МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Кваліфікаційна наукова
праця на правах рукопису

ДОЛГИХ Сергій Миколайович

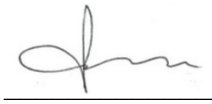
УДК 004.7:004.032.26

ДИСЕРТАЦІЯ
ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ РОЗПІЗНАВАННЯ МЕРЕЖЕВИХ ДАНИХ ІНТЕРНЕТ НА ОСНОВІ ГЕНЕРАТИВНИХ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ

Спеціальність 05.13.06 – «Інформаційні технології»

Подається на здобуття наукового ступеня кандидата технічних наук.

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело


Сергій ДОЛГИХ

Науковий керівник – доктор технічних наук, професор
Приставка Пилип Олександрович

Київ – 2023

АНОТАЦІЯ

Долгих С.М.: Інформаційна технологія розпізнавання мережевих даних Інтернет на основі генеративних нейромережевих моделей – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 «Інформаційні технології». – Національний авіаційний університет, МОН України, Київ, 2023.

Задачі застосування штучного інтелекту, тобто прийняття складних рішень автоматичними системами стають все більш актуальними в різних галузях технології та суспільства через наростаючі фактори кількості (обсягу, масовості), точності, стабільності, обмеження часу, та інших суттєвих обмежень на прийняття рішень фахівцем або навіть експертом.

У багатьох випадках, успіх застосування традиційних систем машинного інтелекту залежить від наявності значних наборів навчальних даних, які повинні задовольняти деяким суттєвим умовам:

– навчальні дані пов'язані з відомим результатом, який характеризує проблему, таким як заздалегідь відомі класи, категорії;

– високий рівень надійності в асоціації вхідних даних з відомим результатом;

– представництво, тобто високий рівень відповідності даних набору розподілу реальних даних; достатній розмір навчальних наборів; та інши.

У деяких областях, наприклад, аналіз та класифікація даних у мережах Інтернет існують серйозні перешкоди для виконання цих умов і, як наслідок, успішного навчання та застосування загальноприйнятих методів машинного інтелекту.

По-перше, наявність надійних і сучасних наборів навчальних даних може бути обмежена; по-друге, можуть існувати проблеми з представництвом наборів у конкретних галузях застосування, наприклад, зразків конкретних додатків Інтернет; і нарешті, як було зазначено у кількох опублікованих результатах, характеристичні параметри мережного трафіку мають тенденцію до зміни як у часі (тимчасовий дрифт), так і в залежності від мережі джерела (дрифт джерела), і з цієї причини використання доступних наборів може не забезпечити бажаного рівня точності машинних систем у задачах розпізнавання та класифікації звичайними методами.

Ці обмеження спонукають до пошуку методів навчання машинних систем, які меншою мірою залежать від підготовлених значних наборів навчальних даних і здатні використовувати систему характеристичних параметрів і концептів, що

відповідає конкретним наборам даних та середовищам використання. Таким чином, дисертаційна робота присвячена розв'язанню важливої науково-прикладної задачі: обґрунтування, розробки та експериментальної перевірки методів навчання систем машинного інтелекту здатних навчатися з мінімальними наборами даних Інтернет, при збереженні або покращенні характеристик успіху навчання, включаючи точність, стабільність та інші.

Наукова новизна одержаних результатів. На основі виконаних теоретичних і експериментальних досліджень розв’язано важливу науково-прикладну задачу розробки та експериментальної перевірки методів машинного інтелекту здатних навчатися з мінімальними наборами даних Інтернет, при збереженні або покращенні характеристик успіху навчання у порівнянні з традиційними методами розпізнавання даних Інтернет. При цьому отримано такі наукові результати:

1. *Удосконалено* концептуальну і математичну модель розподілів даних пакетів трафіку Інтернет, яка за рахунок використання неконтрольованих генеративних представлень, дозволяє підвищити ефективність виділення інформативних факторів даних.

2. *Отримали подальший розвиток* методи теорії генеративних представлень у напрямку розробки методів аналізу структури щільності розподілів даних пакетів трафіку Інтернет в представленнях генеративних моделей глибокого навчання, що за рахунок застосування оригінальної архітектури автоенкодера з різким стисненням розмірності шару кодування, забезпечило підвищення ефективності навчання за характеристиками зниження помилки та відтворення даних при неконтрольованому генеративному навчанні з використанням даних Інтернет та інших типів даних.

3. *Вперше* доведена теорема про категоризацію генеративних представлень, що на підставі методів варіаційного аналізу забезпечує теоретичне обґрунтування методів навчання на основі генеративної структури представлень.

4. *Вперше* розроблено методи навчання призначені для розпізнавання відомих класів даних пакетів трафіку Інтернет, які за рахунок розроблених концептуальної та математичної моделі, методів теорії генеративних представлень та теореми про категоризацію генеративних представлень, забезпечують: достатню точність розпізнавання; зменшення залежності від джерела отримання навчальних даних; зменшення обсягу навчальних даних в порівнянні з відомими методами контрольованого навчання.

5. *Вперше* розроблено методи навчання призначені для розпізнавання натуральних концептів даних пакетів трафіку Інтернет, які за рахунок визначення структури щільності генеративних представлень, дозволяють реалізовувати розпізнавання натуральних концептів даних Інтернет без використання маркованих даних, з точністю на рівні відомих методів контрольованого навчання.

6. *Вперше* визначено, формалізовано та виконано прототипну реалізацію інформаційної технології розпізнавання класів даних трафіку Інтернет, що за рахунок застосування запропонованих методів визначення структури щільності генеративних представлень, дозволяє автоматизувати процес навчання та використання запропонованих моделей і методів розпізнавання класів трафіку Інтернет.

Практичне значення одержаних результатів визначається тим, що запропоновані моделі і методи є науково-методологічною основою для розробки інформаційної технології обробки даних та навчання з наборами даних пакетів трафіку Інтернет мінімального обсягу на основі визначення структури щільності генеративних представлень.

1. Проведено порівняльний аналіз сучасних методів розпізнавання та класифікації даних пакетів трафіку Інтернет. Отримано порівняльні результати ефективності та недоліків розглянутих методів.

2. Отримано програмну реалізацію моделей генеративного навчання автоенкодера з різким стиском розмірності кодуючого шару кодування у програмному середовищі Python, дозволяє отримати інформативні представлення даних Інтернет суттєво зниженої розмірності.

3. Отримано програмне втілення масивів даних пакетів трафіка Інтернет та зображень у програмному середовищі Python.

4. Отримано програмну реалізацію методів визначення структури щільності генеративних представлень у програмному середовищі Python (методи кластеризації за щільністю, багатовимірних гістограм) дозволяє стабільне визначення структури щільності генеративних представлень даних Інтернет, з успішністю генеративного навчання та визначення структури щільності вище 80%.

5. Отримано програмну реалізацію методів розпізнавання відомих класів даних пакетів трафіку Інтернет з використанням структури щільності генеративних представлень у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання. Дозволяє стабільне навчання розпізнавання відомих класів даних Інтернет при зменшення обсягу навчальних даних, у 10–100 разів в порівнянні з відомими методами контрольованого навчання.

6. Отримано програмну реалізацію методу навчання розпізнавання натуральних концептів даних пакетів трафіку Інтернет з використанням структури щільності генеративних представлень, у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання та обробки даних. Дозволяє стабільне навчання розпізнавання натуральних типів (концептів) даних Інтернету без вимоги відомих даних навчання.

7. Отримано програмне прототипне виконання інформаційної технології розпізнавання даних пакетів трафіку Інтернет при наборах навчання мінімального обсягу з використанням структури щільності генеративних представлень у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання.

Результати дисертації використовуються у науково-дослідній діяльності НДЛ протидії кіберзагрозам авіаційної галузі.

Результати роботи підтверджуються ретельним аналізом теоретичних основ, досконалою експериментальною перевіркою та рецензованими публікаціями в українських та міжнародних наукових виданнях.

Особистий внесок здобувача. Результати що становлять основний зміст дисертації, отримані здобувачем самостійно. У роботах, виконаних у співавторстві, у дисертаційній роботі використовуються результати, що отримані особисто здобувачем: у роботі [2] – здобувачу належить концепція застосування методів та моделей неконтрольованого генеративного самонавчання в системах та процесах обробки даних аеронавігації, участь у написанні тексту, [9] – здобувачу належить концепція, обробка експериментальних даних, участь у написанні та форматуванні тексту, [12] – здобувачу належить обробка даних, дизайн та виконання моделей машинного навчання, обробка результатів, [14] – здобувачу належить обробка експериментальних даних, підготовка та форматування тексту, [17] – здобувачу належить концепція використання генеративних представлень; участь у написанні та форматуванні тексту.

Апробація результатів. Результати досліджень дисертаційної роботи доповідалися, обговорювалися та отримали позитивну оцінку на міжнародних та національних наукових та науково-практичних конференціях, включаючи: МНТК «Theory and Application of Soft Computing, Computing with Words and Perceptions (ICSCCW)» (Прага, 2019 р.), МНТК «Advanced Computer Information Technologies (ACIT)» (Дегендорф, 2020, 2021), МНТК «Informatics & Data-Driven Medicine (IDDM)» (Ваксйо, 2021), МНПК «ICT in Education, Research and Industry ICTERI-2021» (Херсон, 2021), МНТК «Soft Computing and Pattern Recognition (SoCPaR)» (MirLabs США, 2022); представлялися на наукових семінарах кафедр прикладної математики та аеронавігації Національного авіаційного університету (2019, 2020 р.); наукових семінарах Київського політехнічного інституту імені І. Сікорського (2019 р.), Ужгородського національного університету (2021 р.).

Публікації. Основні положення дисертації опубліковано у 21 науковій праці, в тому числі в – 1 розділі колективної монографії, 17 наукових статях (16 – у міжнародних рецензованих виданнях, що входять до бази даних Scopus, 1 – у вітчизняних фахових наукових журналах категорії Б), а також 3 матеріалах тез доповідей на конференціях.

Структура та обсяг дисертації. Дисертація складається із анотації, вступу, чотирьох розділів, висновків, додатків, списку використаних джерел і має 129 сторінок основного тексту, 29 рисунків, 20 таблиць. Список використаних джерел містить 110 найменувань. Загальний обсяг роботи складає 139 сторінок.

Ключові слова: інформаційна технологія, теорія неконтрольованого навчання, теорія генеративних представлень, кластеризація, штучні нейронні мережі, моделі глибокого навчання.

ABSTRACT

Dolgikh S. Information technology of learning and classification of Internet packet traffic data based on generative neural models. – Manuscript**.**

Dissertation for the degree of Candidate of Technical Sciences degree in specialization 05.13.06 "Information technologies" – National Aviation University, Kyiv, 2023.

Applications of methods of Artificial Intelligence, that is, making complex decisions by automatic systems are becoming more and more relevant in various fields of technology and society due to ever increasing influence of factors of the number (volume, mass, rate) of decisions; accuracy and stability; time constraints; and other significant factors and constraints on decision-making by a specialist or even an expert. In many cases, the success of the application of traditional machine intelligence systems depends on the availability of significant training data sets that must satisfy some essential conditions:

– training data is reliably associated with a known result that characterizes the problem, such as previously known classes, categories;

– representativity of the training sets, that is, a high degree of correspondence of the data in the trainig set to the distribution of real data;

– sufficient size of the training sets, including per known class or category of input data; and others.

In some areas, for example, the analysis and classification of data in the Internet networks, there can be serious obstacles to meeting these conditions and, as a result, successfully learning and applying commonly accepted methods of machine intelligence. First, the availability of reliable and up-to-date training datasets may be limited; secondly, there may be problems with the representation of sets in specific fields of application, for example, samples of specific Internet applications; and finally, as stated in several results, the characteristic parameters of network traffic tend to change both with time (temporal drift) and with respect to the source of traffic (source drift), that may lead to insufficient characteristics of learning success, including accuracy and stability in application of conventional methods and models of machine intelligence.

These limitations encourage the search for methods that are less dependent on prepared large training data sets and are able to use a system of characteristic parameters and concepts that are appropriate for specific data sets and usage environments. Thus, the dissertation work is devoted to the solution of an important scientific and applied problem: justification, development and experimental verification of methods for learning of machine intelligence systems capable of learning with minimal training sets of Internet data, while preserving or improving the characteristics of learning success, including accuracy, stability and others.

For these reasons methods of creating informative generative representations were closely examined and a theorem of categorization in generative representations proven under a number of identified conditions. The theorem establishes a theoretical foundation for definition of methods of learning known classes of Internet packet data with minimal sets of training samples and learning characteristic types (native concepts) in the data, based on the density structure in the latent distributions of data generative representations proposed and developed in the thesis.

Scientific novelty of the results. Based on the performed theoretical and experimental research, an essential scientific and applied problem of developing and experimentally testing machine intelligence systems capable of learning with minimal Internet data sets, while maintaining or improving the characteristics of learning success in comparison with traditional Internet data recognition systems and methods was solved. The following scientific results were obtained:

1. *Improved* the conceptual and mathematical model of distributions of Internet traffic packet data in the spaces of generative representations, which, through the use of unsupervised generative representations, allowed to increase the efficiency of extraction of informative factors.

2. *Further developed* methods of the theory of generative representations in the direction of developing methods for analyzing the density structure of distributions of Internet traffic packet data in informative representations of generative models, via use of an original architecture of an autoencoder with a sharp compression of the dimension

of the coding layer, which proved effective in unsupervised generative learning with the Internet data and others types of data.

3. *Originally* proved the theorem of categorization in generative representations, via application of methods of variational analysis, that provides theoretical ground for methods of learning based on the generative structure of informative representations.

4. *Originally* developed methods of learning for recognition of known classes of Internet traffic data which, based on the developed conceptual and mathematical model, methods of the theory of generative representations and the theorem on the categorization of generative representations, provide: sufficient recognition accuracy; reduced dependence on the source of training data; reduced the amount of training data compared to known methods of supervised learning.

5. *Originally* developed methods of learning natural concepts of Internet traffic packet which, based on determining the structure of the density of generative representations, allow recognition of natural concepts of Internet data without the use of labeled data, with an accuracy at the level of known methods of supervised learning.

6. *Originally* defined, formalized and implemented a prototype implementation of the information technology of recognition of classes of Internet traffic data of the class of deep learning architectures that by applying the proposed methods for determining the structure of the density of generative representations, allows to automate the process of training and using the proposed models and methods for recognizing classes of Internet traffic.

The proposed approach and methods significantly differ from existing methods of supervised machine learning and classification of Internet packet traffic data by minimal requirements for the size of training sets: down to several positive samples of classes, as well as by a significant reduction in dependence on the source network of training data.

The practical significance of the obtained results is determined by the fact that the proposed models and methods provide a scientific and methodological basis for developing information technology of data processing and learning with data sets of Internet traffic packets of minimal size based on determination of the density structure of generative representations.

Main practical results are:

1. Carried out a comparative analysis of modern methods of recognition and data classification of Internet traffic packets was carried out. Comparative results of effectiveness and shortcomings of the considered methods were obtained.

2. Obtained algorithmic and software implementation of the generative learning models of the type of deep autoencoder neural network with sharp compression of the dimension of the coding layer in the programming environment Python, allows to obtain informative representations of Internet data of significantly reduced dimensionality.

3. Developed programmatic implementation of datasets of Internet traffic packets and images in the programming environment Python.

4. Developed programmatic implementation of methods of determination of density structure in generative representations produced with models of deep learning in the programming environment Python, enabled stable determination of the density structure of generative representations of Internet data at the level of success above 80%.

5. Developed programmatic implementation of methods of recognition of known classes of Internet data, by using density structure of generative representations in the Python programming environment. Enabled stable learning of known classes Internet with in applications with general network data, with reduction of training sets by 10–100 times compared to the existing methods of supervised learning.

6. Developed programmatic implementation of methods of recognition of natural concepts of Internet traffic packet data using the density structure of generative representations in the Python programming environment. Allows stable training for recognition of natural types (concepts) of Internet data without requiring known training data.

7. Obtained a programmatic prototype implementation of the information technology of recognition of classes of Internet packet traffic data with minimal training sets using the density structure of generative representations in the Python programming environment using machine learning packages and libraries.

The results of the dissertation are used in the research activities of the Research Laboratory of Countering Cyber Threats to the Aviation Industry.

The proposed approach has a number of essential advantages compared to conventional supervised methods including: flexibility, in learning specific classes and concepts of interest without constraints of confident knowledge of the complete conceptual structure of the data; the ability to learn iteratively starting with minimal known sets; and, in a strong correspondence to the problem of the thesis, reduce to the minimum dependence of the learning success on the source of training data.

The results of the thesis are supported by a thorough theoretical analysis, extensive and detailed experimental verification and peer-reviewed publications in Ukrainian and international scientific literature.

Individual contribution of the author. The results that are related to the main content of the dissertation were obtained by the author individually. From the works that were published with coauthors, only the results obtained by the author individually were used in the dissertation: in [2] - the author owns the concept of applying methods and models of unsupervised generative self-learning in air navigation data processing systems and processes, participation in writing the text; in [9] - the author owns the concept, processing of experimental data, participation in writing and formatting the text; in [12] - the author owns data processing, design and implementation of machine learning models, processing of results; in [14]: the author owns the processing of experimental data, preparation and formatting of the text; in [17]: to the author belongs the concept of using generative representations; participation in writing and formatting the text.

Approbation of the results. The results of the research of the dissertation were presented, discussed and received positive evaluation at international and national scientific and scientific-practical conferences, including: international conferences «Theory and Application of Soft Computing, Computing with Words and Perceptions (ICSCCW)» (Prague Czech Republic, 2019), «Advanced Computer Information Technologies (ACIT)» (Deggendorf Germany, 2020, 2021), «Informatics & Data-Driven Medicine (IDDM)» (Vaxjo Sweden, 2021), «ICT in Education, Research and Industry ICTERI-2021» (Kherson, 2021), «Soft Computing and Pattern Recognition (SoCPaR)» (MirLabs USA, 2022); were presented at scientific seminars of the chairs of applied mathematics and aeronavigation, National Aviation University (2019, 2020 p.); scientific

seminar of Ihor Sikorsky Kyiv Politechinc University (2019 p.), Uzhgorod National University (2021 p.).

Publications. Main results of the thesis were published in 21 scientific works, that included: 1 chapter in collective monography, 17 scientific articles (16 articles – in international scientific peer-reviewed eidtions, 1 – in a national scientific registered edition) of which 16 included in the international scientific database Scopus, as well as 3 materials and theses of conference reports.

Structure and scope of the thesis. The dissertation consists of an abstract, introduction, four chapters, conclusions, appendices, and a list of references and includes 129 pages of the main text, 29 figures, and 20 tables. The list of references includes 110 items. The total volume of the work is 139 pages.

Keywords: information technology; unsupervised learning theory, theory of generative representations; clustering; artificial neural networks, deep learning.

ПУБЛІКАЦІЇ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Dolgikh S., «Spontaneous concept learning with deep autoencoder». *International Journal of Computational Intelligence Systems*, 12 (1), pp. 1–12, 2018. (*Scopus, Q1*)

2. Shmelyova T., Sterenharz A., Dolgikh S., «Artificial Intelligence in Aviation Industries: methodologies, education, applications and opportunities». IGI Global, 2019. (*розділ в колективній монографії*)

3. Dolgikh S. «Spontaneous categorization and self-learning with deep autoencoder models», *Advances in Aerospace Technology* (*Proceedings of National Aviation University*), 2019, 80 (3), p. 51-60. https://doi.org/10.18372/2306-1472.80.14274 (*категорія Б*)

4. Dolgikh S., «Low-dimensional representations in generative self-learning models». *CEUR Workshop Proceedings*, 2020, vol. 2718, pp. 239-245. (*Scopus*)

5. Dolgikh S., «Identifying explosive epidemiological cases with unsupervised machine learning». *CEUR Workshop Proceedings*, 2020, vol. 2753, pp. 1-10. (*Scopus*)

6. Dolgikh S., «Topology of conceptual representations in unsupervised generative models». *CEUR Workshop Proceedings*, 2021, vol. 2915, pp. 150-157. (*Scopus*)

7. Dolgikh S., «Sparsity constraint in unsupervised concept learning». *CEUR Workshop Proceedings*, 2021, vol. 2962, pp. 188-194. (*Scopus*)

8. Dolgikh S. «Generative conceptual representations and semantic communications». *International Journal of Computer Information Systems and Industrial Management Applications*, 14, 2022, pp. 239-248. (*Scopus*)

9. Prystavka P., Dolgikh S., Cholyshkina O., Kozachuk O., «Latent representations of terrain in aerial image classification», *CEUR Workshop Proceedings*, 2021, vol. 3013, pp. 86-95. (*Scopus*)

10. Dolgikh S., «Unsupervised clustering in epidemiological factor analysis». *The Open Bioinformatics Journal*, 14 (1), 2021, pp. 63-72. DOI: 10.2174/1875036202114010063. (*Scopus*)

11. Dolgikh, S. «Categorized representations and general learning». 10th International Conference on Theory and Application of Soft Computing, Computing with

Words and Perceptions ICSCCW 2019. Advances in Intelligent Systems and Computing, 2020, vol. 1095. Springer, Cham. (Online ISBN 978-3-030-35249-3) pp. 93-100 https://doi.org/10.1007/978-3-030-35249-3_11 (*Scopus*)

12. Seddigh N., Nandy B., Bennett D., Ren Y., Dolgikh S. et al., «A framework & system for classification of encrypted network traffic using Machine Learning». 15th International Conference on Network and Service Management (CNSM), 2019, p. 1-5, doi: 10.23919/CNSM46954.2019.9012662. (Electronic ISSN: 2165-963X) (*Scopus*)

13. Dolgikh S., «On unsupervised categorization in deep autoencoder models». Advances in Computer Science for Engineering and Education III. ICCSEEA-2020. Advances in Intelligent Systems and Computing, 2021, vol. 1247, Springer, Cham. (Online ISBN 978-3-030-55506-1), p.155-166. https://doi.org/10.1007/978-3-030-55506-1_23 (*Scopus*)

14. Prystavka P., Cholyshkina O., Dolgikh S., Karpenko D., «Automated object recognition system based on convolutional autoencoder». 10th International Conference on Advanced Computer Information Technologies (ACIT), 2020, p. 830-833. (Electronic ISBN:978-1-7281-6760-2) doi: 10.1109/ACIT49673.2020.9208945. (*Scopus*)

15. Dolgikh S., «Native concept frameworks in unsupervised generative learning». 11th International Conference on Advanced Computer Information Technologies (ACIT), 2021, pp. 748-752, (Electronic ISBN:978-1-6654-1854-6) doi: 10.1109/ACIT52158.2021.9548372. (*Scopus*)

16. Dolgikh S., «Synchronized conceptual representations in unsupervised generative learning». Proceedings of the 13th International Conference on Soft Computing and Pattern Recognition (SoCPaR 2021), Lecture Notes in Networks and Systems, vol. 417 Springer (Online ISBN 978-3-030-96302-6), 2022, p. 23-32. https://doi.org/10.1007/978-3-030-96302-6_2 (*Scopus*)

17. Prystavka P., Dolgikh S., Kozachuk O., «Terrain image recognition with unsupervised generative representations: the effect of anomalies», 12th International Conference on Advanced Computer Information Technologies (ACIT, Ruzomberok, Slovakia), 2022, p. 485-488. (*Scopus*)

18. Doligkh S., «Categorization in unsupervised generative self-learning systems».

International Journal of Modern Education and Computer Science, 13 (3), 2021, p. 68-78. DOI: 10.5815/ijmecs.2021.03.06 (*Scopus*)

19. Dolgikh S., «Unsupervised landscape, complex observations and association learning». 5th International Conference «Computational Intelligence» IntSol-2019, 2019, p.145-147.

20. Dolgikh S., «Parameter-less histogram-scale method of bandwidth estimation in density based clustering». Матеріали XV міжнародної конференції «Контроль і управління в складних системах (КУСС-2020)», м. Вінниця, 8-10 жовтня 2020 р. – Електрон. текст. дані. – Вінниця: ВНТУ, 2020. – Режим доступу: http://ir.lib.vntu.edu.ua//handle/123456789/30662.

