Облікова картка дисертації

I. Загальні відомості

Державний обліковий номер: 0423U100091

Особливі позначки: відкрита

Дата реєстрації: 16-06-2023

Статус: Захищена

Реквізити наказу МОН / наказу закладу:

II. Відомості про здобувача

Власне Прізвище Ім'я По-батькові:

1. Долгих Сергій Миколайович
2. Dolgikh Serhiy Mykolayovych

Кваліфікація:

Ідентифікатор ORCID ID: Не застосовується

Вид дисертації: кандидат наук

Аспірантура/Докторантура: ні

Шифр наукової спеціальності: 05.13.06

Назва наукової спеціальності: Інформаційні технології

Галузь / галузі знань: Не застосовується

Освітньо-наукова програма зі спеціальності: Не застосовується

Дата захисту: 01-06-2023

Спеціальність за освітою: Теоретична ядерна фізика

Місце роботи здобувача: Національний авіаційний університет

Код за ЄДРПОУ: 01132330

Місцезнаходження: проспект Любомира Гузара, буд. 1, м. Київ, 03058, Україна

Форма власності:

Сфера управління: Міністерство освіти і науки України

Ідентифікатор ROR: Не застосовується

III. Відомості про організацію, де відбувся захист

Шифр спеціалізованої вченої ради (разової спеціалізованої вченої ради): Д 26.062.01

Повне найменування юридичної особи: Національний авіаційний університет

Код за ЄДРПОУ: 01132330

Місцезнаходження: проспект Любомира Гузара, буд. 1, м. Київ, 03058, 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 авіаційний університет

Код за ЄДРПОУ: 01132330

Місцезнаходження: проспект Любомира Гузара, буд. 1, м. Київ, 03058, 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, 50.39

Тема дисертації:

1. Інформаційна технологія розпізнавання мережевих даних Інтернет на основі генеративних нейромережевих моделей

2. Information technology of learning and classification of Internet packet traffic data based on generative neural models

Реферат/анотація:

1. У роботі проведено дослідження, теоретичне обґрунтування, розробку, програмне виконання та експериментальну перевірку методів навчання розпізнавання класів даних пакетів трафіку Інтернет та інших типів з даними навчання мінімального обсягу на основі структури щільності генеративних представлень даних та запропоновано інформаційну технологію обробки даних та навчання машинних систем на основі структури щільності генеративних представлень. У теоретичній частині роботи досліджувалися методи створення інформативних генеративних представлень та доведено теорему про категоризацію в генеративних представленнях, що лежить в основі методів навчання з мінімальними наборами відомих даних, запропонованих у роботі. На підставі результатів теоретичної частини та огляду сучасних методів та моделей навчання штучних систем, запропоновані методи навчання з використанням неконтрольованої генеративної структури (ландшафту щільності) представлень даних Інтернет: метод виявлення характерних

типів даних без вимог відомих даних; та метод ітеративного навчання на генеративному ландшафті з мінімальними наборами навчальних даних, до кількох зразків. На основі результатів теоретичних досліджень та експериментальної перевірки запропонованих методів запропоновано інформаційну технологію навчання з використанням неконтрольованої генеративної структури (ландшафту щільності) представлень, яка з'єднує обробку даних, навчання генеративних моделей та виявлення інформаційної структури даних у єдиний процес, який може застосовуватися з даними різних джерел та типів. Результати роботи підтверджуються ретельним аналізом теоретичних основ, досконалою експериментальною перевіркою та рецензованими публікаціями в українських та міжнародних наукових виданнях.

2. In the thesis, a research into theoretical foundations of unsupervised generative learning, architecture of generative models, design and development, implementation and experimental verification was carried out to propose and verify methods and an information technology of training machine intelligence systems with minimal sets of known data based on generative density structure (landscape) of informative representations created by generative models in the process of unsupervised training with minimization of generative error. Developing such methods is an essential challenge in a number of critical applications including analysis and classification of data in computer networks and Internet. As was established in a number of studies, applying conventional methods with standard sets of training data can affect generality and accuracy of methods in practical applications where data in the networks differs significantly from the sources of training data. The proposed methods are based on the informative structure of unsupervised generative representations produced with models of generative self-learning that do not require known data to produce. Completely unsupervised methods of determination of generative structure of informative representations proposed and verified in the thesis can produce additional essential information about the input distributions to a learning model and allow to significantly reduce the requirement for known data to achieve confident learning of both externally known classes and the common general types or “natural concepts” in the data, offering a natural solution to the identified challenges in the stated problem of Internet traffic classification. In the theoretical part of the thesis, methods of creating informative generative representations were investigated and a theorem of categorization in generative representations proven under a number of identified conditions. The theorem establishes a theoretical foundation for introduction and definition of methods of learning characteristic types (native concepts) and known classes of Internet packet data with minimal sets of training samples based on the density cluster structure in the latent distributions of data proposed and developed in the thesis. The methods use the cluster structure of density distributions in the informative low-dimensional generative representations of Internet packet data, created in the process of unsupervised generative learning to produce latent samples associated with natural concepts or a known classes of interest and construct classifiers of classes and natural concepts with improved accuracy results and reduced dependency on the significant amounts of training data. The proposed approach has a number of essential advantages compared to conventional supervised methods of machine intelligence, including: flexibility, in learning specific classes and concepts of interest without the constraints of confident knowledge of the complete conceptual structure of the data; the ability to learn iteratively, starting with minimal known samples (down to a handful of samples) and improve learning results when new data becomes available without full retraining of the generative model; massively reduced requirement for prior known training data; and, in a strong correspondence to the stated problem of the thesis, reduce to the minimum the dependence of the learning success on the source of training data via employing natural generative structure of the latent distributions of the data in the network. As well, the proposed methods have interesting parallels to learning of biological systems that is characterized by flexibility and ability to learn successfully with minimal data as and when it becomes available. On the base of methods proposed and verified in the thesis, the information technology of minimal sample learning based on density structure (landscape) of informative generative representations was developed. The technology combines the stages of: data processing; selection and training of generative models in an unsupervised process; determination of the density structure of latent representations and learning based on the identified generative structure (landscape) of generative representations into a single information process that can be generalized and extended to data of different types and origin in different domains and problem areas. The results of the thesis are

supported by a thorough review of the theoretical foundations of the problem and the existing approaches in Internet data analysis and classification, comprehensive design of the models based on solid theoretical foundations, extensive and comprehensive experimental verification; presentation and positive acceptance of the results by the research community at international and Ukrainian scientific conferences and seminars;and peer-reviewed publications in Ukrainian and international scientific literature.

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. Приставка Пилип Олександрович
2. Prystavka Pylyp Oleksandrovych

Кваліфікація: д. т. н., 05.13.06

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. Терейковський Ігор Анатолійович

2. Tereikovskyi Ihor Anatolievych

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

Власне Прізвище Ім'я По-батькові:

1. Мирутенко Лариса Вікторівна
2. Myrutenko Larisa Viktorivna

Кваліфікація: к. т. н., 05.13.06

Ідентифікатор ORCID ID: Не застосовується

Додаткова інформація:

Повне найменування юридичної особи:

Код за ЄДРПОУ:

Місцезнаходження:

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Рецензенти

VIII. Заключні відомості

Власне Прізвище Ім'я По-батькові голови ради Луцький Максим Георгійович

Власне Прізвище Ім'я По-батькові головуючого на засіданні Гнатюк Сергій Олександрович

Відповідальний за підготовку облікових документів

Реєстратор

Керівник відділу УкрІНТЕІ, що є відповідальним за реєстрацію наукової діяльності



Юрченко Т.А.