimensional noisy data in geospatial analysis and environmental monitoring. The aim of the thesis is to develop and improve mathematical methods of deep learning, which are based on convolutional deep neural networks and differ in the initialization of the initial weights of networks with unlabeled data utilization based on sparse coding, which leads to improving accuracy in geospatial analysis tasks. In the dissertation the following new scientific results were first received: 1. For the first time a mathematical method for unifying multidimensional noisy geospatial data has been developed based on the sparse input unlabeled data encoding that provides the possibility to build a single classification model for large volumes of input data, that allows to obtain higher classification accuracy. 2. Deep learning method based on convolutional neural networks has been improved, which, unlike the existing ones, not randomly initialize initial weights, but learn features extraction from the large volumes of unlabeled multidimensional noisy time series of data and provide significant enhancements of the overall classification accuracy. 3. For the first time filtration method for obtained classification maps of geospatial data has been developed for increasing their accuracy, based on object approach, in contrast to commonly accepted methods based on the sliding window principle, that allows to save the shape of objects on the map. 4. Proposed methods for noisy multidimensional geospatial data classifying have had further development through the implementation as a workflow using Amazon cloud based platform, which reduced data processing time through efficient data access and parallelization.

Державний реєстраційний номер ДіР:

Пріоритетний напрям розвитку науки і техніки:

Стратегічний пріоритетний напрям інноваційної діяльності:

Підсумки дослідження:

Публікації:

Наукова (науково-технічна) продукція:

Соціально-економічна спрямованість:

Охоронні документи на ОПІВ:

Впровадження результатів дисертації:

Зв'язок з науковими темами:

VI. Відомості про наукового керівника/керівників (консультанта)

Власне Прізвище Ім'я По-батькові:

1. Новіков Олексій Миколайович
2. Novikov Oleksii M.

Кваліфікація: д. т. н., 05.13.21

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

VII. Відомості про офіційних опонентів та рецензентів

Офіційні опоненти

Власне Прізвище Ім'я По-батькові:

1. Крак Юрій Васильович
2. Krak Yurii V.

Кваліфікація: д. ф.-м. н., 01.05.04

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Власне Прізвище Ім'я По-батькові:

1. Глибовець Микола Миколайович
2. Glibovec Mikola Mikolaiyovich

Кваліфікація: д. ф.-м. н., 01.05.03

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Рецензенти

Власне Прізвище Ім'я По-батькові:

1. Макаренко Олександр Сергійович

2. Makarenko Olexandr S.

Кваліфікація: д. ф.-м. н., 01.05.02

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Власне Прізвище Ім'я По-батькові:

1. Бідюк Петро Іванович

2. Biduk Petro I.

Кваліфікація: д. т. н., 05.13.03

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради Павлов Олександр Анатолійович

Власне Прізвище Ім'я По-батькові головуючого на засіданні Павлов Олександр Анатолійович

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності



Юрченко Т.А.