21. Dolgikh S., «Characteristics of categorized latent representations in unsupervised generative Learning». 9th World Congress «Aviation in the XXI Century» National Aviation University (Київ, Україна) 2020. https://conference.nau.edu.ua/index.php/Congress/Congress2020/paper/viewFile/7602/6495.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

ГНМ – глибокі нейронні мережі

КР – контрастна розбіжність

МГВ – мережа глибокої віри

МГК – метод головних компонентів (англ. principal component analysis, PCA)

ОМБ – обмежена машина Больцмана

СГ – стохастичний градієнт

СПВ – середньоквадратична помилка відхилення

ШІ – штучний інтелект

DBSCAN – просторове кластеризування додатків із шумом на основі щільності (англ. density-based spatial clustering of applications with noise)

DNN – глибока нейронна мережа (англ. Deep Neural Network)

HTTP – протокол передачі гипертексту (англ. HyperText Transfer Protocol)

kNN – класифікатор найближчих сусідів (англ. k-nearest neighbor method)

ВСТУП

Обґрунтування вибору теми дослідження. Методи штучного інтелекту розвивалися прискореними темпами в останні роки, да досягли значного прогресу в різних областях: аналіз зображень та відео потоків [1, 2, 3, 4, 5]; медичної діагностики [6, 7]; лінгвістики [8]; моніторингу і аналізу стану комп'ютерних та телекомунікаційних мереж [9, 10] та кібербезпеки [11, 12, 13, 14]; прийняття рішень у критичних галузях, включаючи безпеку та військові додатки [15, 16, 17]; виявлення нових ознак (novelty detection) [18, 19]; ігри, такі як шахи, Ґо та комп'ютерні ігри [20, 21]; творчі програми, такі як створення творів мистецтва [22]; навчання з посиленням (reinforcement learning) [23] та багатьох інших [24].

У більшості згаданих випадків та використаннях успіх застосування систем машинного інтелекту залежить від наявності значних наборів навчальних даних, які повинні задовольняти деяким суттєвим умовам:

– навчальні дані пов'язані з відомим результатом, який характеризує проблему, таким як заздалегідь відомі класи, категорії;

– високий рівень надійності в асоціації вхідних даних з відомим результатом;

– представництво, тобто високий рівень відповідності даних набору розподілу реальних даних;

– достатній розмір навчальних наборів.

У деяких областях, наприклад, аналіз та класифікація даних у мережах Інтернет [25, 26, 27] існують серйозні перешкоди для виконання цих умов і, як наслідок, успішного застосування загальноприйнятих методів машинного інтелекту.

По-перше, наявність надійних і сучасних наборів навчальних даних може бути обмежена [10]; по-друге, можуть існувати проблеми з представництвом наборів у конкретних галузях застосування, наприклад, зразків конкретних додатків Інтернет; і нарешті, як було зазначено у кількох результатах [28, 29], характеристичні параметри мережного трафіку мають тенденцію до зміни як у часі (тимчасовий дрифт), так і в залежності від мережі джерела (дрифт джерела), і з цієї причини використання доступних наборів може не забезпечити бажаного рівня

точності машинних систем у задачах розпізнавання та класифікації звичайними методами.

Ці обмеження спонукають до пошуку методів навчання машинних систем, які меншою мірою залежать від заздалегідь підготовлених значних наборів навчальних даних і здатні використовувати систему характеристичних параметрів і концептів, що відповідає конкретним наборам даних та середовищам використання.

З іншого боку, значне число можливих характеризуючих ознак мережного трафіку, більше 200 [30] і значно більше у разі інших типів даних, таких як наприклад, дані зображень, робить застосування аналітичних методів аналізу даних не практичним. Таким чином, обмеження задачі відповідно підводять до пошуку підходів та методів навчання машинних систем, які можуть використовувати натуральну структуру та особливості даних, не вимагаючи значних наборів даних навчання.

Такий напрямок є відомим у теорії штучного інтелекту як *теорія неконтрольованого навчання* [31], в якій важливу роль відіграють моделі, методи та теорія нейронних мереж не в останню чергу завдяки їх властивості універсальної апроксимації [32]. Ця властивість забезпечує успіх застосування нейромережевих моделей з даними різних типів, складності та походження.

Класичні задачи розпізнавання. Завдання, що розглядаються в роботі, відносяться до наступних класичних завдань розпізнавання [33]:

1. Розробка апріорного словника ознак. Це завдання адресується обґрунтуванням, розробкою, виконанням та експериментальним випробуванням систем виділення інформативних параметрів на основі вихідного набору даних у процесі неконтрольованого навчання.

2. Опис класів апріорного алфавіту класів. Модифікація постановки цього завдання в роботі полягає в тому, що при підході неконтрольованого навчання апріорні набори класів можуть не бути відомими заздалегідь, а повинні визначатися навчальною системою в процесі неконтрольованого навчання та обробки представлення вихідних даних.

3. Необхідною складовою процесу виділення неконтрольованих або природних класів у цьому підході є "переведення" ознак класів у просторі подання в ознаки простору вхідних, або спостережуваних даних, тобто розбиття апріорного простору ознак на області відповідні визначеним у процесі навчання класам.

4. Розробка методів навчання на неконтрольованому ландшафті щільності, що виникає в процесі неконтрольованого навчання відноситься до галузі класичної задачі розпізнавання алгоритмів, який полягає у виборі, що забезпечують ефективний розподіл розпізнаваного зразка до встановлених класів.

5. Нарешті, при апробуванні розроблених методів порушується питання вибору методів оцінки ефективності запропонованих підходів, що стосується класичної задачі методів оцінки ефективності розповсюдження.

Таким чином, загальний напрям роботи, завдання розв'язувані як в теоретичній так і в експериментальній частині близько пов'язані з класичними завданнями розпізнавання які досліджувалися в літературі.

Методи неконтрольованого навчання. У кількох опублікованих рецензованих наукових виданнях у галузі систем машинного навчання та дослідженнях по структурі генеративних представлень, створених моделями неконтрольованого генеративного навчання, у тому числі на основі глибоких нейронних мереж, було встановлено тісний зв'язок між інформаційними структурами у просторах представлень що виникають у процесі неконтрольованого навчання концепціями або категоріями у вхідних даних [34, 35, 36]. Як обговорювалося вище, такий зв'язок може використовуватися для підвищення ефективності навчання в задачах та галузях з дефіцитом надійних навчальних даних.

Методи неконтрольованого машинного навчання мають давню історію. Ці методи включають ряд підходів, таких як: обмежені машини Больцмана і мережі глибокого переконання [37, 38]; методи кластеризації, у тому числі неконтрольованої [39, 40]; мережі, що само-організуються (карти Кохонена [41, 42]); моделі глибоких нейронних мереж із кодуванням та генерацією, включаючи різні типи архітектур нейронних мереж автоенкодера [43, 44, 45]; генеративні моделі протистояння (generative adversarial networks, GAN) [46, 47] тощо.

Результати ряду робіт пов'язують представлення вхідних даних, створені із застосуванням неконтрольованих методів навчання, які не вимагають будь-яких відомих даних або дуже невеликих їх обсягів, з покращеною ефективністю класифікації на основі неконтрольованих представлень [48].

У ряді результатів такий взаємозв'язок виявлявся як структури розподілу даних у просторі представлень, наприклад, структури щільності, які виникають в результаті неконтрольованого навчання до застосування будь-яких відомих даних, що дозволяє визначити, виміряти і в перспективі, використовувати такі інформаційні структури для створення методів і моделей з більш ефективним та спонтанним навчанням.

Правильний вибір підходу та методів неконтрольованого навчання та створення інформативних представлень вхідних даних требує детального огляду та аналізу сучасних методів, моделей та стану досліджень у галузі неконтрольованого машинного навчання.

Основна гіпотеза роботи. Основна гіпотеза роботи полягає в тому, що генеративна структура, наприклад, структура щільності даних інформативних представлень даних пакетів трафіку Інтернет, створених генеративними нейромережевими моделями, може бути використана для підвищення ефективності методів навчання щодо: 1) необхідних відомих даних навчання, тобто навчання з мінімальними наборами відомих даних та 2) зниження залежності успіху навчання від джерела навчального набору.

Сучасни результати опубліковані раніше в науковій літературі [35, 36], вказують на можливість кореляції структури генеративних представлень з характерними типами, концептами у розподілі вхідних даних, що може дозволити використовувати такі структури, включаючи структури щільності генеративних представлень, для успішного навчання машинних систем із сильно обмеженими наборами. відомих даних.

З огляду на вищезазначене, розробка і дослідження нових ефективних методів забезпечення стабільності та ефективності навчання систем машинного інтелекту з мінімальними даними істини з використанням природної структури даних

Інтернет є актуальною науково-практичною задачею, що має теоретичне і практичне значення.

Зв'язок роботи з науковими програмами, планами, темами. Результати дисертаційних досліджень реалізовані в рамках держбюджетної теми №247-ДБ19 «Розроблення та виготовлення програмно-апаратних засобів цільового навантаження для повітряного спостереження та альтернативної навігації літального апарату» (№ держреєстрації: 0119U100553), № 421-ДБ22 «Інтелектуалізована система захищеного передавання пакетних даних на базі розвідувально-пошукового безпілотного літального апарату» (№ держреєстрації: 0122U002361) і виконання тематичних наукових планів досліджень кафедри прикладної математики факультету кібербезпеки, комп’ютерної та програмної інженерії Національного авіаційного університету 2021 р.

Мета і зада чі дослідження. Метою дисертаційної роботи є розробка моделей та методів неконтрольованого глибокого навчання процесу розпізнавання даних пакетів трафіку Інтернет на основі використання структури щільності генеративних представлень для розв'язання завдання впевненого розпізнавання категорій даних Інтернет у областях та додатках, де значні масиви навчальних даних не доступні, або не можуть використовуватись внаслідок значної залежності успіху навчання традиційними методами від джерела даних, на основі інформативної структури представлень моделей генеративного навчання.

Досягнення наведеної мети в рамках цієї роботи можна поділити на розв'язання нижченаведених задач:

1. Провести огляд та оцінку сучасних моделей та методів неконтрольованого навчання розпізнаванню даних пакетів трафіка Інтернет.

2. Створити теоретичну математичну модель розподілу даних пакетів трафіка Інтернет, включаючи розподіл щільності в просторах генеративних представлень. Провести теоретичне обґрунтування гіпотези про виникнення структурованих представлень моделей генеративного навчання.

3. На основі теоретичної моделі та відомих моделей генеративного навчання розробити та імплементувати моделі неконтрольованого генеративного навчання з

даними Інтернет, зокрема нейромережевих кодуючих архітектур глибокого навчання.

4. Зібрати, обробити та підготувати для використання масиви даних пакетів трафіка Інтернет та інших типів для навчання генеративних моделей.

5. Провести оцінку та аналіз характеристик розподілу даних пакетів трафіка Інтернет у просторах інформативних генеративних представлень.

6. Розробити, імплементувати та верифікувати методи навчання розпізнавання відомих класів даних пакетів трафіка Інтернет з наборами навчальних даних мінімального обсягу (до кількох зразків) на основі генеративних представлень.

7. Розробити, імплементувати та верифікувати повністю неконтрольовані методи навчання розпізнавання натуральних концептів даних пакетів трафіка Інтернет без вимог даних навчання відомих класів на основі генеративних представлень.

8. Сформулювати, формалізувати та виконати концептну реалізацію інформаційної технології обробки даних та навчання розпізнаванню відомих класів даних пакетів трафіку Інтернет з наборамі навчальних даних мінімального обсягу (до кількох зразків) на основі генеративної структури представлень.

Об'єктом дослідження є процеси створення та застосування інформаційних представлень даних Інтернет, створені моделями неконтрольованого генеративного навчання.

Предметом дослідження є моделі, методи та засоби створення та аналізу неконтрольованих генеративних представлень даних Інтернет та інших типів, включаючи: нейромережеві моделі неконтрольованого генеративного навчання; методи аналізу розподілу даних, включаючи багатовимірні розподіли; методи кластеризації.

Методами дослідження є теорія інформації, статистичні методи, теорія неконтрольованого навчання та представлень – для розробки методів створення інформативних представлень даних Інтернет із значно зниженою розмірністю; моделі та методи неконтрольованого навчання, включаючи генеративні моделі нейронних мереж автоенкодера – для створення інформативних представлень

даних зі значно зниженою розмірністю; методи кластеризації, включаючи кластеризацію за щільністю – для аналізу структури неконтрольованих генеративних представлень; методи статистичного аналізу – для обробки результатів експериментів та верифікації ефективності розроблених методів.

Наукова новизна результатів дисертаційної роботи:

1. *Удосконалено* концептуальну і математичну модель розподілів даних пакетів трафіку Інтернет, яка за рахунок використання неконтрольованих генеративних представлень, дозволяє підвищити ефективність виділення інформативних факторів даних.

2. *Отримали подальший розвиток* методи теорії генеративних представлень у напрямку розробки методів аналізу структури щільності розподілів даних пакетів трафіку Інтернет в представленнях генеративних моделей глибокого навчання, що за рахунок застосування оригінальної архітектури автоенкодера з різким стисненням розмірності шару кодування, забезпечило підвищення ефективності навчання за характеристиками зниження помилки та відтворення даних при неконтрольованому генеративному навчанні з використанням даних Інтернет та інших типів даних

3. *Вперше доведена теорема* про спонтанну категоризацію генеративних представлень, що на підставі методів варіаційного аналізу забезпечує теоретичне обґрунтування методів навчання на основі генеративної структури представлень.

4. *Вперше* розроблено методи навчання призначені для розпізнавання відомих класів даних пакетів трафіку Інтернет, які за рахунок розроблених концептуальної та математичної моделі, методів теорії генеративних представлень та теореми про категоризацію генеративних представлень, забезпечують: достатню точність розпізнавання; зменшення залежності від джерела отримання навчальних даних; зменшення обсягу навчальних даних в порівнянні з відомими методами контрольованого навчання.

5. *Вперше* розроблено методи навчання призначені для розпізнавання натуральних концептів даних пакетів трафіку Інтернет, які за рахунок визначення структури щільності генеративних представлень, дозволяють реалізовувати

розпізнавання натуральних концептів даних Інтернет без використання маркованих даних, з точністю на рівні відомих методів контрольованого навчання.

6. *Вперше* визначено, формалізовано та виконано прототипну реалізацію інформаційної технології розпізнавання класів даних трафіку Інтернет, що за рахунок застосування запропонованих методів визначення структури щільності генеративних представлень, дозволяє автоматизувати процес навчання та використання запропонованих моделей і методів розпізнавання класів трафіку Інтернет.

Практичне значення отриманих результатів. У роботі розроблено практичні технології обробки даних Інтернет та розпізнавання відомих класів даних пакетів трафіку Інтернет, таких як додатки Інтернет з мінімальними наборами навчальних даних та натуральних концептів даних, на основі структури щільності генеративних представлень.

1. Проведено порівняльний аналіз сучасних методів розпізнавання та класифікації даних пакетів трафіку Інтернет. Отримано порівняльні результати ефективності та недоліків розглянутих методів.

2. Отримано програмну реалізацію моделей генеративного навчання автоенкодера з різким стиском розмірності кодуючого шару кодування у програмному середовищі Python, дозволяє отримати інформативні представлення даних Інтернет суттєво зниженої розмірності.

3. Отримано програмне втілення масивів даних пакетів трафіка Інтернет та зображень у програмному середовищі Python.

4. Отримано програмну реалізацію методів визначення структури щільності генеративних представлень у програмному середовищі Python (методи кластеризації за щільністю, багатовимірних гістограм) дозволяє стабільне визначення структури щільності генеративних представлень даних Інтернет, з успішністю генеративного навчання та визначення структури щільності вище 80%.

5. Отримано програмну реалізацію методів розпізнавання відомих класів даних пакетів трафіку Інтернет з використанням структури щільності генеративних представлень у програмному середовищі Python з використанням пакетів та

бібліотек машинного навчання. Дозволяє стабільне навчання розпізнавання відомих класів даних Інтернет при зменшення обсягу навчальних даних, у 10–100 разів в порівнянні з відомими методами контрольованого навчання.

6. Отримано програмну реалізацію методу навчання розпізнавання натуральних концептів даних пакетів трафіку Інтернет з використанням структури щільності генеративних представлень, у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання та обробки даних. Дозволяє стабільне навчання розпізнавання натуральних типів (концептів) даних Інтернету без вимоги відомих даних навчання.

7. Отримано програмне прототипне виконання інформаційної технології розпізнавання даних пакетів трафіку Інтернет при наборах навчання мінімального обсягу з використанням структури щільності генеративних представлень у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання.

Результати дисертації використовуються у науково-дослідній діяльності НДЛ протидії кіберзагрозам авіаційної галузі.

Особистий внесок здобувача. Результати що становлять основний зміст дисертації, отримані здобувачем самостійно. У роботах, виконаних у співавторстві, у дисертаційній роботі використовуються результати, що отримані особисто здобувачем: у роботі [2] – здобувачу належить концепція застосування методів та моделей неконтрольованого генеративного навчання в системах та процесах обробки даних аеронавігації, участь у написанні тексту, [9] – здобувачу належить концепція, обробка експериментальних даних, участь у написанні та форматуванні тексту, [12] – здобувачу належить обробка даних, дизайн та виконання моделей машинного навчання, обробка результатів, [14] – здобувачу належить обробка експериментальних даних, підготовка та форматування тексту, [17] – здобувачу належить концепція використання генеративних представлень; участь у написанні та форматуванні тексту.

Апробація результатів. Результати досліджень дисертаційної роботи доповідалися, обговорювалися та отримали позитивну оцінку на наукових та

науково-практичних конференціях: МНТК «Theory and Application of Soft Computing, Computing with Words and Perceptions (ICSCCW)» (Прага, 2019 р.), МНПК «Advances in Computer Science for Engineering and Education (ICCSEEA)» (Київ, 2020 р.), МНТК «Advanced Computer Information Technologies (ACIT)» (Дегендорф, 2020, 2021), МНТК «Information Technologies - Applications and Theory (ITAT)» (Хелпа, 2020, 2021), МНТК Informatics & Data-Driven Medicine (IDDM)» (Ваксйо, 2021), МНПК «Information Society and University Studies (IVUS)» (Каунас, 2021), МНПК «ICT in Education, Research and Industry ICTERI-2021» (Херсон, 2021), МНТК «Soft Computing and Pattern Recognition (SoCPaR)» (MirLabs США, 2022); представлялися на наукових семінарах кафедр прикладної математики та аеронавігації Національного авіаційного університету (2019, 2020 р.); наукових семінарах Київського політехнічного інституту імені І. Сікорського (2019 р.), Ужгородського національного університету (2021 р.).

Публікації. Основні положення дисертації опубліковано у 21 науковій праці, в тому числі в – 1 розділі колективної монографії, 17 наукових статях (16 – у міжнародних рецензованих виданнях, що входять до бази даних Scopus, 1 – у вітчизняних фахових наукових журналах категорії Б), а також 3 матеріалах тез доповідей на конференціях.

Структура та обсяг дисертації. Дисертація складається із анотації, вступу, чотирьох розділів, висновків, додатків, списку використаних джерел і має 129 сторінок основного тексту, 29 рисунків, 20 таблиць. Список використаних джерел містить 110 найменувань. Загальний обсяг роботи складає 139 сторінок.

РОЗДІЛ 1.

АНАЛІЗ СУЧАСНИХ МЕТОДІВ ТА МОДЕЛЕЙ НЕКОНТРОЛЬОВАНОГО НАВЧАННЯ ТА РОЗПОДІЛІВ ДАНИХ У БАГАТОВИМІРНИХ ПРОСТОРАХ ПРЕДСТАВЛЕНЬ

1.1. Визначення основних понять

У даному розділі дисертаційної роботи наведені визначення ключових понять, які часто використовуються як у теорії систем Штучного Інтелекту (ШІ), так і в даній роботі. Такими є підходи до створення принципів та методів навчання систем ШІ.

Системи контрольованого навчання. Системі даються приклади вхідних даних та бажаних результатів, наданих "учителем" або навчальним набором даних. Мета навчання полягає в тому, щоб вивчити загальні правила, які дозволяють асоціювати вхідні дані з бажаним або "правильним" результатом на виході [49]. Після фази контрольованого навчання система здатна успішно передбачати або класифікувати вхідні дані без зазначених результатів (міток істини), оскільки характер (тобто суттєві параметри розподілу) вхідних даних залишається постійним.

Системи неконтрольованого навчання. Система не має доступу до маркованих даних, позначених істинним результатом і повинна самостійно встановити та визначити структуру вхідних даних, таку як, наприклад, суттєві характеристики розподілів. Самостійне навчання може бути самоціллю (виявлення прихованих закономірностей у даних) або засобом досягнення мети (вивчення особливостей розподілу даних) [31].

Одним із широковідомих підходів до створення систем неконтрольованого навчання є генеративне навчання, засноване на принципі відтворення вхідних даних, при якому система спочатку створює внутрішній прообраз (представлення) вхідних даних і далі вчиться відтворювати з нього вхідний розподіл у процесі неконтрольованого навчання зі стимулом зниження помилки відтворення. Іноді такі системи називаються системами *із самоконтрольованим навчанням* (self-supervised learning) [50].

Системи напівконтрольованого навчання (semi-supervised learning) – це системи та методи, які поєднують незначну кількість маркованих даних з великою кількістю немаркованих даних у процесі навчання [51]. При цьому марковані дані можуть використовуватися, щоб позначати характерні структури або категорії у вхідних даних, або представленнях створених на їх основі.

Системи з посиленням навчання (reinforcement learning). Система взаємодіє з динамічним середовищем, в якому вона повинна виконувати певну мету (наприклад, керувати транспортним засобом або грати в гру проти суперника) [52]. Середовище забезпечує зворотний зв'язок зазвичай у діапазоні винагорода – штраф, під час руху проблемним простором система вчиться вибирати оптимальний маршрут у ньому зі стимулом максимізації виграшу. Такі системи можна також відносити до класу напівконтрольованих систем, оскільки навчальний вплив існує не у формі заданих даних, а правил і умов взаємодії та поведінки у просторі задачі.

До цього типу належать багато самонавчальних ігрових систем (шахи, Ґо, комп'ютерні ігри та інші).

Навчання представлень (Representation Learning). При навчанні представлень система створює модель вхідних даних навчання на основі інформативних параметрів (features), що визначаються в процесі навчання [53]. Представлення можуть використовуватися у всіх типах навчальних систем і часто дозволяють підвищити ефективність навчання.

Особливим типом навчання представлень є неконтрольовані представлення, що створюються в процесі неконтрольованого навчання без маркованих наборів відомих класів.

Оскільки напрям роботи тісно пов'язаний з системами неконтрольованого навчання без використання або з мінімальним використанням даних істини, у наведених нижче розділах наводиться опис систем і методів неконтрольованого навчання.

1.2. Огляд сучасних методів, підходів та моделей неконтрольованого навчання

У цьому розділі міститься детальний опис сучасних методів із додатком до теоретичних засад та експериментів, проведених у роботі. Область застосування та таксонометрія методів неконтрольованого навчання показана на рис. 1.1.

Рис. 1.1. Підходи та методи неконтрольованого навчання

1.2.1. Методи аналізу розподілу даних

Метод головних компонентів (МГК). У роботі перетворення головних компонентів [54] використовувалося для визначення характеристик вхідного розподілу, таких як число лінійних компонентів вхідних параметрів з найбільшою варіацією. Ці характеристики можуть бути істотними щодо оптимальної розмірності простору представлення моделей генеративного навчання і з цієї причини важливі для побудови ефективних моделей.

Для визначення характеристик головних компонентів, набір загального розподілу розміром від 0.1–0.3 повного неконтрольованого набору даних оброблявся перетворенням виділення головних компонентів, що дозволило визначити кількість лінійних компонентів, які вносять суттєвий внесок у варіаційну

складову розподілу вхідних даних. Дані аналізу головних компонентів для наборів Інтернет та зображень наведено в розділі 3.

Метод багатовимірних гістограм. Метод багатовимірних гістограм [55] може бути корисним при аналізі неконтрольованих розподілів у багатовимірних просторах представлень. Визначення параметрів розподілів та їх залежності від масштабу гістограми дає можливість зробити істотні висновки про характеристики та характер розподілів, їх розміри, щільність та інші істотні фактори.

Важлива перевага методу гістограм полягає в тому, що він не вимагає попередніх знань про розподіл, або припущення щодо їх характеру. Це дозволяє застосовувати його у випадках розподілів складних реальних даних, коли зробити обґрунтовані представлення про характер розподілу параметрів апріорі не завжди можливо. За допомогою гістограмування можна визначити суттєві параметри розподілів як загального немаркованого зразка, так і наборів відомих класів, таких як:

– розміри області розподілу та її статистичні параметри характеристики, такі як моменти розподілу;

– параметри розподілу щільності загального розподілу даних та розподілів відомих класів;

– характеристики залежності розподілу щільності від параметрів масштабування гістограми та інші.

Використання багатовимірної візуалізації разом із методами гістограм дозволяє помітити цікаві особливості розподілів у просторах представлень, створених моделями неконтрольованого навчання. На рис. 1.2 представлений приклад розподілу щільності представлення створеного моделлю глибокого автоенкодера з даними Інтернет.

Методи багатовимірного гістограмування використовувалися значною мірою під час експериментального аналізу та визначення характеристичних параметрів розподілів, створених моделями неконтрольованого навчання для перевірки теоретичних результатів роботи.

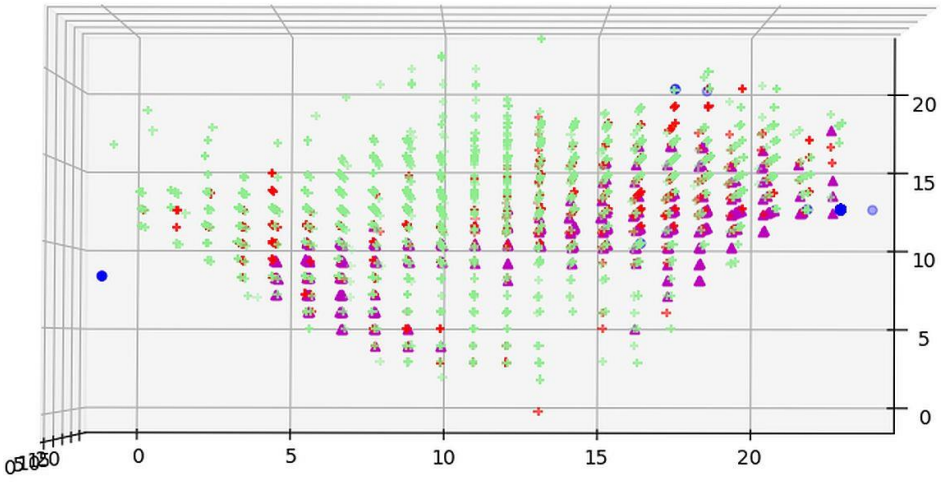


Рис. 1.2. Гістограма розподілу щільності в генеративному просторі представлення даних Інтернет

Методи ядерної щільності. Методи ядерної щільності широко використовуються в статистиці для моделювання функції або форми невідомого розподілу даних. Вони тісно пов'язані з методами гістограм, долаючи деякі їхні проблеми, такі як сильна залежність від параметрів масштабу.

Моделювання невідомого розподілу *f(X)* ядерним естіматором $F_h(x)$ визначається за формулою Парцена-Розенблату [56]:

$$F_h = \frac{1}{n}\sum_{i=0}^{n} K_h\,(x - x_i)$$

де h – параметр ширини смуги, є зовнішнім параметром методу; $\mathrm{K}_h(\mathrm{X}) = 1/h\ \mathrm{K}(X/h)$. Вибір значення параметра ширини смуги істотний для оптимального моделювання розподілу і може вимагати подальшого аналізу для наборів даних з невідомим розподілом.

Застосування методів ядерної щільності у багатьох випадках допомагає встановити закон, що лежить в основі розподілу емпіричних даних, дозволяючи згладити емпіричну функцію щільності розподілу та оцінити її суттєві параметри.

Існує значна кількість широко застосовуваних функцій ядра, таких як Гауссовська, Епанечнікова, Flat та інші [57, 58].

Для оцінки параметра ширини смуги використовувалися наступні методи:

1. Емпірична формула наближення нормального розподілу [59]:

$$h = 0{,}9\ min\ (\sigma, IQR/1{,}34) \times n^{-\frac{1}{5}},$$

де σ, IQR – статистичні параметри розподілу; n – чисельність набору даних.

2. Методи оцінки ширини смуги, реалізовані у програмному пакеті sklearn-kit.

3. Одночасне використання методу гістограм та оцінки ядерної щільності.

Одним з поширених застосувань методів ядерної щільності є їх використання в методах кластеризації за щільністю маркованих даних, що не вимагають і дозволяють оцінити розподіл щільності в довільних масивах даних з невідомим заздалегідь законом розподілу. Адаптивні методи оцінки ядерної щільності з варіацією параметра ширини смуги є цікавим можливим напрямом для поліпшення характеристик безперервності та згладжування розподілів представлень створених моделями неконтрольованого навчання.

Комбінація методів гістограм та ядерної щільності дозволяє подолати обмеження цих методів, таких як залежність від зовнішніх параметрів масштабу кошика (метод гістограм) або ширини смуги (методи ядерної щільності) та встановити характер розподілу та його параметри з необхідною точністю.

1.2.2. Методи кластеризації

Метою кластеризації є визначення внутрішніх груп чи концепцій, категорій, кластерів у розподілі даних. Різні методи кластеризації ґрунтуються на критерії близькості між зразками даних, які можуть суттєво відрізнятися між методами [60, 61]. Було показано, що не існує абсолютного критерію, який би забезпечував кращий результат незалежно від кінцевої мети кластеризації, тому кластерні методи відрізняються за типом та характером критерію близькості [62]. Існує багато типів методів кластеризації, включаючи:

– K-means, Fuzzy K-means;
– Hierarchical clustering;
– Mixture of Gaussians та інші.

Звичайні методи кластеризації, такі як K-Means не можуть застосовуватися безпосередньо з моделями неконтрольованого навчання, оскільки вони вимагають

апріорного знання про структуру інформації, яке може не бути відомо заздалегідь у даних загального характеру, наприклад, кількість характерних структур або кластерів, яке саме є невідомим параметром розподілу та вимагає визначення. З цієї причини в роботі використовувалися методи неконтрольованої кластеризації за щільністю, які не залежать від знання структури розподілу даних.

Методи кластеризації за щільністю. Кластеризація на основі щільності відноситься до неконтрольованих методів навчання, які ідентифікують відмінні групи/кластери в даних, ґрунтуючись на ідеї, що кластер у просторі даних є безперервною областю з високою густиною точок, відокремленою від інших безперервними областями більш низької щільності. Перевага методів неконтрольованої кластеризації полягає в тому, що їх можна використовувати для визначення структури щільності розподілу даних без попереднього знання про характер розподілу та концепції у вихідних даних, що має принципове значення для аналізу представлень, створених у процесі неконтрольованого навчання.

Розповсюдженими методами кластеризації за щільністю, які не спираються на попереднє знання про характер даних, такі як очікуване число кластерів (K-Means), закон розподілу (змішаний гаусівський) або функцію близькості (ядерні методи) є DBSCAN і MeanShift.

Метод DBSCAN був запропонований в [63] і ґрунтується на обчисленні ядер і пов'язаних компонентів обчислюваних на підставі графи околиці. Параметрами методу є ε (околиця пошуку) та minPts (мінімальна кількість точок, що визначає кластер). Характеристики методу:

– не вимагає інформації про кількість кластерів (на відміну від K-Means);

– здатний визначати кластери довільної форми. Наприклад, DBSCAN здатний знайти кластер, повністю оточений (але не пов'язаний) іншим кластером;

– відрізняє дані фону (шум) та стійкий до викидів (outliers);

– вимагає лише двох зовнішніх параметрів і в загальному нечутливий до порядку даних у масиві;

– може ефективно використовуватися з масивами у базах даних.

З іншого боку, результат застосування методу суттєво залежить від вибору параметрів околиці, визначення яких може бути неочевидним та вимагати додаткового аналізу під час використання невідомих даних.

Метод MeanShift [64] заснований на процесі визначення максимумів функції щільності розподілу довільних даних, представлених дискретним набором проб. Метод має два зовнішні параметри: функція ядра та ширина смуги. Однак одне значення функції ядра, "плоске ядро" (flat kernel) не передбачає обмежень на характер розподілу і може використовуватися з довільними даними.

Характеристики методу:

– не включає припущень про характер розподілу, не прив'язаний до певної галузі використання і може застосовуватися для аналізу довільних даних;

– не включає припущень про форму та геометричні розміри кластерів даних;

– застосування методу визначається одним зовнішнім параметром – шириною смуги (розмір вікна);

– параметр ширини смуги має ясну фізичну інтерпретацію, на відміну параметрів деяких інших методів кластеризації.

Результат застосування методу може істотно залежати від вибору значення параметра ширини смуги h, що може не бути тривіальним для довільних даних і вимагати попереднього аналізу. Як зазначалося вище, існує кілька практичних та емпіричних методів визначення значення ширини смуги.

У роботі використовувалась реалізація методу MeanShift у програмному пакеті *sklearn-kit*. Застосування методу дозволило стійко визначити структуру розподілу щільності у просторі представлення моделей і розробити на основі визначеної структури (ландшафту) щільності ефективні методи ітераційного навчання з мінімальними наборами даних. На діаграмі (рис. 1.3) представлено загальний розподіл даних пакетів трафіку Інтернет у просторі генеративного представлення з визначеними методом MeanShift кластерами щільності.

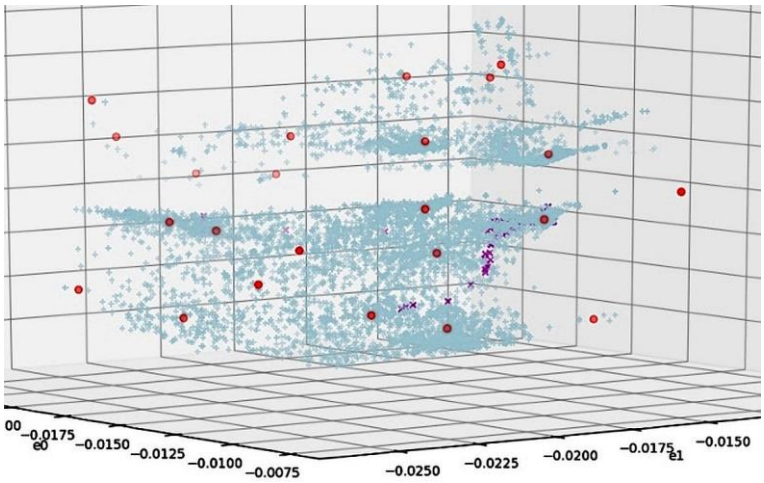


Рис. 1.3. Розподіл даних пакетів трафіку Інтернет у просторі представлення з визначеними кластерами щільності (метод MeanShift)

Також можна відзначити методи непараметричного визначення функції щільності, розроблені останнім часом [65].

1.2.3. Глибокі нейронні мережі

Застосування глибоких нейронних мереж (ГНМ) досягло значних успіхів у галузі контрольованого навчання даних у різних галузях та різних типів. Наприклад, були зроблені значні просування у сфері класифікації зображень, зокрема розпізнавання осіб [66]. Потужними інструментами зі значними здібностями до узагальнення їх роблять здатність до ефективного навчання масивними наборами даних, включаючи складні типи, такі як зображення або потоки медіа даних.

Посилаючись лише на деякі з багатьох досягнень в обробці даних, активації, оптимізації та інших етапів проектування та навчання глибоких нейронних мереж з дуже великими наборами даних та діапазонами класифікаційних категорій, Д. Кінгма та співавтори [67] розробили розширене навчання стохастичної оптимізації з адаптивною швидкістю навчання, що дозволило поліпшити точність і час навчання. Дж. Чжан та співавтори [68] досягли успіху у навчанні глибоких нейронних мереж до тисячі шарів із значно покращеною точністю розпізнавання зображень. Стохастичні алгоритми глибинного навчання тепер дозволяють навчати глибокі нейронні мережі з більш ніж 1200 шарами, значно скорочуючи час

навчання та досягаючи виняткової точності в класифікації зображень, порівнянної та переважаючої здатності середньої людини.

В даний час глибокі нейронні мережі стали одним з найважливіших інструментів досліджень штучного інтелекту, як у контрольованому так і неконтрольованому навчанні не в останню чергу завдяки їх природним перевагам, зокрема, досягнута за минулі десятиліття можливість ефективно навчати мережі практично необмеженої глибини, що відповідає даним найскладніших типів (наприклад, недавні моделі глибокого нейронного автоенкодера, що використовувалися для неконтрольованого навчання на масиві зображень, мали більше ста мільярдів параметрів навчання).

До того ж наейронні мережі мають властивість універсальної апроксимації [32, 69], тобто теоретично здатні моделювати будь-який розподіл із заданою точністю, що поділяє їх природним інструментом аналізу та класифікації складних даних.

На рис. 1.4 представлена діаграма глибокої нейронної мережі контрольованого навчання AlexNet [70] ImageNet-2012, яка виграла змагання за класифікацією зображень набору з 22 тисяч категорій.

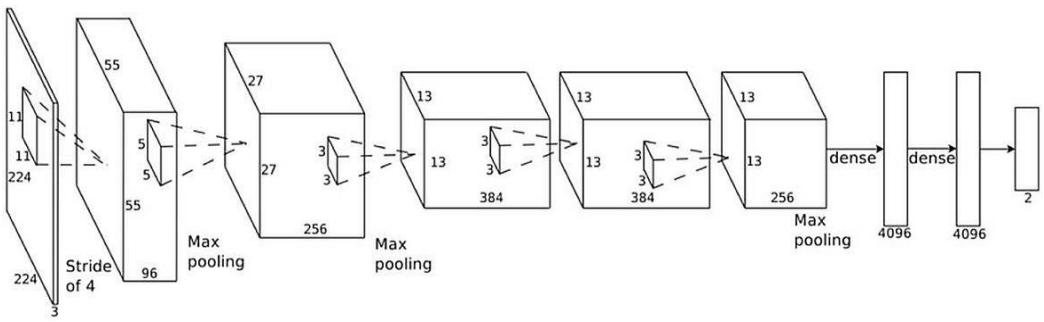


Рис. 1.4. Модель глибокої нейронної мережі AlexNet для обробки та класифікації зображень

Для навчання ГНМ часто використовуються методи зниження помилки за стохастичним градієнтом (СГЗ) [71]. Ці методи застосовуються до моделей нейронних мереж контрольованого та неконтрольованого навчання.

Розроблені та використовувані в роботі моделі глибокого автоенкодера так само, як і багато моделей з цитованих результатів, засновані на методах побудови, навчання та аналізу глибоких нейронних мереж. Зокрема, при навчанні моделей глибокого автоенкодера зі стисненням розмірності центрального шару використовуваних у роботі використовувався алгоритм зниження за стохастичним градієнтом на основі помилки відтворення навчального немаркованого набору вихідних даних розвинений у дослідженнях глибоких нейронних мереж.

1.2.4. Моделі неконтрольованого навчання

Класичні моделі та методи. Моделі обмежених машин Больцмана (ОМБ) та мереж глибокої віри (МГВ) [37, 38] були введені у практику машинного навчання наприкінці 20-го століття і є потужними інструментами виділення інформативних параметрів та створення представлень. Принцип навчання моделі заснований на зменшенні помилки відтворення вихідного розподілу, створюваного у зворотному ході моделі шляхом налаштування параметрів моделі за шарами. Оскільки в моделях обмежених машин Больцмана та заснованих на них МГВ зв'язки існують лише між нейронами сусідніх шарів, але не між елементами одного рівня, це дозволяє застосувати ефективні методи навчання, такий як метод контрастної розбіжності (КР), запропонованого Г. Хінтоном [72]. Вони дають змогу значно скоротити час навчання моделі без втрати точності.

МГВ – графічні моделі, які вчаться отримувати глибоке ієрархічне подання навчальних даних у повністю неконтрольованому режимі. Таким чином, МГВ моделюють спільний розподіл між вхідними даними X та іх прихованим розподілом h. Структура та принципи навчання ОБМ наведено на Рис. 1.5.

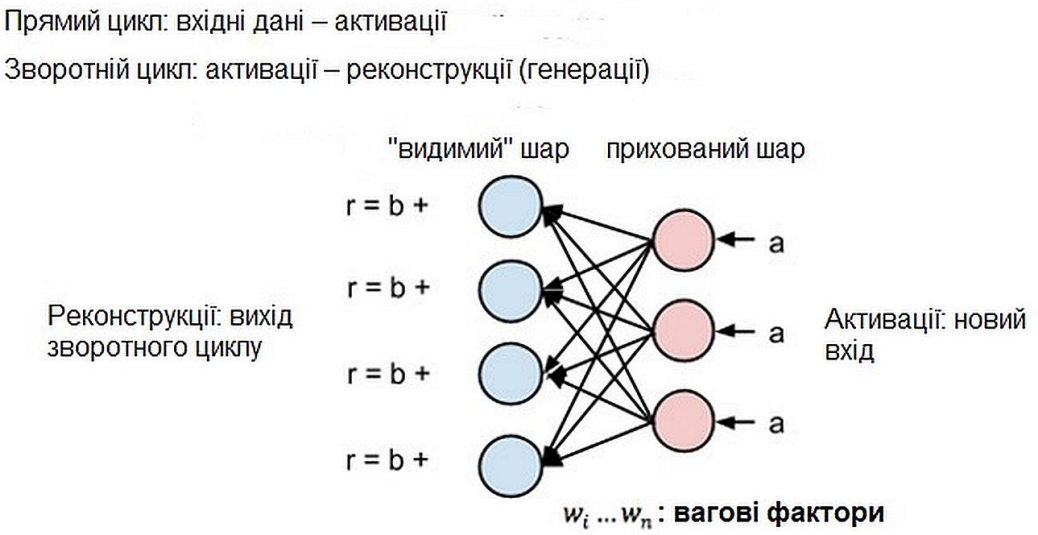


Рис. 1.5. Структура та навчання моделі ОМБ

В ряді робіт в різних галузях додатку відзначається підвищення ефективності подальшого контрольованого навчання при використанні параметрів, виділених за допомогою неконтрольованих моделей РБМ та МГВ. Ці результати добре узгоджуються з теоретичними висновками роботи щодо можливості підвищення ефективності класифікації при використанні категоризованих представлень вхідних даних, які виділяють ключові концептуальні структури, які можуть бути у кореляції з відомими категоріями даних.

В даний час моделі, засновані на ОМБ в основному поступилися місцем різним варіантам глибоких нейронних мереж автоэнкодера, які показують переважаючі результати як у створенні неконтрольованих представлень, так і в класифікації.

Моделі генеративного навчання. Задачею неконтрольованого генеративного навчання є відтворення вихідних даних з максимальною точністю по тренувальному набору, тобто:

$$mean_{\text{набір\_навчання}}\,(Measure \mid X, X'|) \rightarrow min \qquad (1.1)$$

де X, X’ – набір навчання та його відтворення моделлю, що навчається; міра відхилення може бути евлідівним квадратичним відхиленням координат $|X - X'|$, або більш складною функцією X, X'. У таких моделей критерій точності відтворення вихідної інформації означає, що представлення зберегло значну частину інформації про вхідний розподіл.

Відображення простору вхідних даних в простір представлення генеративної моделі називається "кодуюче перетворення" або "енкодер":

$$r = R(X), \tag{1.2}$$

тоді як відображення у зворотному напрямку, тобто з простору представлення в простір вхідних даних, “генеруюче перетворення”, перетворення відтворення, генератор або декодер (рис. 1.6):

$$X' = G(r) \tag{1.3}$$

Рис. 1.6. Перетворення кодування та генерації

Можна помітити, що позиція образу вхідного зразка X у просторі представлення r і перетворення генерації G повністю визначають його відтворення X' у просторі вхідних даних, як і відхилення відтворення $\varepsilon(X) = |X, X'|$. Істотний висновок, який випливає з цього спостереження полягає в тому, що параметри розподілу в просторі представлень, як і параметри відображення генеративної частини моделі, можуть мати істотне значення для визначення здатності моделей до ефективного навчання, включаючи здатність до категоризації даних у процесі неконтрольованого навчання.

Моделі нейронних мереж автоенкодера. Моделі автоенкодера стали популярними нині з кількох причин:

– по-перше, вони поєднують кодуюче і генеративне перетворення в єдиному процесі неконтрольованого навчання;

– по-друге, при навчанні автоенкодерів можуть використовуватися стандартні методи побудови та навчання нейронних мереж, в яких за минулий період були

зроблені значні досягнення, не в останню чергу такі, як можливість ефективно навчати мережі практично необмеженої глибини;

– по-третє, представлення, створені просунутими моделями автоенкодера в ряді робіт показали ефективність, яка перевуршує рівень раніше розглянутих моделей ОБМ та мереж глибокої віри.

Так в експерименті Google Labs було досягнуто вражаючих результатів з неконтрольованого розпізнавання зображень [36].

Загальна схема архітектури автенкодера наведена на рис .1.7.

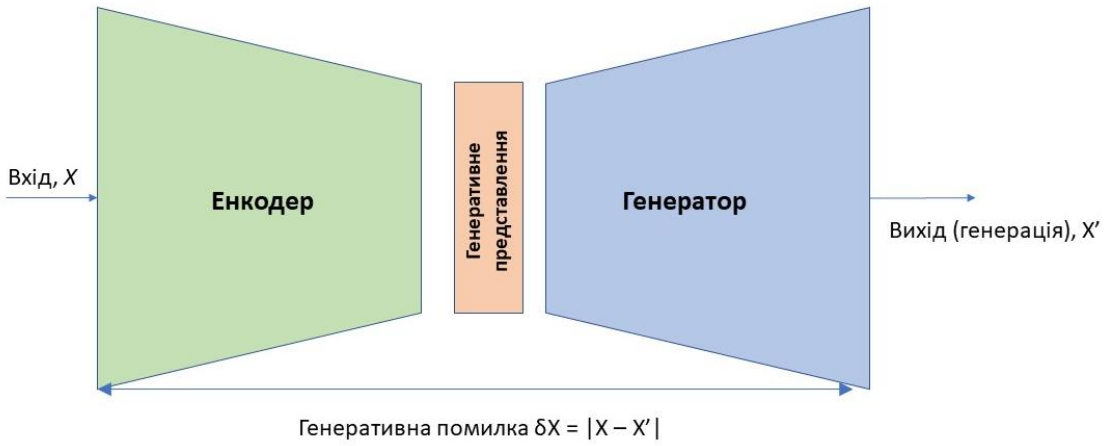


Рис. 1.7. Структурна схема архітектури автоенкодера

Навчання автоенкодера проводиться застосуванням методів зниження відхилення генерації від вихідного розподілу за стохастичним градієнтом. Розроблені за минулий період методи та алгоритми підвищення ефективності навчання, такі як алгоритми оптимізації, просунуті функції активації, пошарова нормалізація та багато інших, можуть використовуватися практично без обмежень та можуть суттєво підвищити ефективність навчання автоенкодерів.

Одним з важливих напрямків роботи з удосконалення автоенкодерів за минулі десятиліття стало створення моделей, що підвищують ефективність генералізації при навчанні за рахунок пригнічення можливостей перепідгонки даних (overfitting) та здатності до апроксимації складних розподілів реальних даних, таких як відеопотоки реального часу. В огляді робіт можна знайти велику кількість варіантів і архітектур автоенкодера з детальним описом [43, 44].

Важливим питанням є вибір конкретного типу, архітектури та структури моделі автоенкодера серед багатьох типів, які розглядалися. Наприклад, в даний час моделі розрідженого та багаторівневого автоенкодера широко використовуються для створення інформативних представлень даних зображень у додатках медицини, безпеки, відео та інших складних типів даних.

Теоретичне обґрунтування генеративних нейромережевих моделей, що використовувалися в роботі дано в розділі 2. Детальна структура моделей глибокого автоенкодера, що використовувалися в роботі, описана в розділі 3.1.

Інші типи моделей неконтрольованого навчання. Існує багато відомих методів неконтрольованого навчання та подання даних, такі як мережі Хопфілда [73], самоорганізуючі мережі [40, 41] та інші [74]. Детальний опис цих методів виходить за рамки цієї роботи.

1.2.5. Моделі та методи машинного навчання, засновані на нейробіології

В останні роки область машинного навчання, заснованого на біології, дуже стрімко розвивалася. Було досягнуто ряд значних результатів, які привели до ефективності та впевненості у навчанні систем машинного навчання і, зокрема, глибоких нейронних мереж, у кількох сферах застосування, таких як розпізнавання зображень, аналіз часових рядів, ігри та інші, в галузі людських здібностей або навіть переважаючи їх.

Детальний огляд основних сучасних розробок у галузі машинного навчання заснованого на біології, із застосуванням досягнень та результатів у галузі нейробіології для машинного інтелекту можна знайти в Д. Хасабіса та співавторів [75], особливо у застосуванні до загального навчання та спонтанного навчання, моделям безперервного навчання (Цікон-Ган, Хаясі-Такагі та інші.), імовірнісним моделям та моделям глибокого навчання (Лейк, Резенде та інші), прогресивному навчанню, концептуальним представленням та іншим, у той час як основні поняття, результати, перспективи та проблеми у застосуванні моделей машинного навчання були досліджені у Й. Бенджіо [41].

Важливо відзначити, що результати розвитку методів і моделей в машинному навчанні, засновані на нейробіологічних принципах, були отримані паралельно з недавніми досягненнями в експериментальних дослідженнях біологічних сенсорних мереж [76, 77], які продемонстрували спільність низькорозмірних нейронних представлень сенсорних даних, включаючи візуальні, звукові, нюхові при обробці сенсорної інформації в мозку ссавців, включаючи людину. Ці результати пропонують інтригуючі паралелі у процесах навчання штучних та біологічних систем.

1.3. Методи та критерії визначення інформаційної відповідності розподілів випадкових величин

Відомо кілька заходів спільної кореляції розподілів, які можуть вказувати на ступінь залежності або збереження інформації між розподілами випадкових величин, які, можливо, приймають значення в різних просторах.

1.3.1. Статистичні критерії

Коваріація розподілів випадкових величин X і Y визначається як [78]:

$$Cov(X,Y) = \sum_{x \in X, y \in Y} p(x,y)(x - X_{mean})\,(y - Y_{mean})$$

Коваріація розподілів X і Y має розмірності $[X]\times[Y]$ та показує ступінь координації між варіаціями розподілів X і Y. Можна помітити, що для повністю незалежних розподілів $p(x,y) = p(x) \times p\,(y)$ і коваріація зникає.

Кореляція розподілів X і Y може бути визначена через коваріацію та такі параметри розподілів як:

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\,\sigma(Y)}$$

Зауважимо, що на відміну від коваріації вказаної вище кореляція безрозмірна, і її значення нормоване в інтервалі [-1, 1], що дозволяє уявити не лише наявність чи відсутність, а й відносне значення залежності між розподілами. Так, значення, що дорівнює нулю спостерігається для повністю незалежних розподілів, тоді як

значення близьке до 1 по абсолютній величині означає сильну залежність між розподілами.

Також для векторних змінних можуть бути визначені матриці коваріації та кореляції, зі значеннями рівними коваріації або кореляції між параметрами розподілів. Розмірність матриць коваріації та кореляції розподілів X, Y дорівнює (m, n), де m і n представляють відповідно розмірності розподілів X і Y.

1.3.2. Інформаційні критерії

Спільна ентропія. Спільна ентропія розподілів X і Y визначається за формулою Шеннона:

$$H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} p(X,Y)\log p(X,Y).$$

При аналізі представлень, співвіднесених з вихідними розподілами є деякі суттєві спостереження. Якщо розподіли корельовані, $p(X, Y) \sim p(X)$ x 1, тобто розподіл Y визначається X, то спільна ентропія X та Y дорівнює ентропії розподілу X. На протилежному кінці спектра, якщо розподіли незалежні, $p(X, Y) = p(X) \times p(Y)$ і спільна ентропія є добутоком ентропій розподілів:

$$H(X,Y) = H(X) \times H(Y) \geq H(X).$$

Таким чином, порівнюючи спільну ентропію розподілу та представлення з ентропією вхідного розподілу можна визначити ступінь залежності, або збереження суттєвої інформації між вхідним розподілом та його представленям.

З іншого боку, недоліком цього методу як і методу взаємної інформації нижче є те, що він вимагає явного знання спільного розподілу, тобто вибірки кожної точки у вихідному розподілі з кожним зображенням у просторі представлення. Для складних типів даних з великою кількістю параметрів закони, розподіл яких невідомий заздалегідь, обчислення спільного розподілу може бути непрактичним або навіть нерозв'язним завданням.

Взаємна інформація. Взаємна інформація – величина, яка вимірює залежність між двома розподілами, які вимірюються разом. Зокрема, вона може повідомити скільки інформації передається в середньому в одній випадковій змінній щодо

іншої, тобто, якою мірою інформація про один розподіл може повідомити про інше. Взаємна інформація двох розподілів визначається на основі дивергенції Кулбека-Лейблера [79]:

$I(X,Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\ log\frac{P(x,y)}{P(x)\,P(y)}$для дискретного розподілу або у безперервній формі:

$$I(X,Y) = \int_{x\in X}\int_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)\,p(y)}\ dx\,dy$$

Так само як і у випадку спільної ентропії, взаємна інформація може дати представлення про залежність розподілів. Можна побачити, що у разі повністю незалежних розподілів $p(X,\ Y) = p(X) \times p(Y)$ та взаємна інформація розподілів дорівнює нулю, тобто за значенням одного розподілу не можна зробити жодних висновків про інше. У протилежному випадку, коли Y детерміновано визначається X легко помітити, що взаємна інформація розподілів X і Y зводиться до ентропії розподілу X, тобто X містить всю інформацію про спільний розподіл.

Іншими представлважливими властивостями взаємної інформації є симетричність та невід'ємність.

$$I(X,Y) = I(Y,X);\ I(X,Y) \geq 0$$

Нарешті, слід зазначити, що хоча ці визначення зручні в теоретичних дослідженнях, вони завжди ефективні у практичних додатках через необхідність повного знання спільного розподілу у просторі вихідних даних, що може вимагати значних ресурсів для великих масивів даних із значною кількістю параметрів.

Оцінка ентропії емпіричного розподілу. Відносно простий та безпосередній інформаційний критерій оцінки ступеня впорядкованості в загальному наборі даних без відомого характеру розподілу заснован на обчисленні емпіричної ентропії даних та порівнянні її з абсолютно невпорядкованим розподілом, таким як рівномірний розподіл.

У простому випадку даних, представлених N однаковими параметрами (каналами), такими як зображення, представлене набором ідентичних пікселів (A,

B) з числовими значеннями, що представляють кольори, оцінка емпіричного та граничного значення ентропії даних за формулою Шеннона може бути наступною:

$$E_{emp} = \sum_{N} \sum_{k=1..B} -p_k \log p_k$$

де p_k – емпірична ймовірність значення параметра у k-му кошику; B – кількість кошиків вибірки та граничне значення ентропії $E_{lim} = N \times \log B$.

Таким чином, для даних загального типу навіть без суттєвого знання про характер та семантичний зміст даних, обчислення "фактора порядку" $\frac{E_{emp}}{N \times logB}$, заснованого на емпіричному значенні ентропії набору при розумному виборі дозволу, тобто кількості кошиків в інтервалі розподілу, може дати представлення про ступінь порядку даних і можливості їх ефективного стиснення з виділенням відносно невеликої кількості інформативних категорій.

Принцип виведення Байєса в машинному навчанні. Важливою теоретичною основою досліджень як у контрольованому, так і неконтрольованому навчанні систем машинного інтелекту є принцип байєсівського висновку [80], його наслідки та додатки. Він являє собою платформу для розуміння механізмів навчання систем та їх взаємозв'язку, як, наприклад, зв'язок методів навчання моделей зниження за стохастичним градієнтом (СГЗ) та контрастною розбіжністю (КР) серед багатьох інших.

Однак у практичних додатках спроби прямого застосування методу часто стикаються з серйозними обчислювальними перешкодами. Прикладом є проблема обчислення коефіцієнта нормалізації ("проблема статсуми" [81, 82]), яка в багатьох практичних додатках, особливо при складних розподілах та типах даних, виявляється нерозв'язною.

Ряд методів було запропоновано та успішно використано для подолання обчислювальних труднощів методами апроксимації, таких як вибірки Гіббса, статистичні вибірки типу ланцюга Маркова, Монте-Карло (ЛММК, англ. MCMC), варіаційні байєсівські методи та інші. У роботі варіаційні методи

використовувалися під час доказу теореми про категоризацію генеративних представлень (розділ 2).

Методи енергетичного навчання. Моделі енергетичного навчання фіксують залежності, пов'язуючи скалярну енергію (міру сумісності) з кожною зміною конфігурацією. Засноване на енергії навчання забезпечує уніфіковану основу для багатьох пробелістичних та неймовірних підходів до навчання, особливо для неймовірного навчання графічних моделей та інших структурованих моделей (Ранзато та співавтори [83]).

Наприклад, для моделей, заснованих на енергії, таких як ОВМ та глибоких нейронних мереж енергетична функція може бути записана наступним чином:

$$F(x) = -\log \sum_{h} e^{-E(x,h)}$$

де *x, h* – зразок вхідного, латентного розподілу; *E(x, h)* – енергетична функція стану, що визначається параметрами моделі, такими як вага і зміщення, і значеннями вхідних та латентних вибірок.

Вільна енергія моделі є функцією конфігурації її моделі та вхідної вибірки, яка визначається як міра варіації регенерації вхідних даних. Для деяких моделей, таких як ОВМ можна отримати явну формулу вільної енергії конфігурації через параметри моделі [84]:

$$F(x, w) = -b'x - \log \sum_{h_i} e^{h_i (c_i + w_i x)}$$

де *b, c, w* – параметри моделі, такі як зміщення вхідних та латентних елементів та ваг, *x* – вхідний зразок. Ряд моделей і наближень було запропоновано для енергетичної функції глибоких нейронних мереж [85].

В останні роки значна увага приділялася встановленню зв'язку між навчанням моделей з алгоритмами мінімізації помилок, такими як СГ, КР та методами статистичної термодинаміки систем, що навчаються. Так, згідно з результатами [82] було встановлено при певних припущеннях, що навчальні алгоритми є наближенням байєсівського виведення на масиві навчальних даних.

1.4. Методи, програмні продукти, бібліотеки обробки даних та машинного навчання

У практиці розробки моделей машинного навчання, у тому числі нейромережевих, широко використовується програмне середовище Python [86] завдяки легкості та інтуїтивності роботи і великої кількості спеціалізованих бібліотек, що надають найширші можливості обробки даних та швидкої імплементації навіть складних моделей.

При роботі зі складними моделями нейронних мереж, такими як глибокі нейронні мережі поширені програмні продукти Keras, Tensorflow [87], що дозволяють швидко та гнучко розробляти нейромережні моделі найвищої складності з широким вибором спеціалізованих компонентів та функцій, такими як функції активації, вартості, компоненти оптимізації та інші. Також для обробки даних, обробки, представлення та візуалізації результатів використовувалися пакети sklearn-kit та інші модулі Python, що дозволило спростити та прискорити підготовку, проведення та обробку результатів експериментальної частини та подання результатів.

1.5. Експериментальні результати у неконтрольованому навчанні

Перший напрямок, це відомі з 1990 р. – середини 2000 р. результати, які демонструють, що попереднє навчання деяких моделей машинного навчання в неконтрольованому режимі, що забезпечує хорошу якість представлення (тобто хороше відтворення вихідних даних з побудованого представлення), може призвести до відчутного поліпшення точності класифікації ([46] та аналогічні). Найпростішим поясненням цього ефекту є те, що такі представлення мають структури, пов'язані з характеристичними типами вхідних даних, що полегшує навчання та підвищує його ефективність.

Наступна група результатів була отримана в експериментах зі сприйняття візуальних стимулів, тобто зображень у середині – наприкінці 2010-х [35, 36]. Наприклад, командою Google Mind було продемонстровано повністю неконтрольоване навчання розпізнавання класів зображень моделлю глибокого

розрідженого автоенкодера на великих масивах зображень. Хоча точність ідентифікації не була достатньо високою для класифікації на впевненому рівні, експеримент продемонстрував, що певні ефекти спонтанної кластеризації за категоріями високого рівня можуть бути результатом повністю неконтрольованого спостереження та обробки інформації в процесі генеративного навчання.

Прямі результати кластеризації із спонтанної кластеризації даних (категоризації) при неконтрольованому генеративному навчанні були представлені в роботах по структурі щільності генеративних представлень даних зображень, Інтернет, епідеміологічних та інших типів. Було продемонстровано, що внутрішній простір представлень може мати структуру, корелювану з відомих категорій вхідних даних, розроблено методи визначення, вимірювання, аналізу та візуалізації розподілів даних у представленнях моделей генеративного навчання [88, 89].

Насамкінець, не можна не відзначити результати експериментальної нейробіології, згадувані в розділі 1.2.5, які вказують на важливість низькорозмірних представлених сенсорних даних при обробці інформації біологічними організмами.

Огляд стану галузі досліджень, методів та моделей неконтрольованого навчання дозволяє говорити про те, що використання генеративних нейромережевих архітектур може бути ефективним підходом у створенні інформативних представлень даних та просунутися до розв'язання поставленого завдання успішного навчання штучних систем інтелекту з мінімальними відомими даними. Для подальшого просування потрібен розгляд теоретичних основ систем генеративного навчання.

РОЗДІЛ 2.
ДОСЛІДЖЕННЯ МЕТОДІВ НАВЧАННЯ З ВИКОРИСТАННЯМ ГЕНЕРАТИВНОЇ СТРУКТУРИ НЕКОНТРОЛЬОВАНИХ ПРЕДСТАВЛЕНЬ

У даному розділі розглянуто теоретичні основи моделей та методів у галузі неконтрольованого навчання неконтрольованих представлень моделей самокодування та генерації, даних та методів навчання таких моделей, що використовувалися суттєво або мають істотне відношення до напрямку дослідження, методів та результатів роботи.

У розділі наведено визначення основних концепцій теорії неконтрольованого навчання, розглянуто принципи та методи побудови та навчання моделей неконтрольованого навчання, створення та аналізу неконтрольованих представлень та доведено ключові результати для обґрунтування принципів та методів побудови моделей неконтрольованого навчання, що використовувалися в експериментальній частині роботи.

На підставі отриманих теоретичних результатів обґрунтовано та розроблено методи навчання систем на основі структури щільності (ландшафту) генеративних представлень даних пакетів трафіку Інтернет: метод розпізнавання натуральних концептів даних, що не вимагає відомих даних навчання; та метод метод навчання розпізнавання відомих класів даних з мінімальними вимогами до навчальних наборів маркованих даних (до одиничних зразків).

У заключній частині розділу проведено аналіз отриманих результатів, зроблено висновки, що пов'язують результати роботи з іншими областями теорії систем машинного навчання та неконтрольованого навчання, а також запропоновано гіпотези та зроблено припущення для експериментальної перевірки.

2.1. Визначення основних понять

Вхідні (сенсорні) дані. Розглядаються деякі вхідні дані довільного типу, що характеризуються явними або "спостережуваними" параметрами: $X(t)$. Це можуть бути дані різних типів: текст, числові значення, колірні канали пікселів зображення, багатовимірні представлення потоків зображень тощо. У роботі

розглядаються дані пакетів Інтернет, представлені вектором числових значень характеристик пакетів, і дані зображень, представлені двомірним числовим вектором значень яскравості $X(t) = (a_1, a_2, .. a_N)$.

Представлення. Представлення вхідних даних визначається кодуючим перетворенням:

$$E: I \rightarrow R, x(t) \rightarrow y(r)$$

де: x, r – параметри, або координати у просторах вхідних даних та представлення I, R. Таким чином, кожному розподілу даних $X(t)$ у просторі вихідних даних I відповідає розподіл $Y(r)$ у просторі представлення R, яке визначається параметрами представлення та перетворенням кодування.

Приховані параметри. Для підмножини в просторі даних X приховані параметри $H = \{ h_k \}$, $X = \{ x(h_k) \}$ визначаються як "натуральні" або неявні координати, що характеризують підмножину даних. Наприклад, у наборі даних зображень можна виділити підмножину зображень облич, яке характеризуються специфічним значенням параметрів, характерних тільки для облич (очі, брови, ніс і т. д.), які в цьому випадку можна вважати прихованими параметрами підмножини облич у просторі загальних зображень.

Таким чином, можна розглядати розподіли концепцій класів, типів даних як у вхідному просторі X, так і у просторі представлення R, які визначаються прихованими параметрами h_k:

$$X_k = X(h_k);\ r_k = r(h_k) \qquad (2.1)$$

Слід зазначити, що приховані параметри: 1) не спостерігаються, тобто не виявляються безпосередньо чи явно у вхідних даних; 2) невідомі апріорі; тобто явна формула або алгоритм, який пов'язує вхідну вибірку з прихованим параметром (або "витягує" приховані параметри з вхідної вибірки) може не бути відомий до, і в процесі навчання генеративної моделі.

Кодування та генерація. Як зазначалося вище, кожній вибірці в просторі вхідних даних перетворення кодування ставить у відповідність її образ у просторі

представлення. Таким чином, можна сказати, що перетворення кодування дозволяє представляти вхідні дані новим набором параметрів у просторі представлення.

Як правило, число параметрів, тобто розмірність простору представлення менше чи значно менше простору вхідних даних. Наприклад, у разі набору даних зображень, використовуваного в роботі, розмірність вхідного простору становила до 4,000 параметрів (при розмірі зображень 64 × 64 пікселів), у той час як простір представлення мав розмірність 3 або три прихованих координати – параметри представлення, тобто йдеться про стиснення з фактором більше 1,000.

Генеративна модель визначається як відображення з простору представлення назад у простір вхідних даних:

$$G: R \to I;\ Y(t') = G(y(r)) \tag{2.2}$$

так що принаймні для підпростору *r* в просторі представлення *R* генерований образ *G(r)* належить вхідному розподілу *X(t)*.

При перетвореннях кодування та генерації *R*(*x*), *G*(*y*) далі у роботі *R(x)* називатиметься прообразом, *G(y)*, генерацією або відтворенням зразка *x* у просторі вхідних даних. Таким чином, процес навчання генеративних моделей – ОВМ або автоенкодер ґрунтується на зменшенні помилки відтворення навчального набору *X*(*t*) у вхідному просторі. У випадку моделей заснованих на нейронних мережах, таких як автоенкодер, навчання може здійснюватися стандартними методами навчання нейронних мереж, такими як зворотне розповсюдження (англ. backpropagation, [90]), СГЗ та аналогічними.

Неконтрольоване генеративне навчання з мінімізацією помилки відтворення. Моделі машинного навчання навчаються мінімізувати помилку прогнозування на основі набору навчальних даних і наперед відомих "правильних" результатів. З кожною партією навчальних даних параметри моделі оновлюються, щоб мінімізувати похибку її прогнозу порівняно з "правильним" значенням. Ця помилка вимірюється функціоналом розбіжності або втрат *L*, що визначається вибором функції вимірювання відхилення, або метрикою в просторі передбачень, яка дозволяє обчислити значення помилки при поточних параметрах моделі на наборах

ітерацій навчання X_k. Процес оновлення параметрів моделі потім повторюється до досягнення бажаної точності передбачення:

$$L = L(X_k, W, f(X_k, P(W, X_k))) \quad (2.3)$$

де X_k – навчальний набір k-й ітерації навчання; W – поточні параметри моделі, f – функція метрики у просторі передбачень; $P(W, X_k)$ – передбачення моделі на навчальному наборі ітерації.

У випадку моделей генеративного навчання простір передбачень збігається з простором вхідних даних і функціонал помилки визначається метрикою в просторі вхідних даних. Необхідно відзначити, що ця метрика не обов'язково має бути евклідовою, і в більш складних моделях часто буває відмінною від неї або модифікованою додатковими складовими.

Широко використовуваними прикладами функцій помилок є середньоквадратична помилка відхилення (СПВ) заснована на евклідовій метриці у просторі навчальних даних; категорична і бінарна крос-ентропія [90] в контрольованому навчанні з маркованими даними. В методах неконтрольованого навчання, при невідомих заздалегідь явних класах або категоріях вихідних даних, мова може йти про міри відхилення між вхідними даними та результатом, генерованим моделлю на включаючи вже назване середньоквадратичне відхилення, відхилення за абсолютною величиною, додаткові, що забезпечують розрідженість [91] та інші функції втрат.

2.2. Категоризовані представлення

Особливий інтерес становлять представлення, у яких розподіл даних складається з кінцевого набору “добре визначених” розподілів (уточнення буде зроблено далі), які вважатимуться характерними типами даних, або “натуральними концептами”. На відміну від задач контрольованого навчання, такі концепти є зазвичай “прихованими”, тобто невідомими заздалегідь до і в процесі навчання. Представлення, які відповідають цим умовам будуть називатися "категоризовані представлення". Метою або однією з цілей неконтрольованого генеративного навчання може бути виділення та визначення прихованих концептів вхідних даних.

Цей напрямок може збігатися із завданням роботи, оскільки визначення структур натуральних типів даних може надати додаткову інформацію про розподіл вхідних даних, у процесі якої не вимагає значних обсягів відомих даних та підвищити впевненість та якість навчання навіть з мінімальними відомими даними.

При аналізі неконтрольованих представлень суттєве значення мають такі критерії:

1) точність відтворення, тобто збереження моделлю суттєвої інформації вхідного розподілу, що забезпечує значну точність відтворення даних зі стисненого представлення; зокрема, у генеративних моделях може йтися про точність відтворення вхідний партії навчання, яка може бути виміряна середньою помилкою відтворення партії;

2) загальність, тобто стабільність точності відтворення при розширенні або зміні набору навчання, оскільки його характеристики зберігаються;

3) стиснення вхідних даних, наприклад фізично (за розмірністю простору представлень) або логічно (за обмеженнями накладеними на простір представлень), що забезпечують виділення важливих інформативних структур у вихідних даних.

Як було зазначено у ряді опублікованих результатів, деякі моделі неконтрольованого навчання, у тому числі генеративні моделі, можуть створювати категоризовані представлення вхідних даних у процесі неконтрольованого навчання не вимагаючи відомих даних навчання.

2.3. Теоретичні підходи у неконтрольованому генеративному навчанні

Припустимо, що існує перетворення простору вхідних даних в деякий ефективний простір представлення, яке відображає області відповідні деякому набору прихованих концептів вхідних даних в компактні регіони в просторі представлення без значного перетину. Таке перетворення буде називатися "категоризуючим перетворенням".

Формальна постановка задачі. Припустимо задані набори даних: D, довільного розміру з представленими відомими категоріями $C = \{ C_k \}$ та набір

навчальних даних D_0. За умовою завдання або середовища застосування, або a) розмір набору D_0 недостатній для застосування традиційних методів контрольованого навчання; b) навчальні дані набору D_0 не цілком відповідають розподілу вхідних даних (це можливо наприклад у випадках, коли набори D, D_0 отримані з різних джерел). Таким чином, виникає обмеження що застосування стандартних методів контрольованого навчання на наборі D_0 може не забезпечити бажаного рівня точності, як було зазначено у ряді опублікованих результатів у додатках до аналізу даних мереж Інтернет.

Потрібно визначити методи: $M(x, C)$: розпізнавання відомих категорій (класів) $C(x) = \mathrm{M}(x \in X, C)$ на просторі вхідних даних $D \in X$, який задовольняє умові прийнятної точності: $e(x \in D) \leq \varepsilon_{max}$, e: середня помилка розпізнавання; та $P(x)$: визначення та розпізнавання натуральних (прихованих) концептів даних, $n(x) = P(x)$, де $N = \{ n_k \}$ – набір натуральних концептів визначений методом на основі набору вхідних даних D, також з умовою прийнятної точності.

Підходи до розв'язання задачі можуть бути встановлені на основі теорії генеративного навчання, яка дозволяє створювати інформативні представлення вхідних даних методами неконтрольованого навчання з мінімізацією помилки відтворення навчальних даних, що не вимагають значних наборів маркованих відомими категоріями.

Припустимо, існує модель генеративного навчання, яка здійснює перетворення кодування та відтворення вихідних даних X у просторі представлення зниженої розмірності R. При виконанні умов 1. Точності відтворення, що вимірюється певною метрикою у вихідному просторі (функція втрат); 2. Зниження розмірності простору представлення, та 3. Загальності, тобто незалежності середньої точності відтворення від конкретного набору даних, значна кількість сучасних опублікованих результатів дають підстави очікувати, що структура розподілу даних у просторі представлення може бути у кореляції з характерними типами вхідних даних представлених вибіркою навчального набору. Далі, значне зниження розмірності простору представлення порівняно з розмірністю вихідних даних значно полегшує розпізнавання інформаційної структури представлення

методами кластеризації, у тому числі неконтрольованими, тобто такими що не вимагають зразків відомих концептів. Неконтрольоване навчання генеративних моделей, таких як нейромережеві моделі типу автоенкодера може проводитися стандартними методами навчання, такими як методи зворотного розповсюдження, стохастичного градієнта та іншими, з представницькими наборами даних без асоціації з відомими класами або категоріями.

2.3.1. Категоризація при контрольованому навчанні

У разі контрольованого навчання, коли набір категорій відомий заздалегідь, легко довести наступну лему:

Лема 1: Категоризуюче перетворення при контрольованому навчанні (класифікації).

Твердження: наявність категоризуючого перетворення T для деякого набору вхідних даних D із відомим наборім категорій, або класів $\{ C \}$ еквівалентно наявності моделі, здатної класифікувати D в C із середньою помилкою нижче деякого фактора ε при будь-якій підмножині даних з D.

Доведення. Якщо перетворення категоризації T існує для всіх підмножин концептів (класів) вхідних даних P_k, розглянемо образи підмножин класів $R_k = \{ C_k \}$ в просторі представлення R, яке визначається $T: I \rightarrow R$.

Відповідно до визначення категоризуючого перетворення, вони є компактними і розділеними регіонами в R. Тоді згідно з результатами універсальної апроксимації нейронних мереж [69], існують нейронні мережі S_k, які апроксимують T та відображають $P_к$ у простір класів C_k з будь-якою заданою точністю для кожного класу. Приклад таких мереж може бути побудований, наприклад, за допомогою алгоритму степ-функції [92] хоча вони можуть бути не мінімальними та не найефективнішими. Тоді можна збудувати комбінацію мереж $S_к\, S(I) \rightarrow \{ C \}$ яка забезпечує класифікацію вхідних даних в C з будь-якою заданою точністю.

Зворотне твердження може бути доведено безпосередньо, оскільки категоризуючим перетворенням у цьому випадку може бути сама модель $S(I)$, яка

забезпечує класифікацію вхідних даних у простір відомих класів C, а простір представлення R у цьому випадку можна вважати ідентичним простору C:

$$T: S(I) \rightarrow C;\ R = C,$$

що завершує доказ леми.

Незважаючи на те, що наведений доказ досить простий, з леми можна зробити цікавий висновок: властивість категоризованості, тобто стійкого та передбачуваного розподілу вихідних даних за відомими категоріями є властивістю самих даних, і моделі штучного навчання, такі як глибокі нейронні мережі, є реалізацією цієї внутрішньої властивості.

При неконтрольованому навчанні, коли набір прихованих концептів вхідних даних не є відомим заздалегідь, знаходження розбиття представлення на добре визначені регіони розподілу концептів аналогічного доведеної леми є нетривіальним завданням неконтрольованого навчання.

2.3.2. Категоризація при неконтрольованому навчанні

Неконтрольована категоризація означає здатність деяких моделей неконтрольованого навчання групувати вибірки даних у компактні структури, такі як кластери у просторі представлень на основі схожості у просторі вхідних даних. Як показують опубліковані результати такі натуральні інформаційні структури можуть виникати у генеративних представленнях при неконтрольованому навчанні з мінімізацією помилки відтворення і можуть бути корельовані з характеристичними типами вхідних даних.

Відомим прикладом спонтанної категоризації подібного типу є вищезгадані результати дослідницької групи Google Labs з неконтрольованої обробки зображень [34, 35], в яких було виявлено спонтанне виникнення структур у нейронних мережах глибокого розрідженого автоенкодера чутливих до певних категорій при неконтрольованому навчанні зображень, а також значна кількість інших результатів.

Встановлення зв'язку між інформаційними структурами, що виникають у представленнях створених у процесі неконтрольованого навчання та характерними

категоріями вхідних даних є одним з ключових напрямів досліджень у сфері неконтрольованого навчання. Цей висновок є теоретичною основою експериментальних спостережень прикладів спонтанної категоризації даних різних типів у кількох роботах, включаючи моделі, використані в даній роботі.

Доказ затвердження аналогічного Лемі 1 для випадку неконтрольованого навчання, тобто без заздалегідь відомого набору категорій вхідних даних є значно більш складним. Тим не менш, за певних припущень та умов, доказ леми про спонтанну категоризацію неконтрольованих представлень наведено в роботі [93]. У цьому розділі ми наведемо стисле зведення доказу теореми про категоризовані представлення при неконтрольованому генеративному навчанні.

Теорема: Категоризовані представлення при неконтрольованому генеративному навчанні.

Твердження: в ансамблі моделей неконтрольованого генеративного навчання, які відповідають умовам успішного навчання**:**

1. Мінімізації генеративної помилки навчання: $\overline{L_{gen}(B)} \leq \varepsilon$, L_{gen} – функціонал помилки; B – набір навчання вхідних даних;

2. Значного зниження розмірності даних: $d_{пред} \ll d_{вх}$, розмірності просторів генеративного представлення, вхідного;

3. Загальності: виконання умови 1 при будь-якому наборі навчання $B \in I$; та при додатковому припущенні про склад вхідних даних: $I = U\, O_k + N$, де O_k – кінцева кількість характерних типів (концептів) вхідних даних; N – випадковий шум, статистично переважають моделі з генеративними представленнями "хорошої категоризації", де під "хорошою категоризацією" розуміються розподіли з мінімальним перетином областей розподілу концептів:

$$\sum_{i,j} \dim O_{i,j} = (H_i \cap H_j) \to min \qquad (2.4)$$

Доведення. Припустимо, що розподіли представлень концептів O_k визначаються параметрами розподілу h_k в просторі представлень, наприклад k-мірних кластерів. Нагадаємо, що при неконтрольованому навчанні модель не має

даних про розподіли концептів, включаючи їх число, області розподілу тощо. З умови теореми випливає висновок про склад простору представлення:

$$R = \bigcup_k H_k + N,$$

де H_k – латентні області розподілу концептів, N – розподіл випадкової складової даних (випадковий шум), не пов'язана з будь-якими характерними типами (тобто незалежний випадковий розподіл за кожним параметром).

Далі розглянемо параметри $\{ g_k \}$ генеруючої моделі G, що діє з простору представлення у вхідний простір, наприклад, вага і зміщення (w, b) у моделях нейронних мереж:

$$G: R \rightarrow I;\ G(r) = F(g_k, r) \tag{2.5}$$

При визначених таким чином латентних розподілів концептів H_k та генеративних параметрів g_k генеративна конфігурація може бути визначена як відображення:

$$P_{gen} = (\{H_k\} = (\{ h_k\}), \{ g_k \}): X \in I \tag{2.6}$$

тобто параметри латентних розподілів концептів та генеративні параметри навченої моделі за умовою 1 теореми повністю визначають невипадкову складову вхідного розподілу.

Припустимо далі, що існує генеративна конфігурація K, яка відповідає та визначена деякою генеративною моделлю ансамблю, яка максимізує категоризацію латентних розподілів концептів $\{ H_k \}$, максимізує відділення латентних областей розподілу концептів згідно умови (2.4). Спробуємо показати, що така конфігурація також мінімізує вільну енергію відображення генерації, і відповідно до принципів та навчання з мінімізацією енергії має більш високу ймовірність спостереження порівняно з іншими конфігураціями.

У загальному вигляді зв'язок між функцією втрат та енергією конфігурації моделі із заданим зразком даних можна записати у вигляді (2), [83]. Адаптована на випадок генеративного навчання, функція втрат $L(X, g)$ представляє помилку

генерації при генеративному відображенні з прихованими концептами H_k (ґрунтуючись на властивості асоціативності суми):

$$L(X, g) = \frac{1}{S} \sum E(h_k, g) + {}^{1}/_{\beta}\ log \sum V(h_k,\ g), \tag{2.7}$$

де g – параметри генеративного відображення; $E(H_k,\ g) = E\ (h_k,\ g)$ – енергія генеративного відображення латентного розподілу концепту H_k; S, β – константи; $V(h_k,\ g)$ – можна інтерпритувати як локальну варіацію латентного розподілу H_k: $V(h_k, g) = \int_{y \subset H_k} e^{-\beta y}\, dy$

Далі, щоб визначити напрями змін функцій енергії та помилки при варіації параметрів регенеративної конфігурації, допустимо варіацію параметрів генеративної конфігурації K: $\delta p = \{\delta h_k,\ \delta g_i\}$ та оцінимо відповідний диференціал функції втрат та функції вільної енергії перетворення генерації ${}^{\partial L}/_{\partial p}\ |_K$, ${}^{\partial E}/_{\partial p}\ |_K$. Зауважимо, що згідно з [83] навчання зі зменшенням помилки еквівалентне зменшенню вільної енергії, так що при малих варіаціях можна припустити що:

$$\delta E = {}^{\partial E}/_{\partial p}\, \delta p = {}^{\partial E}/_{\partial L} \times {}^{\partial L}/_{\partial p}\, \delta p = \alpha\ {}^{\partial L}/_{\partial p}\ \delta p \sim \delta p\ {}^{\partial L}/_{\partial p} \tag{2.8}$$

де $\alpha = {}^{\partial E}/_{\partial L}\ |_{\ p = K}$ – позитивна константа.

Починаючи з параметрів латентного розподілу $\{ h_k \}$ можна помітити, що якщо область розподілу концепту H_k у просторі представлення збільшується, в той час як генеративні параметри моделі залишаються постійними, вибірки за межами вихідної області розподілу концепту можуть бути зіставлені перетворенням генерації з областю у вхідному просторі зайнятою іншим концептом, збільшуючи помилку типу "хибний негатив", δL_1. Також, зразки різних “чужих” концептів можуть бути присутніми в розширеній області концепту δH_k і можуть також відображатися у неправильну концептуальну область у вхідному просторі, збільшуючи помилку "хибний позитив" δL_2, і як наслідок, загальну помилку генерації:

$$\delta L = \delta L_1 + \delta L_2\ \geq 0$$

На ілюстрації рис. 2.1 області розподілу концепту в генеративному представленні H_k та вхідному просторі I_k позначені зеленим кольором, області розподілу інших концептів, відповідно, H_n, I_n.

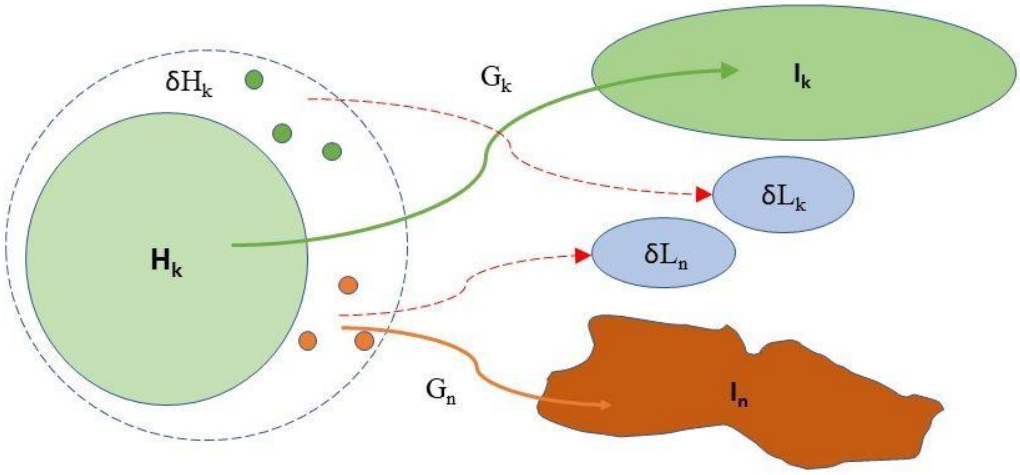


Рис. 2.1. Збільшення помилки відтворення при варіації області розподілу концепту

При варіації параметрів розподілу області H_k частина зразків трансформуються у області вхідних розподілів інших концептів (області δL_k, δL_n), збільшуючи помилки хибних негативів та хибних позитивів, і в результаті, загальну помилку генерації. Таким чином, можна зробити висновок, що варіація області розподілу від стану максимальної категоризації призводить до збільшення помилки відтворення.

Очевидно, що цей аналіз вірний для кожного концепта H_j і, повторюючи аргумент вище, можна дійти висновку, що конфігурація мінімуму функції вільної енергії досягається, якщо області розподілу для всіх прихованих концепцій компактні з мінімальним перетином, тобто задовольняють умови (2.4) теореми. Таким чином,

$$\delta L = \delta L_k + \delta L_i = \sum_{i \in K} \delta L_i \geq 0$$

і згідно (2.8):

$$\frac{\partial E(X,g)}{\partial h_k} \sim \frac{\partial L(X,g)}{\partial h_k} \geq 0 \quad (2.9)$$

Тепер розглянемо варіацію параметрів генеративної моделі $G = \{ g_i \}$ при фіксованих параметрах розподілів прихованих концептів. Нагадаємо, що параметри конфігурації моделі з проведеним навчанням зводять до мінімуму функцію втрат генерації, тому будь-яка зміна значень параметрів моделі збільшить значення функції втрат, і як наслідок,

$$\frac{\partial E(X,g)}{\partial g_i} \sim \frac{\partial L(X,g)}{\partial g_i} \geq 0 \quad (2.10)$$

Таким чином, з (2.9) та (2.10) можна зробити висновок, що для всіх параметрів $P = \{ h_k, g_k \}$ генеративної конфігурації K:

$$\frac{\partial E(X,g)}{\partial p}\Big|_{\text{K}} \sim \frac{\partial L(X,g)}{\partial \text{p}}\Big|_{\text{K}} \geq 0 \quad (2.11)$$

Звідси випливає, що конфігурація з максимальною категоризацією представлення мінімізує як вільну енергію, так і помилку генерації моделі.

Доказ теореми тепер випливає безпосередньо з результатів [83] про зв'язок між енергетичним навчанням та принципом мінімізації вільної енергії, згідно з якими конфігурації навчання слідують розподілу Больцмана-Гіббса [94], тобто:

$$p(K) \sim e^{-E(K)} = e^{-E(h_k,\ g_k)}$$

де $p(K)$ – статистична можливість спостереження конфігурації K у результаті навчання. Дійсно, оскільки категоризовані представлення мають нижчу енергію генеративного відображення в конфігураціях з максимальною категоризацією, вони будуть статистично кращими відповідно до закону розподілу Больцмана-Гіббса, і при виконанні умов теореми, тобто умов точності відтворення; загальності та зниження розмірності представлення, максимум ймовірності конфігурації збігається з максимумом категоризації прихованих концептів у просторі представлення.

Доведений результат можна проілюструвати на простому прикладі. Припустимо, що у вихідних даних присутні зразки двох прихованих концептів. Розглянемо два приклади розподілів концептів у просторі представлення: максимально категоризований, при якому розподіли концептів займають компактні

та ізольовані області у просторі представлення; і “розмите”, де області розподілу концептів розподілені по простору представлення і перекриваються значною мірою.

Можна помітити, що перший, категорований тип розподілу може бути успішно модельований моделлю нейронних мереж, яка апроксимує плавне і безперервне генеративне перетворення з простору представлення у вихідний простір даних з контрольованою точністю *G* незалежно від розміру набору даних, що випливає з теореми про універсальну апроксимацію [69] (рис 2.2, верх). Отже, такий розподіл міг би задовольняти раніше названим умовам точності відтворення, спільності та стиснення при наборах даних будь-якого розміру.

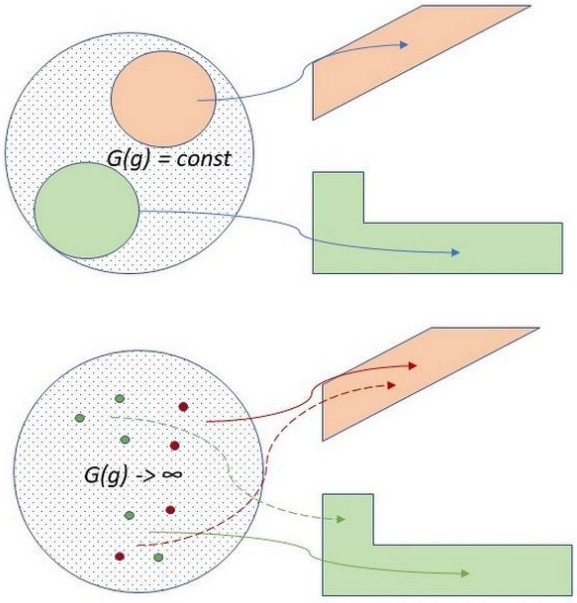


Рис. 2.2. Обмеження точності відтворення та узагальнення

З іншого боку, при другому типі розподілу збільшення розміру набору даних призведе до більшої щільності обох вибірок у латентній області перетину розподілів концептів, і з кожною ітерацією навчання генеративне перетворення має бути все більш різким, щоб забезпечити точне відображення у правильні області вхідного простору. При досить великому наборі даних такий розподіл неминуче призведе до конфлікту між обмеженнями точності відтворення та загальності,

внаслідок чого або точність не може бути збережена, або кількість параметрів моделі необхідно буде збільшувати без межі (рис. 2.2, внизу).

Ці аргументи підтверджують висновок, що категоризовані представлення є єдиними сумісними з вищезгаданими обмеженнями точності відтворення; загальності та стиснення у просторі представлень.

Іншим суттєвим наслідком доведеного ефекту категоризації для практичного застосування систем навчання є значне зменшення обсягу латентних даних при збереженні інформації про вхідний розподіл. Наприклад, у прикладі візуальних даних, що використовувалися в даній роботі, розмір зразка даних був зменшений з 1,024–4,096 до трьох числових параметрів, що означає стиснення обсягу даних у 300–1,000 разів, представляючи важливу і в деяких випадках критичну перевагу для систем з обмеженими ресурсами пам'яті та обчислювальної потужності.

2.3.3. Припущення та обмеження доказу

За доказом було зроблено низку неявних припущень, які будуть обґрунтовані в цьому розділі.

Перше – існування категорованих представлень. При доказі неявно передбачалося, що у просторі всіх можливих генеративних конфігурацій існує конфігурація з найкращою категоризацією, яку можна описати параметрами розподілів прихованих концептів. Вважаємо, що вищезгадані експериментальні результати, також емпіричний досвід із практиками розпізнавання осіб, малювання та абстрактних концепцій загалом забезпечують обґрунтування того, що такі представлення можуть існувати принаймні у деяких моделях генеративного навчання.

Друге важливе припущення – це значущість вхідних параметрів. Дійсно, вхідні параметри повинні бути релевантними і досить конкретними, щоб розрізняти зразки суттєво різних прихованих концептів. Розглянемо простий приклад: допустимо набір даних зображень особи, що використовується для навчання моделей глибокого навчання має тільки один параметр, наприклад, стать

людини. Зрозуміло, що єдина можлива категоризація в цьому випадку може бути за значенням цього параметра і точніша категоризація осіб була б неможливою.

Нарешті, важливим фактором є представництво прихованих концептів у наборі даних. Як показують результати досліджень, найвищий ступінь категоризації і відповідно можливості неконтрольованого навчання мають концепції з великим населенням у навчальному наборі даних. Штучно створені набори даних можуть порушити баланс між прихованими концептами, що може суттєво вплинути на навчання моделі та результуючу інформаційну структуру представлення.

2.3.4. Категоризація та здатність до узагальнення

Аргументи, представлені при доведенні теореми про категоризацію генеративних представлень можуть дати представлення про зв'язок між обмеженнями точності та загальності в моделях як контрольованого, так і неконтрольованого навчання. Зокрема можна поставити запитання: до якої межі можна стиснути даний набір даних, щоб результат навчання продовжував задовольняти вимогам точності відтворення та загальності?

Як випливає з доведеної теореми категоризації, стиснення простору параметрів даних із збереженням суттєвого інформаційного змісту веде до виникнення структури інформативних латентних представлень, пов'язаної з характерними типами (концептами) даних. Цей процес може продовжуватися аж до межі, коли розмірність простору представлення досягнє розмірності варіацій розподілів прихованих концептів зі значним змістом у вхідних даних, тобто кількості незалежних параметрів або ступенів свободи розподілів прихованих концептів H_k.

За цією межею подальше стиснення латентного простору з одночасною підтримкою обмежень точності відтворення та загальності не є можливим: примус моделі підтримувати точність призведе до надмірного припасування, тобто до буквального запам'ятовування-кодування навчальних даних, у той час як спроба використовувати інші набори аналогічних даних може знизити точність

класифікації в моделях контрольованого навчання, або точність відтворення вхідних даних у неконтрольованих моделях за рахунок того, що модель не зможе відобразити всі характерні варіації розподілів прихованих концептів.

Таким чином, можна припустити, що оптимальна ефективна розмірність простору представлення в моделях може бути обрана на основі очікуваного або спостережуваного спектра варіацій прихованих концептів у вхідних даних, які за деяких умов можуть бути апроксимовані відомими класами даних.

2.4. Методи аналізу розподілів даних у представленнях

2.4.1. Статистичні методи аналізу розподілів даних

Припустимо в k-мірному просторі R заданий розподіл даних X:

$$X = \{ X_l, l = \overline{1,N}\} = \{(x_{l,i}), i = 1,k;\ l = \overline{1,N}\}$$

Кластеризація $S(X)$ є розподілом даних X по непустому числу класів (кластерів) S_j , що задовольняють умові:

$$S_j = \{ X_i, i = \overline{1,N_j}\};\ X = \cup S_j;\ S_i \cap S_j = \emptyset \tag{2.12}$$

при неідеальній кластеризації умова вище може бути полегшена на наступне:

$$X = \cup S_j + N;\ \ \dim(S_i \cap S_j) \ll \dim(S_i, S_j)$$

де N – компонента випадкового шуму.

Якщо склад кластера S_j є відомим, то можна визначити статистичні параметри розподілу даних кластера, такі як центроїд і варіабельність (дисперсійно-коваріаційна матриця, [3]):

$$E^{(j)} = \{ \bar{X}_i, i = \overline{1,k}; X \in S_j \} \tag{2.13}$$

$$D_{i,l}^{(j)} = \{ cov(X_i, X_l), i, l = \overline{1,k}; X \in S_j \}$$

На підставі цих статистичних величин кластера S_j, а також параметрів загального розподілу даних X можна робити висновки про характер розподілу даних кластера, таких як компактність та інші. Для оцінки ступеня кластеризації добре відомі величини, що визначаються відомими функціоналами [3]:

1. Сума, у тому числі зважена, внутрішньокластерних дисперсій, обчислюється за формулою:

$$Q_1(S) = \sum_{j=1}^{M} \sum_{l=1}^{N_j} d^2\,(X_l^{(j)}, \bar{X}^{(j)})$$

де $\bar{X}^{(j)}$ – центр кластеру S_j.

2. Сума попарних внутрішньокластерних дистанцій:

$$Q_2(S) = \sum_{j=1}^{M} \sum_{l=1}^{N_j-1} \sum_{h=l+1}^{N_j} d(X_l^{(j)}, X_h^{(j)})$$

Зручність цієї величини полягає в тому, що її мінімізація забезпечує максимізацію суми міжкластерних дистанцій.

3. Загальна внутрішньокластерна дисперсія:

$$Q_3(S) = \det\left(\sum_{j=1}^{M} N_j\, V_j\right)$$

або:

$$\acute{Q}_3(S) = \prod_{j=1}^{M} \det\left(V_j\right)^{N_j}$$

де V_j – матриця коваріації кластеру S_j; $d(X, Y)$ – метрика у просторі представлення. У роботі евклідова метрика $d(X,Y) = \sqrt{\sum_{l=1}^{k}(X_l - Y_l)^2}$, але існують інші метрики – Міньковського, Чебишева, Манхеттен та інші.

Нарешті можна відзначити методи кластеризації за щільністю, що обговорювалися в розділі 1.2.2. Вони можуть використуватися для визначення структур щільності в просторах генеративних представлень при повністю неконтрольованому навчанні.

Результатом застосування методу кластеризації за щільністю P є розбиття області розподілу даних у просторі представлення R на множину кластерів щільності $D = \{ D_j \}$ з можливістю визначення належності вхідного зразка $x \in X$ до визначеного кластера D_k:

$$D_k = P(R) = P(E(X)) \quad (2.14)$$

де *E(x)* – перетворення кодування з вхідного простору X в простір представлення, (1.2).

При визначенні кластерної структури щільності, методи аналізу розподілу даних, описані вище в цьому розділі, можуть бути повністю застосовні до кластеризації за щільністю. Така кластерна структура отримана в результаті застосування методів неконтрольованої кластеризації називатиметься далі ландшафт щільності представлення.

2.4.2. Контрольовані та неконтрольовані методи аналізу розподілу даних

Існують два принципово різних напрями визначення структури розподілів даних у просторах представлень: контрольоване, тобто з використанням представницьких наборів відомих концептів; і неконтрольоване, що не требує знання асоціації даних із відомими концептами. Відповідно, використовуються два типи вибірок у просторах представлень:

1. Вибірки категорій, тобто набори X_j відомих зразків цієї категорії C_j перетворених у простір представлення перетворенням кодування E (2.2): $r_j = E(X_j)$.

2. Загальна вибірка, тобто набір немаркованих даних Y у вхідному просторі, який може використовуватися для вимірювання розміру, форми та інших параметрів розподілу даних $E(Y)$ у просторі представлення, а також неконтрольованої кластеризації та визначення генеративного ландшафту, як обговорювалося вище. Типовий розмір загальної вибірки у роботі становив 0.1 – 0.2 від набору даних із чисельності.

Далі наводиться короткий опис методів і підходів контрольованого та неконтрольованого аналізу латентних розподілів які дозволяють визначити загальну картину та детальні параметри розподілів даних у просторах генеративних представлень.

Контрольовані методи аналізу розподілів. Аналіз латентних розподілів наборів відомих категорій X_j може включати:

1. Визначення розмірів латентних вибірок категорій $r_{j.}$

2. Визначення статистичних параметрів латентних розподілів вибірок категорій, таких як середнє, варіація та вищі моменти розподілу.

3. Кореляційний та коваріційний аналіз латентних вибірок категорій.

4. Аналіз латентних розподілів категорій методами багатовимірних гістограм.

Параметри, що характеризують розподіл латентних вибірок категорій можуть включати:

1) розмір, в координатах простору представлення, відносно розміру загальної вибірки*: E(Y)*: $D(C_j)_l = Max\left(E(X_j)\right)_l \;/\; Max(E(Y))_l$;

2) форму, форма області розподілу категорії виявлена при візуалізації розподілу у просторі представлення;

3) розповсюдження (Spread), або відносний обсяг вибірки категорії до обсягу загальної вибірки: $S_{pr}(C_j) = Vol\,(E(X_j)) \;/\; Vol\,(E(Y))$.

4) структуру – визначається як число компонентів, кластерів у розподілі латентної вибірки категорії, які можна визначити візуально або іншими методами: $S_{tr}(C_j) = Card(E(X_j)) \,/\, Card(E(Y))$;

5) щільність (Density), характерна щільність даних у латентної області розподілу категорії. Розраховується як відношення кількості зразків у вибірці категорії та її обсягу: $D(C_j) = Card(E(C_j)) \,/\, Vol(E(C_j))$.

Неконтрольовані методи аналізу розподілу. Аналіз латентного розподілу загального набору *r = E(Y)* може включати:

1. Визначення розмірів та форми латентної області розподілу загальної вибірки.

2. Визначення статистичних параметрів латентного розподілу загальної вибірки, такі як середнє, варіація та вищі моменти розподілу.

3. Визначення характеристик розподілу щільності загального набору в просторі представлення, діапазону варіації щільності та регіонів з максимальною та мінімальною щільністю.

4. Аналіз структури латентного розподілу загальної вибірки за допомогою методів багатовимірних гістограм.

5. Визначення структури щільності латентного розподілу загальної вибірки методами неконтрольованої кластеризації з виділенням основних кластерів щільності.

6. Характеристики розподілу загальної вибірки за визначеними основними структурами/кластерами щільності. Може використовуватися параметр концентрації C_{on}, визначається як частка загальної вибірки в основних структурах / кластерах щільності та інші.

Вимірювання характеристик та аналіз розподілів загальної вибірки та вибірок категорій дає можливість порівняти ефект категоризації для різних типів даних та дослідити зв'язок між інформаційною структурою латентних розподілів та відомими класами даних.

2.5. Теоретичне обґрунтування методів та моделей роботи

2.5.1. Теоретичне обґрунтування архітектури нейронних мереж автоенкодера зі стисненням шару кодування

Результати по спонтанній категоризації дозволяють запропонувати простий, але ефективний тип моделі нейронних мереж глибокого автоенкодера з суттєвим "фізичним" стиском центрального шару (тобто зниження розмірності центрального шару, що створює латентне представлення порівняно з розмірністю вхідних даних).

У пропонованій нейромережевій архітектурі істотне стиснення інформації досягається послідовністю шарів розширення і кодування, зі значним перепадом розмірностей. Перша група шарів розширення дозволяє досягти точності кодування вихідних даних, у той час як значне скорочення розмірності центрального шару створює латентне представлення сприяє виділенню прихованих концептів даних. Схема архітектури моделей наведена в розділі 1, який містить детальний опис моделей що використовувався в роботі.

Завдяки наявності додаткових шарів і спеціальних функцій активації та обробки, таких як пошарова нормалізація даних, такі моделі мають достатню

глибину для моделювання складних типів даних зі значним числом вхідних параметрів. Також вони мають якості, які роблять їх сумісними зі висновками теореми про категоризацію генеративних представлень (розділ 2):

1. Моделі автоенкодера є моделями генеративного навчання, що створюють інформативні представлення даних навчання.

2. Навчання проводиться методом зниження за стохастичним градієнтом сумісного з принципом енергетичного навчання.

3. Стиснення простору вхідних даних досягається обмеженням розмірності центрального шару моделі.

На додаток, наявність фізичного шару кодування значно полегшує дослідження параметрів розподілів у генеративних представленнях створюваних моделями у процесі неконтрольованого генеративного навчання, включаючи безпосередню візуалізацію розподілів у просторах представлень.

2.5.2. Теоретичне обґрунтування методів навчання з використанням структури щільності генеративних представлень

Результати теоретичної частини роботи, включаючи теорему про категоризацію генеративних представлень, дозволяють запропонувати та обґрунтувати методи навчання моделей з використанням структури генеративних представлень, яка виникає в результаті неконтрольованого навчання генеративних моделей (генеративного ландшафту представлень). Наслідки теореми дозволяють просунутися у розумінні деяких суттєвих питань, наприклад, якою мірою можна стискати інформацію у просторі генеративного представлення, але вона недостатня, щоб перейти безпосередньо до задачі неконтрольованого розпізнавання концептів, оскільки області та параметри розподілу концептів можуть бути невідомі.

Для подальшого просування необхідно зробити додаткові припущення, а саме: що простір вхідних даних представлений навчальним набором містить кінцеве число основних концептів зі значним представництвом у наборах, які використовуються для генеративного навчання. Це припущення виправдане у

випадках роботи, таких як дані Інтернет або спеціалізовані набори зображень. Це ключове припущення дозволяє очікувати, на підставі теореми категоризації, що простір представлень успішних моделей генеративного самонавчання може мати виражену структуру щільності, з областями підвищеної концентрації даних у районах відповідних латентним регіонам розподілів характерних типів даних (концептів) у просторі представлення:

$$max\, D(y) \sim H_k(y)\,,$$

де $max(D(y))$ – локальні максимуми функції щільності $D(y)$ розподілу даних у просторі представлення; H_k – області розподілу прихованих концептів.

Вищезазначені аргументи дозволяють визначити інформативну структуру представлень – генеративний ландшафт L сформований в результаті успішного генеративного навчання та застосування методів неконтрольованої кластеризації у просторі представлень, згідно висновками теореми категоризації. Існують відомі методи кластеризації, такі як кластеризація за щільністю (DbScan, Optics, MeanShift), методи багатовимірних гістограм та інші, що дозволяють аналізувати структури щільності в багатовимірних просторах без відомих зразків концептів, тобто в повністю неконтрольованому режимі:

$$L(D) = \{\, d_j \,\}, j = \overline{1,n} \;\; = K_d(E(D)),$$

де $\{\ d_j\ \}$ – структури щільності структури щільності розподілу даних, такі як кластери щільності у просторі генеративного представлення; K_d – алгоритм кластеризації за щільністю; E – перетворення кодування.

Методи кластеризації за щільністю для даних D можна визначити як (метод DbScan):

$$n_{\varepsilon}(p) = \{\, q \;\in D \mid dist(p,q) \;\leq\; \varepsilon \,\}$$

$$C_{\boldsymbol{\varepsilon},P_t} = \{\, p \;\in D \mid n_{\varepsilon}(p) \;\geq\; P_t \,\},$$

де ε – радіус околиці (параметр методу); $n_{\varepsilon}(p)$ – функція чисельності околиці; $C_{\varepsilon,\mathrm{Pt}}$ – ядра щільності.

При досить репрезентативному загальному наборі D, ґрунтуючись на припущенні загальності навчання, можна очікувати що генеративний ландшафт представлення $L(D)$ описує загальну структуру вхідних даних. Необхідно

зазначити, що ні процес генеративного самонавчання моделі, ні визначення генеративної структури представлення не вимагають відомих даних концептів і таким чином, є повністю неконтрольованими.

На основі визначеної структури щільності (ландшафту) представлень, навіть при мінімальних наборах відомих даних, можна сформувати навчальні набори класів методом вибірки:

$$d_p, d_n(D_c) = \{\, d \in L \mid E(D_c) \in d_p \,\}, \{\, d \in L \mid d \neq d\_p \,\}$$

$$(T_p, T_n) = \left(\{\, y \in d_p \,\}, \{\, y \in d_n \,\} \right)$$

де D_c – набір (зразок) відомих представників класу; d_p, d_n – структури ландшафту, відповідні розподілу класу та даним поза класом; T_p, T_n – набори зразків класу і поза класом у просторі представлення.

Методи навчання на генеративному ландшафті щільності використовують кластерну структуру, таку як кластеризація за щільністю для навчання розпізнавання відомих концептів при мінімальних даних навчання або в повністю неконтрольованому режимі, без навчальних даних.

Обидва різновиди методу використовують кластерну структуру щільності загального латентного розподілу $D = \{ D_j \}$, яку можна визначити методами, що обговорювалися в попередніх розділах.

На першому етапі, повністю неконтрольованому, з урахуванням кластерної структури щільності D визначається структура “натуральних концептів” $H = \{H_k, k=1,m\}$ у латентному розподілі загального набору наступним процесом:

1. Вибирається домінантний кластер K_1 в D, наприклад, за чисельністю (тобто максимальною концентрацією даних загального набору).

2. Створюється вибірка елементів $P_{K1} \in K_1$ яка представляє клас концепта H_1.

3. Створюється вибірка елементів N_{K1} із елементів інших домінантних кластерів $K_j \subset D, K_j \neq K_1$.

4. На маркованому наборі (P_{K1}, N_{K1}) з бінарним маркуванням (True, False) навчається бінарний класифікатор B_1 звичайного типу, наприклад, типу найближчого сусіда (kNN) [95], нейромережевий, опорних векторів (SVM) [96] або інших типів. Важливо відзначити, що хоча навчання класифікаторів виробляється

контрольованим процесом, вони не вимагають даних навчання маркованих заздалегідь відомими класами.

5. Процес повторюється для наступного домінантного кластера і так далі, до межі мінімальної чисельності.

Отриманий в результаті процесу набір класифікаторів визначає структуру «натуральних концептів» у просторі генеративного представлення:

$$K = \{H_j,\ B_j\} \tag{2.15}$$

Слід зауважити, що визначена таким чином концептуальна структура в просторі представлень дозволяє співвіднести вхідний зразок з його натуральним концептом:

$$x \in R \rightarrow H_j = B_j(E(x)) \tag{2.16}$$

з визначенням концепту $H_j(x)$ з результатів класифікаторів за умовою вищої впевненості чи аналогічною.

Відповідно до висновків теореми про категоризацію генеративних представлень, визначена таким процесом латентна структура повинна відповідати характерним типам вхідного розподілу. Це обґрунтовує підхід навчання розпізнаванню натуральних концептів без вимоги маркованих наборів відомих класів.

У різновиді методу навчання на генеративному ландшафті з мінімальними даними навчання, потрібна невелика кількість зразків $X_j = \{x^{(j)}_k,\ k = 1, \text{n}\}$ відомого класу C_j у вхідному просторі. На відміну від традиційних методів контрольованого навчання можлива мінімальна кількість зразків, до єдиного, $n = 1$ [97]; також важливо відзначити, що метод не вимагає зразків всіх відомих концептів відразу і може навчатися ітеративно в міру доступності навчальних даних.

Далі, латентні образи набору X_j у просторі представлення $e_j = E(X_j)$ можуть бути асоційовані з домінантними кластерами генеративного ландшафту D, створюються позитивна та негативна вибірки та проводиться навчання класифікатора відомого класу C_j, аналогічно до описаного вище процесу неконтрольованого методу визначення натуральних концептів. Обґрунтування цього методу залежить від суттєвого припущення про сильний зв'язок між

відомими типами даних (класами) та головними натуральними концептами інформативних представлень. Воно обґрунтоване у разі даних пакетів трафіку Інтернет завдяки суттєвій відмінності мережевої поведінки додатків Інтернету, і можливо, задовольняється для інших типів даних.

Необхідно відзначити, що запропоновані методи навчання на генеративному ландшафті по суті ітеративні, тобто дозволяють поліпшити результат навчання послідовністю ітерацій в процесі навчання. В результаті емпіричного досвіду початковий навчальний набір X_j може бути розширений новими зразками ($X_j(1)$ і так далі), що покращує точність класифікаторів класів.

Зведення обґрунтованих у цьому розділі методів навчання на генеративному ландшафті наведено в табл. 2.1.

Таблиця 2.1

Методи навчання з використанням генеративного ландшафту представлень

Метод	Мета	Вхідні дані	Результат
Визначення генеративної структури (ландшафту) представлень даних Інтернет	Визначення генеративної структури даних представлень	Загальна вибірка сенсорних даних	Структура латентних кластерів пов'язаних із натуральними концептами
Неконтрольоване навчання розпізнаванню натуральних концептів даних Інтернет	Визначення характерних типів (концептів) вхідних даних	Генеративна структура (ландшафт) представлення	Структура натуральних концептів даних; класифікатори концептів
Ітеративне навчання розпізнаванню відомих класів даних Інтернет	Навчання розпізнаванню відомих класів з мінімальними наборами даних	Ландшафт представлення, мінімальні набори класів	Класифікатори відомих класів; необхідний рівень точності та впевненості класифікації

2.6. Зв'язок з теоріями та методами навчальних систем

2.6.1. Зв'язок з теоріями неконтрольованого та напівконтрольованого навчання

Висновки теоретичної частини роботи засновані на, і тісно пов'язані з раніше отриманими результатами в теорії систем неконтрольованого навчання. Суттєвою частиною та основою отриманих результатів із категоризації генеративних

представлень послужили підходи та висновки теорії неконтрольованого генеративного навчання, теорії представлень і енергетичного навчання.

Результати, отримані за допомогою моделей побудованих на принципах і результатах теорії неконтрольованих систем дають цікаві і важливі приклади категоризації за концептами сенсорних даних, що виникає при неконтрольованому навчанні генеративних систем з різними типами даних і архітектур. Експериментальні результати наведені в розділі 4, а також отримані раніше (розділ 2.2.4, та інші) підтверджують висновки теоретичної частини роботи про закономірний характер виникнення категоризованих представлень генеративних моделей за зазначених умов.

Результати, отримані в теоретичній частині роботи, також мають суттєві зв'язки з теорією напівконтрольованого навчання. Наприклад, при доказі теореми про категоризацію генеративних представлень суттєво застосовувалися припущення безперервності, кластеризації та розподілу прихованих концептів, які широко використовуються в теорії напівконтрольованого навчання [49].

Інша причина інтересу до підходів та методів напівконтрольованого навчання застосування біологічними системами, які здатні вчитися швидко та ефективно з мінімальною кількістю даних та дослідів. Наприклад, значна частина завдань навчання людини включає невелику кількість даних істини таких як пряме навчання, у поєднанні з великою кількістю неконтрольованого досвіду (наприклад, спостереження за об'єктами в навколишньому середовищі без додаткового знання про їх значення або природу). Так, існують експериментальні підтвердження, що діти чутливі до зображень об'єктів, що часто зустрічаються в їхньому оточенні: зображення собак, кішок і осіб дорослих людей [98].

Новизною отриманих результатів та розроблених методів є, по-перше, виділення суттєвих інформаційних структур щільності в генеративних представлень повністю неконтрольованими методами, без вимог відомих даних а також розробка методів ітеративного навчання на основі визначеної структури (ландшафту) генеративних представлень. Застосування цих методів дозволило

суттєво підвищити ефективності та гнучкість методів навчання як показано в експериментальних роботи, навіть із мінімальними наборами навчальних даних.

2.6.2. Зв'язок з нейробіологічними моделями навчальних систем

Багато ключових поступів у сфері систем штучного інтелекту ґрунтувалися на моделях і принципах навчання біологічних систем [76]. У контексті теоретичних та експериментальних результатів роботи можна провести цікаві паралелі зі стратегіями обробки інформації біологічними системами. Висока точність відтворення даних при їх суттєвому стисканні може мати прямі та очевидні переваги для біологічних систем в ефективному використанні обмежених ресурсів обробки інформації та пам'яті, правильної інтерпретації та реакції на зміни у навколишньому середовищі, які можуть мати вирішальне значення для виживання.

Можна помітити, що системи здатні до ефективного навчання повинні задовольняти всім обмеженням теореми категоризації, а саме, точності відтворення зовнішньої інформації, від якої прямо залежить ефективність відповіді на зовнішній стимул; здатності до узагальнення що забезпечує стабільність кореляції стимулу та відповіді; та суттєве стиснення інформації у зв'язку з обмеженими фізичними ресурсами системи. З цієї причини виникнення категоризованих представлень у таких системах можна пояснити ефектом категоризації обґрунтованим у цій частині роботи.

При такому підході структура інформативних представлень сенсорних даних, отримана в ході навчання без вчителя послідовністю дослідів взаємодії з навколишнім середовищем, відіграє роль основи для ітеративного навчання процесом проб і помилок, в результаті якого виробляється більш точне розпізнавання суттєвих для системи сигналів довкілля. Ці результати можуть вказати цікаві напрямки для подальших досліджень у галузі нейробіології та інформаційних принципів навчання як біологічних так і штучних систем.

В результаті теоретичного обґрунтування методів та моделей неконтрольованого генеративного навчання, створення інформативних низькорозмірних представлень даних пакетів трафіку Інтернет та визначення

інформативної структури представлень, проведеного в цьому розділі, виникає можливість використання інформативної структури генеративних представлень для розробки методів навчання зі значно зменшеними вимогами до наявності заздалегідь відомих даних навчання.

РОЗДІЛ 3.
ДОСЛІДЖЕННЯ АРХІТЕКУТРИ, МОДЕЛЕЙ ТА МЕТОДІВ

У даному розділі дисертаційної роботи наведений детальний опис моделей, даних та методів, використаних в роботі.

3.1. Модель глибокого автоенкодера з різким зниженням розмірності представлення

Архітектурна модель, використана в дослідженні для отримання інформативних представлень даних пакетів трафіку Інтернет, а також інших типів, представляє собою нейронну мережу глибокого автоенкодера з майже симетричною структурою, з різким стиском центрального (кодуючого) шару.

Модель складалася з трьох або більше основних компонентів, включаючи: вхідний та вихідний шари ідентичної розмірності рівної розмірності вхідних даних; центрального шару кодування; та глибоких прихованих шарів з одного або декількох шарів нейронів для придбання інформаційних параметрів даних (features). При використанні даних зображень використовувалися додаткові шари адаптації для придбання параметрів масштабу, таки як шари згортки та розширення стандартні в роботі з даними зображень.

Структуру архітектури моделей глибокого автоенкодера зі стисненням кодуючого шару показано на рис. 3.1.

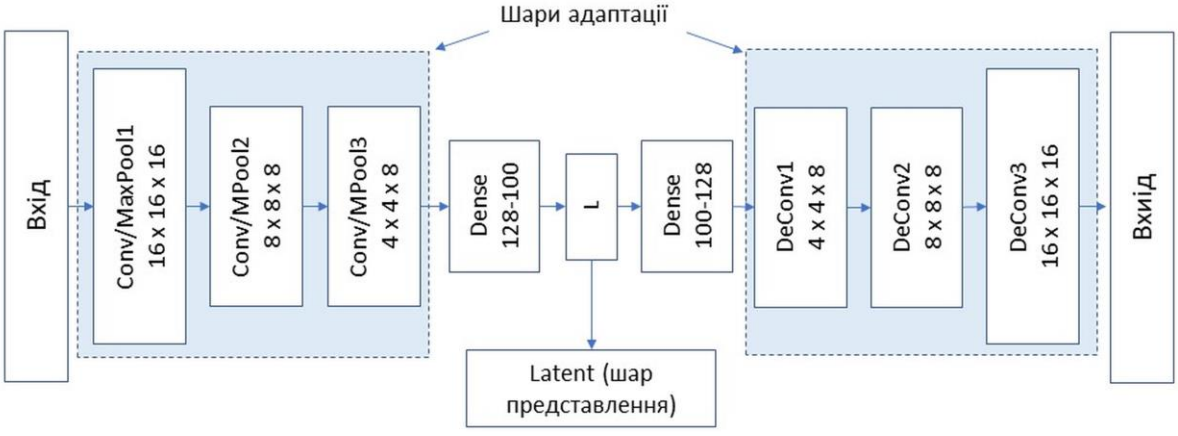


Рис. 3.1. Архітектурна діаграма моделі глибокого автоенкодера

У представленій архітектурі симетричні шари "Dense" є прихованими шарами моделі, в яких здійснюється виділення інформативних параметрів подання

вихідних даних при навчанні на вибірці немаркованих даних з мінімізацією помилки відтворення. У моделі використовувалися один або кілька шарів з ефективними функціями активації та погруповою обробкою даних, таких як погрупова нормалізація. Кодування вихідних даних створювалося активацією нейронів центрального шару моделі ("Latent").

Характерною особливістю розроблених і використовуваних в експериментальній частині роботи моделей глибокого автоенкодера є суттєве стиснення розмірності простору подання в 1-3 десяткових порядку величини в порівнянні з вхідним шаром. Така структура виправдана щодо зв'язку принципу мінімізації вільної енергії з ефективністю категоризації при неконтрольованому навчанні з відтворенням даних (розділ 2).

Загалом, залежно від значень гіперпараметрів, таких як розмірності прихованих шарів, моделі використані в роботі мали від 8,000 (дані Інтернет) до 96,000 (дані зображень) параметрів навчання, як описано в табл. 3.1. Детальна діаграма архітектури моделей використовуваних у роботі наведена в Додатку А.

Моделі були виконані програмною мовою Python з використанням пакетів моделювання нейронних мереж Keras та Tensorflow [87]. Також використовувалися загальні програмні пакети обробки, перетворення, подання та візуалізації даних, такі як: numpy, scikit, sklearn-kit, matplotlib [99] та інші.

Таблиця 3.1

Параметри архітектури нейромережевих моделей роботи

Шар	**Розмірність**	**Інтервал значень**	**Функція активації**	**Функція вартості (cost function)**
Вхід	F = 22 .. 30 (набор Интернет) 32 × 32 (набор изображений)	[0 .. 1]		-
Dense	M = 10 .. 50	[0, ∞[	Relu	-
Latent	L = 3 .. 10	]-∞, +∞[	Leaky Relu	-
Вихід	O = F	[0 .. 1]	Sigmoid	MSE, CCE (1)

3.2. Обґрунтування вибору моделі

Відповідно до результатів, отриманих у теоретичній частині дисертаційної роботи, успішні моделі генеративного навчання, до яких належить описана вище архітектура автоенкодера, можуть створювати в процесі неконтрольованого навчання структуровані або категоризовані представлення вхідних даних. Вибір моделі глибокого автоенкодера, як засіб для створення структурованого та відповідно до доведених результатів, корелюваного з концепціями вхідних даних представлень ґрунтується на наступних аргументах:

1. Будучи універсальним апроксиматором, нейронні мережі можуть мати практично необмежену універсальність та підходять для моделювання складних типів даних, включаючи дані Інтернет, зображень та інших типів.

2. Ефект виникнення інформативних структур, корельованих відомими класами даних у моделях глибоких нейронних мереж та автоенкодерів був описаний раніше в експериментальних роботах зі штучними та реальними даними [34, 35, 90].

3. Глибокі нейронні мережі широко представлені в біологічних системах, які також відрізняються здатністю до ефективного навчання у широкому спектрі ситуацій при мінімумі заданих даних [75, 76, 77].

4. Результати останніх років [35, 36] та інші підтверджують високу ефективність генеративних моделей глибоких мереж у створенні інформативних представлень сенсорних даних у порівнянні з більш ранніми моделями, можливо, завдяки їх практично необмеженій глибині і відповідно можливостях апроксимації навіть найбільш складних даних.

Ґрунтуючись на цих твердженнях, можна очікувати, що моделі нейронних мереж глибокого автоенкодера можуть бути гарною відправною точкою для вивчення неконтрольованої категоризації представлень і методів навчання на основі неконтрольованої генеративної структури інформативних представлень.

3.3. Вибір параметрів моделі

Важливими гіперпараметрами моделей нейронних мереж у роботі були: розмір прихованих шарів D_h, та шару кодування, D_{enc}. Було здійснено пошук по

решітці значень цих параметрів при D_h в інтервалі 10 .. 100; D_{enc} 3 .. 10 за критеріями ефективності навчання.

Результати пошуку представлені в табл. 3.2, де тип моделі позначений як < розмірність прихованого шару >-< розмірність кодуючого шару >, наприклад "50-3".

Таблиця 3.2

Вибір гіперпараметрів архітектури (набір даних Інтернет)

Модель	Функція втрат, вихід	Точність навчання, вихід	Кількість кластерів щільності	Розміри набору (мін, макс)
50-5	6×10^{-4}	96.8%	47	0.015 / 0.022
25-5	9×10^{-4}	96.6%	46	0.015 / 0.023
10-5	0.001	96.4%	41	0.016 / 0.026
50-3	9×10^{-4}	96.2%	36	0.019 / 0.032
25-3	0.0011	95.7%	34	0.019 / 0.034

*Результати усереднені за набором 20 сесій навчання зменшення ймовірності стататистичної флуктуації. Ступінь упевненості 0.95 при інтервалі ±0.3%.

З наведених результатів можна зробити висновок, що оптимальна розмірність шару представлення у разі даних Інтернет може бути у інтервалі 3 – 5, що також підтверджується даними аналізу головних компонентів. Також можна помітити, що зміна розмірності прихованого рівня у значному інтервалі 25–50 не викликає суттєвої зміни ефективності навчання на наборі даних Інтернет.

У більшості наступних екпериментів розмірність шару представлення була обрана рівною трьом для спрощення вимірювання параметрів розподілів та можливості безпосередньої візуалізації розподілів у просторі представлення програмними пакетами об'ємної візуалізації.

3.4. Набори даних для навчання генеративних моделей та створення генеративних представлень

Для підтвердження загальності встановлених результатів та виключення артефактів конкретного набору або типу даних, у роботі використовувалося два набори даних суттєво різних за характером та типом.

Набір даних пакетів трафіку Інтернет. Дані в дослідженні є інтернет-трафіком, який був отриманий із записів громадської телекомунікаційної мережі в Новій Зеландії. Записи пакетів трафіку Інтернет були отримані з сервера WITS університету Вайкато, Нова Зеландія [100].

Записи пакетів протоколу Інтернет (IP протокол) були зібрані в сесії з комбінації пари (інтернет-адреса, порт) джерела та призначення, які були елементами немаркованого набору даних, що використовувалося у роботі. Параметрами інтернет-сесій набору були статистичні характеристики розподілу розмірів пакетів та інтервалу між пакетами в сесіях. Таким чином, кожен зразок набору представляв екземпляр інтернет-сесії, такий як телефонна розмова, перегляд веб-сторінок, сеанс обміну миттєвими повідомленнями, завантаження файлів тощо, і визначався вектором 22-30 числових параметрів статистик розподілу пакетів сесії за розміром та тимчасовому інтервалу [101], як зазначено в табл. 3.3.

Таблиця 3.3

Параметри набору даних пакетів трафіку Інтернет

Тип параметра	Кількість параметрів	Опис	Тип даних	Попередня обробка
Загальні	6	загальна тривалість сесії, підсумковий розмір даних сесії (за напрямками), кількість пакетів (за напрямками), протокол даних		Масштабування до інтервалу [0,1]
Статистика розміру пакетів	8 -12	мін, макс, середнє, стандартне відхилення, ентропія розміру пакетів, за напрямками		Масштабування, [0,1]
Статистика часового інтервалу	8 -12	мін, макс, середнє, стандартне відхилення, ентропія часового інтервалу, за напрямками		Масштабування, [0,1]

Попередня обробка. Набір даних трафіку Інтернет був оброблений перед використанням для навчання моделей наступним чином:

1. Видалено сесії з мінімальним набором пакетів, тобто менше 2 по одному з напрямків, або менше 3 у сумі по сесії. Обсяг інформації у таких сесіях є мінімальним, кількість значних параметрів скорочується до одного (розмір пакета),

використання таких даних у навчанні немає сенсу. Таке відсівання може розглядатися як видалення фону.

2. Числові значення за всіма параметрами наведено до інтервалу (0, 1) (масштабування).

Таким чином, оброблений для навчання набір даних Інтернет містить вектори сесій Інтернет розмірності 22-30, з раціональними значеннями в інтервалі (0, 1).

Одним із завдань роботи було дослідження розподілу даних у просторі генеративних представлених даних за типом інтернет-додатку сесій – елементів набору. Для розв'язання цієї задачі частина набору була маркована класами додатків найбільш представлених у наборі на основі добре відомого порту, який є параметром Інтернет-протоколу (табл. 3.4).

Таблиця 3.4

Програми, додатки у наборі даних Інтернет

Додаток Інтернет	**Тип**	**Чисельність набору додатку**	**Примітка**
DNS	мережевий протокол	2500	
NTP	мережевий протокол	1500	
Telnet	віддалена сесія	500	
Messenger (MS)	месенджер	1300	дані протоколу та змісту не розділені
Messenger (Orbit)	месенджер	600	дані протоколу та змісту не розділені
Gmail	емейл	1000	дані протоколу та змісту не розділені
XBox (MS)	онлайн гра	2200	
Newton (Escale)	онлайн гра	300	
Streaming	стріминг	300	
Bit Torrent	сервіс файлів	2600	дані протоколу та змісту не розділені
Web (HTTP)	Гіпертекст протокол	5500	Дані, марковані протоколом гіпертекст (HTTP), який може містити різні додатки.

Статистичні характеристики набору даних Інтернет

Таблиця 3.5

Статистичні характеристики набору пакетів трафіку Інтернет

Характеристика	Значення	Примітка
Розмір масиву (кількість елементів набору)	130,000	исключены образцы с минимальным числом пакетов
Число маркованих класів	12	класи за додатком Інтернет сесії
Число параметрів (розмірність вхідних даних)	22 - 30	статистичні параметри розподілу розміру та часових інтервалів пакетів сесії, див. Таблицю 3.3
Середнє значення параметрів, діапазон	2.7×10^{-4} – 0.522	
Варіація параметрів, діапазон	0.006 - 0.383	
Головні компоненти (відносний вклад вище 0.1 варіації)	5	
Головні компоненти (відносний вклад вище 0.01 варіації)	14	
Розподіли параметрів	переважають одно- та двохмодальні розподіли з довгим хвостом	
Відносна емпірична ентропія *	0.41	

(*) Емпірична та гранична ентропія розраховані за допомогою гістограми в 1 мільйон кошиків, після чого розраховане значення досягло плато і значно не змінювалося при подальшому збільшенні кількості кошиків.

Приклади розподілу параметрів у вихідному просторі наведено на графіках нижче. Зліва наведено гістограму розподілу загального зразка, праворуч, аппрохімацію розподілу методом ядерної щільності. При застосуванні методів ядерної щільності використовувалися функції ядра Гауссовська та Єпанечників.

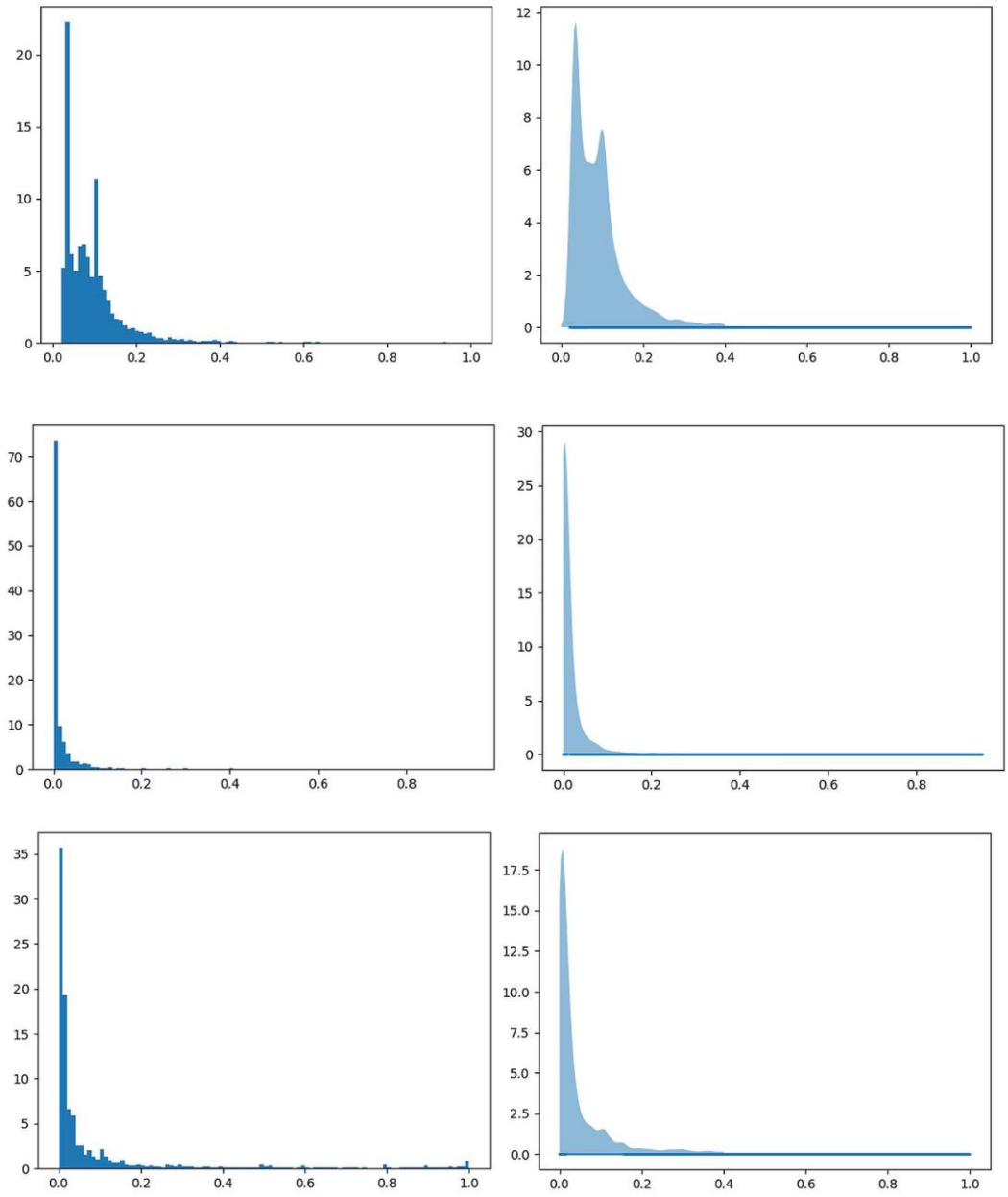


Рис. 3.2. Розподіли параметрів середнього розміру пакета, середнього міжпакетного інтервалу та повного розміру сесії (верх - низ) набору даних трафіку Інтернет

Будучи “живим” записом трафіку у великій громадській мережі Інтернет дані мають широке утримання різних додатків, у наборах даних було зазначено понад 400. З цієї причини видається, що цей набір добре підходить для перевірки правильності розробленого підходу з даними значного різноманіття та варіації.

Набір даних зображень. Другий набір даних, використаний у дослідженні, складався з масиву приблизно 1,100 зображень місцевості та ландшафту, які були

поділені на одинадцять маркованих класів, включаючи Будівлі; Ліс; Поле; Вода; Дороги; Транспортні сліди; Транспорт та інші [102]. Середня кількість зразків класу знаходилася в області 100. Хоча частина зображень набору була маркована відомими класами, вони жодним чином не використовувалися при навчанні неконтрольованих моделей, а лише при вимірюванні параметрів та візуалізації розподілів класів.

Приклади зображень набору наведені на рис. 3.3.

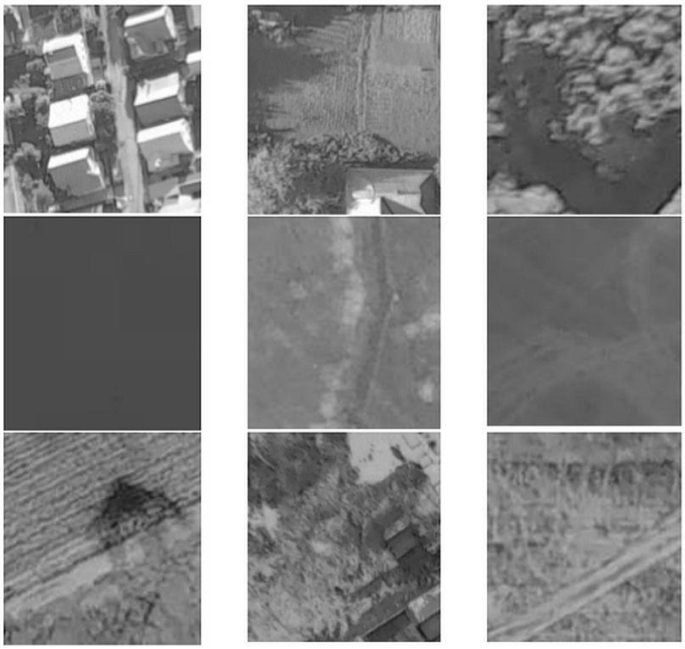


Рис. 3.3. Приклади даних набору зображень

Статистичні характеристики набору даних зображень. Дані набору зображень оброблені етапами згортки / розширення вектор числових значень в інтервалі (0, 1) розмірності 576. Набір статистичних характеристик наведений у табл. 3.6.

Таблиця 3.6

Статистичні характеристики набору даних зображень

Параметр	Значення	Примітка
Розмір масиву	1,100	
Число маркованих класів	9-11	Класи типів ландшафту та місцевості
Число параметрів (розмірність вхідних даних)	576	
Середнє значення параметрів, діапазон	0.139 - 0.757	
Варіація параметрів, діапазон	0.061 - 0.223	
Головні компоненти (сумарна варіація ~ 0.5)	3 / 576	
Головні компоненти (сумарна варіація вище 0.9)	100 / 576	
Розподіли параметрів	переважають одно- та двохмодальні Гауссовские распределения	
Відносна емпірична ентропія	0.116	

Приклади розподілів параметрів у вхідному просторі представлені на рис. 3.4 (ліворуч - гістограма, праворуч наближення методом ядерної щільності).

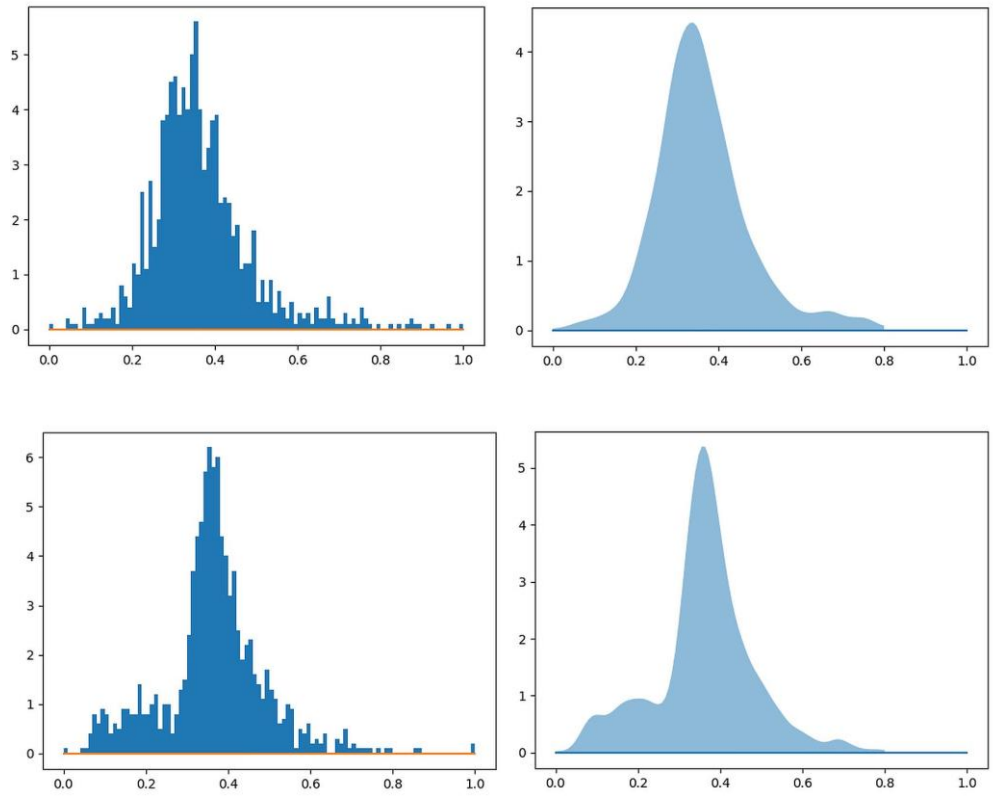


Рис. 3.4. Характеристичні розподіли значень параметрів, набір даних зображень

Дані зображення є більш складним типом даних, ніж дані Інтернет, і відтворення результатів дослідження з даними іншого типу і характеру має важливе значення для усунення можливості побічних ефектів і артефактів специфічних для конкерного набору даних, і підтвердження загальної застосовності результатів теоретичної та експериментальної частин роботи.

3.5. Неконтрольоване генеративне навчання

Навчання моделей проводилося як автоматичного генеративного неконтрольованого процесу навчання, метою якого є досягнення гарного відтворення вхідних даних. У навчанні використовувався стандартний у сучасних моделях штучних нейронних мереж метод зниження за стохастичним градієнтом (ЗСГ, SGD).

Оскільки метою навчання генеративних систем є точність відтворення вхідного розподілу, а не точність класифікації, в набір заздалегідь визначених класів, при навчанні моделей на немаркованих масивах даних у роботі використовувалися функції втрат середнього квадратичного відхилення (СКВ, MSE), яка вимірює усереднене по навчальному набору відхилення вектора відтворення евклідовій метриці простору сходових даних та категоричної крос-ентропії (ККЕ, CCE).

При неконтрольованому навчанні використовувалися кілька показників прогресу та успіху генеративного навчання:

– зміна значення функції втрат упродовж сесії неконтрольованого навчання;

– зміна значення функції точності протягом сесії неконтрольованого навчання; функція точності у разі неконтрольованого навчання визначається як збіг значень індексів максимального значення у вхідному та вихідному зразку, тобто міра близькості пов'язана з коваріацією даних входу та виходу;

– кореляція та ковариація між наборами у вхідних даних та результатом генерованим моделлю.

Результати ефективності неконтрольованого навчання моделей на масивах даних Інтернет та зображень, представлені в табл. 3.7. Критерієм закінчення

навчання було досягнення плато, тобто постійного значення функції втрат на верифікаційному наборі даних.

Таблиця 3.7

Результати неконтрольованого навчання, дані Інтернет та зображень

Набір даних	**Число епох навчання**	**Функція втрат, вхід/вихід**	**Точність, вхід/вихід**
Інтернет	20 – 50	$0.01 / 9 \times 10^{-4}$	0.1 / 0.95 (1)
Набір зображень	50	0.96 / 0.0095	0.00 / 0.31(2)

(1) При випадковій генерації ймовірність збігу вихідного і відтвореного зразка була б 1/30 ~ 0.03, тобто більш ніж у 30 разів гірше.

(2) При випадковій генерації ймовірність збігу вихідного і відтвореного зразка була б 1/576 ~ 0.0017, тобто майже в 200 разів гірше.

Також існує можливість відстежувати зміни форми та структури вибірки навчання у просторі представлення у процесі навчання моделі за допомогою функції зворотного виклику у пакеті моделювання нейронних мереж Keras. Результати експериментів представлені в розділі 4.

Результати, представлені в цьому розділі дозволяють зробити висновок, що в процесі неконтрольованого навчання моделі досягали високого ступеня точності відтворення вхідних даних, на що вказують критерії зменшення помилки відхилення та коваріації. Це свідить про те, що створені моделями генеративні предствалення зниженої розмірності зберегли значну частину суттєвої інформації про розподіли вхідних даних.

3.6. Додаткові компоненти та методи генеративного навчання

Відповідно до визначень теоретичної частини роботи в розділі 2, навчена генеративна модель може здійснювати перетворення кодування з простору вхідних даних в простір представлення, $E(X)$ і генеративне перетворення $G(R)$, що діє у зворотному напрямку, тобто з латентного простору в простір вхідних даних (формула 1.2).

У нейромережевих моделях, використовуваних у роботі, кодуюче та генеративне перетворення $E(X)$, $G(R)$ були реалізовані компонентами (суб-моделями) основної архітектури нейронної мережі (рис. 3.1) наступним чином:

1. Кодуюче перетворення $E(X)$: субмодель (нейромережевий тензор) перетворюючий вхідний шар у шар латентного представлення: $L = T_E(X)$, тобто зіставляючий вхідному зразку x вектор активацій латентного шару $l(x) = T_E(x)$.

2. Генеративне перетворення $G(R)$: субмодель (нейромережевий тензор) перетворюючий шар представлення у вхідний шар: $Y = T_G(R)$, тобто зіставляючий зразку l активацій латентного шару (або точки у просторі представлення) вектор активацій вихідного шару моделі $y(l) = T_G(l)$.

Поряд з генеративними моделями нейронних мереж, які містили компоненти кодування та генерування як описано вище, в роботі використовувалися методи кластеризації по щільності, здатні визначати структуру щільності в просторах генеративних представлень повністю неконтрольованим процесом. Після навчання моделі метод кластеризації за щільністю налаштовувався на представницькому наборі немаркованих даних P трансформованого у простір представлення моделі перетворенням кодування. Настроєний метод кластеризації $K(P)$ може визначати кластер K_n відповідний вхідному зразку як:

$$K_n = K(E(x)) \tag{3.1}$$

де $E(x)$ – перетворення кодування навченої генеративної моделі.

В результаті застосування неконтрольованої кластеризації в просторі представлень можна ідентифікувати структуру латентного розподілу даних, представлену кластерами щільності визначеними методом, $K = \{ K_j \}$. Кластер K_j у цьому випадку може інтерпретуватися як натуральний, або прихований концепт вхідного зразка, встановлений моделлю в результаті навчання на неконтрольованих даних, а набір кластерів K як натуральну структуру концептів сенсорних даних.

3.7. Аналіз та візуалізація латентних розподілів

Як зазначалося у роботі, при аналізі розподілів у представленнях генеративних моделей використовувалися два типи вибірок.

Загальна вибірка була отримана застосуванням перетворення кодування до представницького набору вхідних даних *G:* $V_G = T_E(G)$. Отримання загальної вибірки не залежить від наявності і не вимагає маркованих даних відомих класів.

Вибірки відомих класів C_j були отримані застосуванням перетворення кодування до наборів категорій X_j у просторі вхідних даних: $V_j = T_E(X_j)$. Отримання вибірок категорій у просторі представлення очевидно залежить від наявності маркованих даних класів.

Щоб спростити і полегшити безпосереднє спостереження та вимірювання параметрів розподілів у латентному просторі, в моделях, використаних в експериментах, розмірність кодуючого шару, і відповідно генеративних представлень була встановлена рівною трьом, що дозволило здійснювати пряму тривимірну візуалізацію за допомогою доступного програмного забезпечення. Однак питання про оптимальний вибір розмірності простору представлень не є тривіальним, як обговорювалося в теоретичній частині роботи, де були зроблені висновки щодо визначення оптимального значення розмірності представлень на основі параметрів варіації основних концептів даних.

Візуалізація розподілів здійснювалася плотуванням вибірок розподілів класів, а також загальних, немаркованих зразків перетворених у простір представлення застосуванням кодуючого перетворення стандартними пакетами тривимірної графіки такими як matplotlib [99].

Приклади візуалізації декількох розподілів класів даних Інтернет (тобто, додатків Інтернет) включаючи фон загального розподілу в генеративному просторі представлення дані на наступній діаграмі (рис. 3.5). Як можна спостерігати на графіці, латентні розподіли класів перетворюються на добре помітні області в просторі представлень, у повній відповідності з висновками теореми про категоризацію генеративних представлень, хоча параметри категоризації, такі як

форма, розмір, щільність тощо можуть відрізнятися між класами. Цей ефект найкраще спостерігається з класами, що мають значне представництво у вхідних даних.

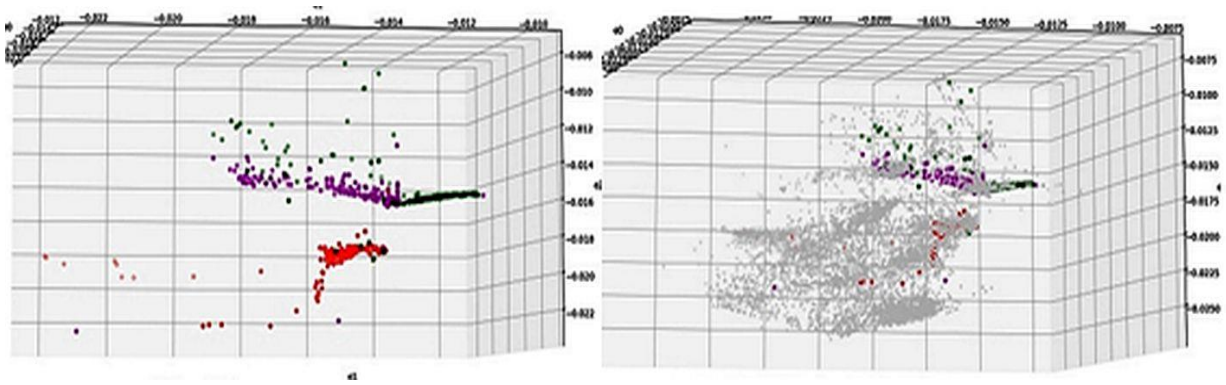


Рис. 3.5. Латентні розподіли вибірок класів даних Інтернет

Комбінація вимірювань латентних розподілів вибірок відомих класів описаними методами з безпосереднім спостереженням за допомогою візуалізації розподілів у латентному просторі дозволяє визначити параметри розподілів та порівнювати їх для різних класів даних.

3.8. Класифікація у просторах генеративних представлень

Ключовий висновок заснований на теоретичних результатах роботи і раніше опублікованих досліджень полягає в тому, що генеративна структура, яка формується, згідно з теоремою категоризації генеративних представлень в процесі навчання з мінімізацією помилки відтворення може відповідати відомим класам вхідних даних і, як наслідок, використовуватися для підвищення ефективності навчання.

При класифікації у латентному просторі, класифікатор відомого класу C_j K_j може навчатися на бінарному маркованому наборі, що включає набір зразків класу з маркою “Істина”, $P_j = \{x_l, l=1,n_p\}$, та відомих зразків які не належать класу, $N_j = \{y_m, m=1,n_n\}$ з маркою “Помилка”, трансформований в латентний простір перетворенням кодування: $K_j = K(E(P_j), E(N_j))$, де K – класифікатор контрольованого навчання одного з відомих типів.

У роботі використовувалися класифікатори геометричного типу, такі як метод найближчого сусіда [95], але підхід не обмежує вибір типу класифікатора, і можливі інші варіанти, такі як SVM, Випадковий ліс (Random Forest), нейромережеві моделі та інші відомі типи.

В результате навчання класифікатор класса K_j може визначити відомий чи "зовнішній" клас C_{ext} зразка x у вхідному просторі:

$$C_{ext} = K_j(E(x)) \tag{3.2}$$

Таким чином, концепт C_{ext} (3.2) та K_{nat} (3.1) визначений раніше методами неконтрольованої кластеризації являють собою відповідно "зовнішню категорію" зразка із заздалегідь відомого набору класів та його натуральну концепцію, отриману в результаті неконтрольованого навчання моделі та категоризації у просторі представлення.

Точність класифікації може бути виміряна на наборі маркованих даних порівнянням класу зразка, визначеного класифікатором (3.2) з його істинним класом. У наведених нижче експериментальних результатах використовувалися стандартні критерії точності класифікації: чутливість (false negatives), що відображає статистичну помилку другого типу (β- помилка) та специфічність (false positives), що відображає статистичну помилку першого типу (α- помилка). У всіх вимірах точності класифікації у роботі вимірювалися обидва критерії помилки. У ряді результатів помилка класифікації показана як сумарна F1-міра помилки, визначена як:

$$F1 = {2}\Big/{\left(\frac{1}{1-\alpha} + \frac{1}{1-\beta}\right)} \tag{3.3}$$

Значення метрики точності F1 знаходиться в інтервалі (0, 1), при цьому близьке значення до 1 вказує на високу точність по обох типах помилки, в той час як значення близьке до нуля – на високу помилку щонайменше одного з типів.

3.9. Методи навчання з використанням генеративної структури (ландшафту) представлення

В результаті обґрунтування методів у теоретичній частині роботи, включаючи теорему про категоризацію генеративних представлень, були розроблені методи навчання, здатні використовувати генеративну структуру (ландшафт) представлень для більш ефективного навчання. Зокрема, мається на увазі більш висока ефективність у таких областях як:

1) значне зниження потреби у заздалегідь відомих даних навчання;

2) спонтанність навчання: навчання може відбуватися в міру наявності навчальних даних і не вимагає великих масивів заздалегідь маркованих даних до початку навчання;

3) значно знижена залежність успіху навчання від джерела навчального набору;

4) гнучкість: вивчення нових концептів без повного перенавчання моделі;

5) на відміну від відомих підходів напівконтрольованого навчання, таких як навчання прототипів (prototype learning [103]), метод не вимагає зразків навчання всіх відомих класів одночасно; навчання може проводитися ітеративним процесом у міру надходження даних.

На основі результатів теоретичної частини роботи з категоризації генеративних представлень та обґрунтованих методів визначення структури щільності даних представлень запропоновані два різновиди методів навчання з використанням структури (ландшафту) генеративних представлень:

1. Повністю неконтрольований метод навчання розпізнавання натуральних концептів вхідних даних без вимог даних навчання відомих класів.

2. Метод навчання розпізнаванню відомих класів на основі генеративного ландшафту з мінімальними наборами навчання класу (метод навчання "сигнальною вибіркою").

Загальну діаграму інформаційних процесів методів навчання з використанням структури щільності генеративних представлень даних трафіку Інтернет зображено на рис. 3.6.

Рис. 3.6. Процес обробки інформації у методах навчання на основі генеративної структури представлень даних Інтернет

3.9.1. Метод розпізнавання відомих класів на основі генеративного ландшафту з мінімальними наборами навчання

Метод ітеративного навчання на основі генеративного ландшафту ґрунтується на ітеративному формуванні класифікатора концепту шляхом визначення ітерацій навчальних наборів концепту у просторі представлення на основі визначеної структури кластерів щільності даних. Загальну діаграму процесу методу подано на рис. 3.6.

Вхідними даними методу є:

– генеративна структура представлення даних, така як набір кластерів щільності, визначена в повністю неконтрольованому процесі як обґрунтовано вище;

– невеликі набори позитивних зразків відомого класу, які можуть складатися з початкового “сигнального” набору та наборів ітерацій навчання.

Результатом навчання є постійний класифікатор класу зі значною точністю розрізнення приналежності до класу як зразків класу так і тих, що не належать до нього.

У конкретній реалізації методу в роботі використовувалися два типи зразків представлених в ітераціях навчального набору: *реальні* та *штучні*. Реальні зразки містяться у початковій вибірці концепту (нульова ітерація або "сигнальна вибірка", набір позитивних зразків класу розміром від кількох зразків) та додаткових даних, отриманих в результаті емпіричного досвіду. Штучні зразки генеруються конкретним алгоритмом методу з урахуванням визначеної латентної кластерної структури, одним з або комбінацією способів:

– як центри кластерів асоційованих із зразками класу;

– обрані випадковим чином елементи кластерів асоційованих із зразками класу;

– як набір елементів на визначеній відстані від центрів кластерів.

Співвідношення реальних та штучних зразків у наборі є параметром методу.

В результаті застосування алгоритму побудови навчальних наборів ітерацій навчання, вже перша ітерація класифікатора може приймати обґрунтовані рішення про належність зразків вхідного простору до класу. При розширенні навчальних наборів класу зі збільшенням числа ітерацій навчання точність класифікатора зростає і може наблизитись до результатів моделей, навчених на значному масиві маркованих даних. Можна зауважити, що такий процес ітеративного навчання нагадує навчання біологічних систем на основі емпіричних спроб та помилок.

Ітеративний процес навчання на генеративному ландшафті ілюстровано на рис. 3.7, що показує зміни розподілу навчального набору класу в просторі представлення у процесі навчання.

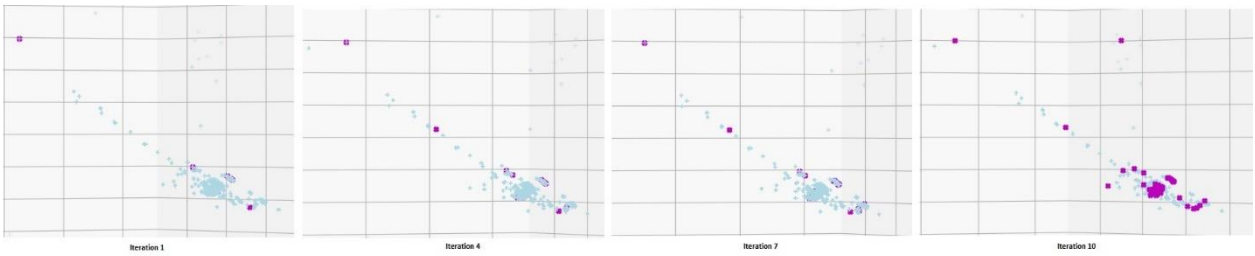


Рис. 3.7. Зміна латентного розподілу навчальних наборів класу при ітеративному навчанні на генеративному ландшафті

На діаграмі область розподілу відомого набору класу позначено блакитним кольором, навчальни набори в ітераціях навчання – малиновим. Видно, як зі зростанням числа ітерацій навчальні набори покривають значну частину області розподілу класу, що веде до підвищення точності класифікації.

У роботі використовувалася найпростіша реалізація методу і можна розраховувати що уточнення, розвиток та оптимізація методу може й надалі підвищити ефективність навчання розпізнавання відомих класів за мінімальних даних навчання.

3.9.2. Метод неконтрольованого розпізнавання натуральних концептів даних на основі генеративного ландшафту

Метод може використовуватися для розпізнавання натуральних типів або концептів вхідних даних на основі встановленої структури генеративного ландшафту щільності. Метод не вимагає наперед відомих даних концептів. Обґрунтування методу було представлено в теоретичній частині роботи (розділ 2).

В основі методу лежить обґрунтоване припущення, що головні концепти даних можуть збігатися зі структурами генеративних представлень. Визначення латентної структури представлень дозволяє визначити основні кластери щільності даних і будувати класифікатори концептів на основі визначеної латентної кластерної структури.

Діаграма на рис. 3.8 ілюструє побудову навчального набору класифікатора концепту з урахуванням кластерної структури щільності. Зразки кластера концепту розглядаються як позитивний набір класу, зразки інших кластерів, як набір поза

класом. Потім класикатор навчається на маркованому бінарному наборі звичайними методами.

Як у випадку методу навчання сигнальною вибіркою, описаного вище, реалізація методу має кілька параметрів, таких як чисельність наборів концепту і поза концептом, метод генерації зразків концепту на основі кластерної структури.

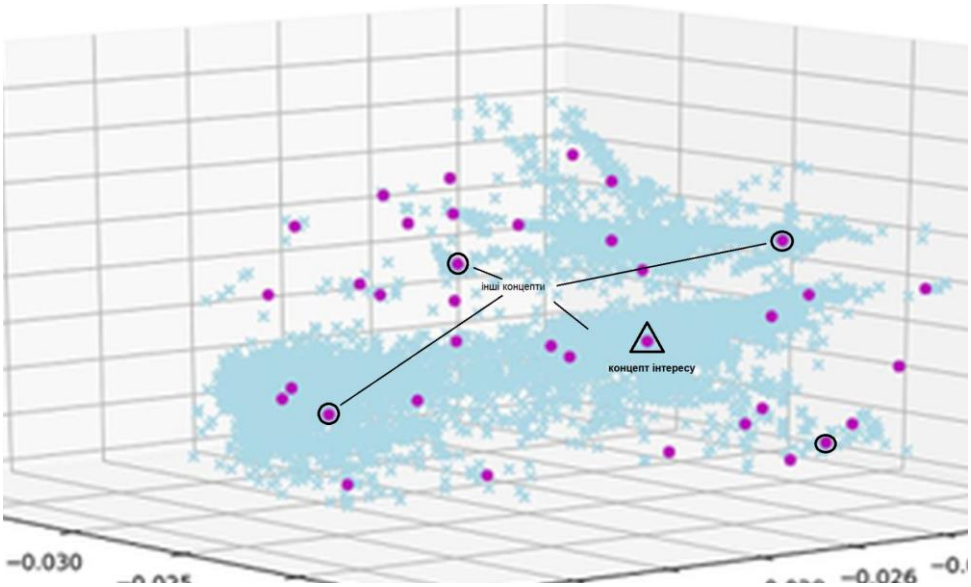


Рис. 3.8. Побудова навчального набору класифікатора концепту на основі генеративного ландшафту кластерів щільності

Встановлена в результаті побудови класифікаторів натуральних концептів структура натуральних категорій даних може бути пов'язаназ відомими зовнішніми класами. Як показано в експериментальних результатах роботи, у деяких випадках спостерігалася сильна кореляція між натуральними категоріями та відомими класами даних у наборах різних типів.

РОЗДІЛ 4.
РЕЗУЛЬТАТИ АНАЛІЗУ ЗПРОПОНОВАНИХ МЕТОДІВ ТА МОДЕЛЕЙ

У даному розділі дисертаційної роботи наведено результати експериментальної перевірки теоретичних висновків та методів навчання з використанням генеративної латентної структури (ландшафту) даних Інтернет, запропонованих у роботі.

4.1. Неконтрольоване генеративне навчання

Архітектура, параметри, методи та критерії неконтрольованого навчання генеративних моделей були описані у розділі 3 дисертаційної роботи.

4.1.1. Формування генеративного ландшафту в процесі неконтрольованого навчання

В експериментах був встановлений ефект структуризації даних у просторі представлених генеративних моделей роботи в процесі неконтрольованого навчання зі зменшенням помилки відтворення.

У сесіях навчання моделей на масиві даних пакетів Інтернет (до 100 епох епох навчання) були виміряні розміри вибірки візуалізації в просторі представлення та кількість видимих структур щільності у вибірці визначені методом кластеризації за щільністю, а також візуальним спостереженням. Розміри вибірки та кількість визначених латентних структур були визначені за допомогою методу зворотного виклику, який виконував вимірювання у просторі представлення моделі після кожної 10-ї епохи навчання.

Таблиця 4.1

Структуризація латентного розподілу даних загального набору у процесі неконтрольованого навчання (дані Інтернет)

Епоха навчання	Кількість кластерів щільності	Розмір вибірки [(1)]
0 (необученная модель)	5	0.975, 0.560, 0.478
20	8	0.027, 0.063, 0.031
40	10	0.024, 0.044, 0.024
60	12	0.024, 0.042, 0.024
80	12	0.024, 0.044, 0.024
100	15	0.0235, 0.042, 0.024

[(1)] У латентних координатах (координатах простору представлення)

Як можна бачити з наведених вище результатів, у міру стиснення вибірки в просторі представлення відбувається виділення характерних структур (кластерів) щільності, які можна означити методами неконтрольованої кластеризації за щільністю. Цей процес спостерігався стійко в багаторазових експериментах і наведені результати дозволяють стверджувати, що структуризація генеративних представлень є необхідним наслідком генеративного навчання зі зниженням помилки у повній відповідності до висновків теоретичної частини роботи. Результати є прямим експериментальним підтвердженням теореми про категоризацію генеративних представлень.

4.1.2. Класифікація у процесі генеративного навчання

В експериментах було встановлено стійке підвищення точності класифікації при навчанні з маркованими даними відомих класів вхідних даних у просторі представлень у процесі неконтрольованого генеративного навчання.

У сесіях навчання моделей на масиві даних Інтернет (до 100 епох навчання) було виміряно середню точність класифікації для кількох класів (тобто додатків Інтернет) у процесі неконтрольованого генеративного навчання. Точність класифікації в латентному просторі подання була визначена за допомогою методу зворотного виклику, який виконував вимірювання в рівні кодування моделі після кожної 10-ї епохи навчання.

Таблиця 4.2

Точність класифікації у процесі неконтрольованого генеративного навчання (дані Інтернет)

Епоха навчання	Точність класифікації (F1- метрика)
10	0.865
20	0.908
40	0.910
60	0.914
80	0.913

Як можна бачити з результатів табл. 4.2, протягом неконтрольованого навчання моделі відбувається стабільне підвищення точності класифікації.

Цікавим є збіг ефектів структуризації в попередньому розділі з підвищенням точності класифікації у ході генеративного навчання.

Ці результати можуть вказувати, що структура, яка виникає у представленнях у ході генеративного навчання може корелювати з відомими типами (класами) вхідних даних. Якщо генеративна структура не мала б значущої кореляції з типами даних, було б важко пояснити позитивну кореляцію між неконтрольованим навчанням моделі, яке не передбачає будь-якого зв'язку з відомими класів та підвищенням точності класифікації в просторі генеративного представлення.

Подані в цьому розділі результати показують, що у ході неконтрольованого навчання відбувається виділення характерних інформаційних структур у просторі генеративних представлень. У зв'язку з одночасним підвищенням точності класифікації це дозволяє припустити кореляцію між генеративними структурами у просторі представлень та відомими типами вхідних даних, що є непрямим підтвердженням висновків теореми про категоризацію генеративних представлень та подальших припущень роботи.

4.2. Розподіли даних у просторах генеративних представлень

4.2.1. Латентні розподіли загальної вибірки

При візуальному спостереженні набору загальної вибірки даних Інтернет у просторі представлень можна ідентифікувати множинні структури щільності, такі як кластери, смуги, відділені області та інші структури. Застосування методів неконтрольованої кластеризації дозволяє встановити та візуалізувати кластери щільності загального розподілу та розподілу категорій, у той час як використання методів гістограм дозволяє безпосереднє спостереження та вимірювання параметрів розподілу щільності загальної вибірки та вибірок категорій. Ці методи дозволяють аналізувати та провести вимірювання наступних та інших параметрів:

– розміри області розподілу загального зразка у координатах простору представлення;

– статистичні характеристики латентних розподілів, такі як середнє значення, варіація, коефіцієнти коваріації та кореляції;

– параметри розподілів щільності даних у просторі представлення як для загального розподілу, так і за категоріями;

– характеристичні параметри структур щільності у загальному розподілі та розподілах категорій.

У нижченаведених розділах представлені характеристики розподілів даних загальної вибірки в генеративних представленнях даних Інтернет та зображень.

Набір даних Інтернет

Таблиця 4.3

Характеристики латентного розподілу загальної вибірки, набір Інтернет

Характеристика	Значення	Примітка
Макс. / мінімальний розмір (1)	0.0394 / 0.0166	абсолютна величина
Макс. / мінімальна щільность	210.1 / 0.0	
Відношення областей високої / низкої щільності	0.077 / 0.907	висока щільність: > 2.5 середньої (1) низька щільність: < 0.5 середньої
Середнє значення, діапазон (2)	0.0253 - 0.0316	
Стандартне відхилення, діапазон (2)	0.022 – 0.035	
Генеративна структура (3)	17 / 36 *	* застосовувався метод MeanShift

(1) Середня (уніформна) щільность: 1.0

(2) По осям пространства представления

(3) Чисельність структур щільності у розподілі вибірки (візуальне спостереження, метод кластеризації за щільністю)

На графіках (рис. 4.1) наведено розподіли щільності даних Інтернет по осях простору представлення. Порівнюючи з розподілами у просторі вихідних параметрів можна помітити суттєве ускладнення структури розподілів по всіх осях координат, які стають багатомодальними з виділенням декількох областей підвищеної щільності.

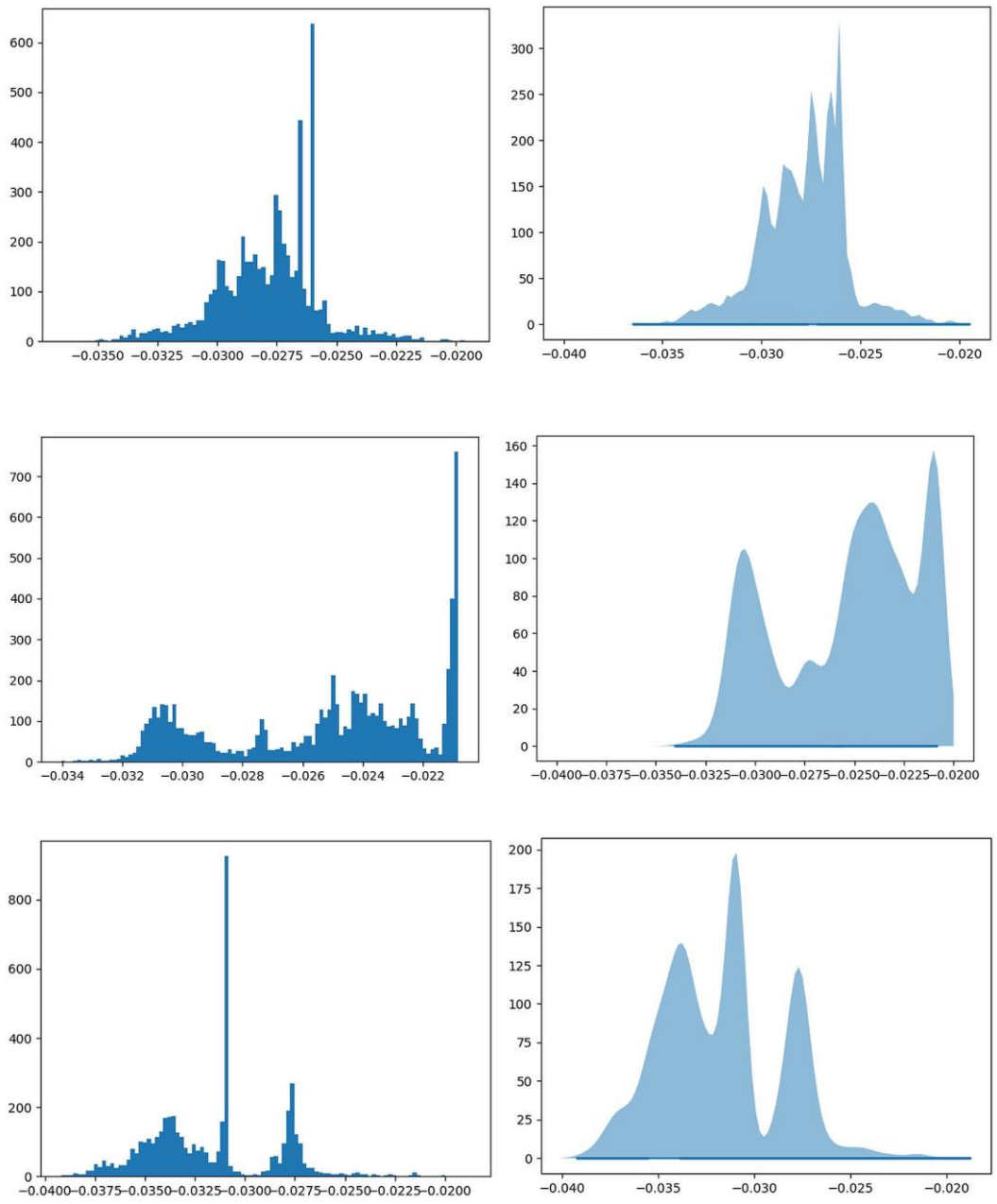


Рис. 4.1. Розподіли щільності даних Інтернет по осях простору представлення (ліворуч: гістограма, праворуч: наближення методом ядерної щільності)

Аналогічний висновок можна зробити з аналізу багатовимірної гістограми розподілу щільності у просторі генеративного представлення даних Інтернет, де проглядається складна структура з численними кластерами щільності. На діаграмі регіони підвищеної щільності відзначені відтінками червоних кольорів, а низького, зеленого.

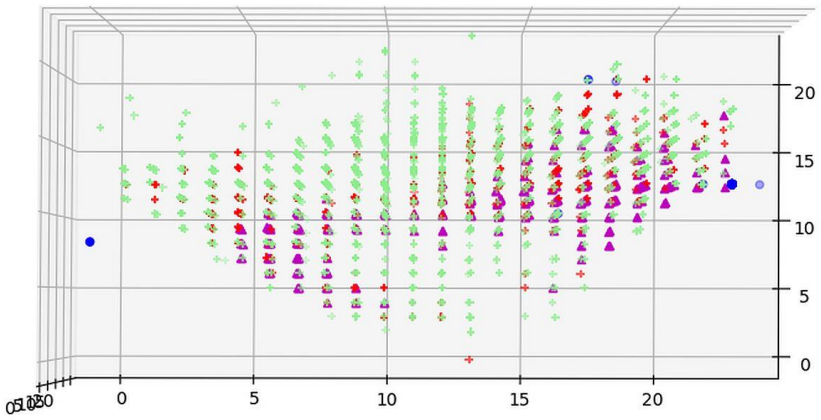


Рис. 4.2. Гістограма розподілу щільності в просторі представлення, дані Інтернет

Набір даних зображень

Характеристики розподілу даних загальної вибірки набіру зображень представлені у табл.4.4.

Таблиця 4.4

Характеристики латентного розподілу загальної вибірки, набір зображень

Характеристика	Значення	Примітка
Макс. / мінімальний розмір	0.095 / 0.0011	абсолютна величина
Макс. / мінімальна щільность	241 / 0.001	
Відношення областей високої / низкої щільності	0.027 / 0.967	висока щільність: > 2.5 середньої низька щільність: < 0.5 середньої
Середнє значення, діапазон	0.0217 - 0.0531	
Стандартне відхилення, діапазон	0.0116 - 0.0145	
Генеративна структура	7 / 16 *	* застосовувався метод MeanShift

На діаграмі 4.3. представлена проекція розподілу загального масиву зображень у просторі представлення. Як і в попередньому випадку даних Інтернету ясно проглядається неоднорідність розподілу з численними структурами підвищеної щільності.

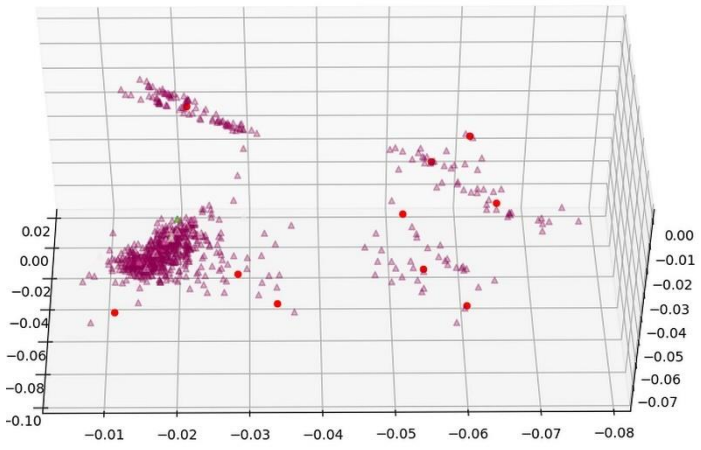


Рис. 4.3. Розподіл даних зображень у просторі представлення з визначеними кластерами щільності

На графіках (рис. 4.4) наведено розподіл щільності даних зображень по осях простору представлення.

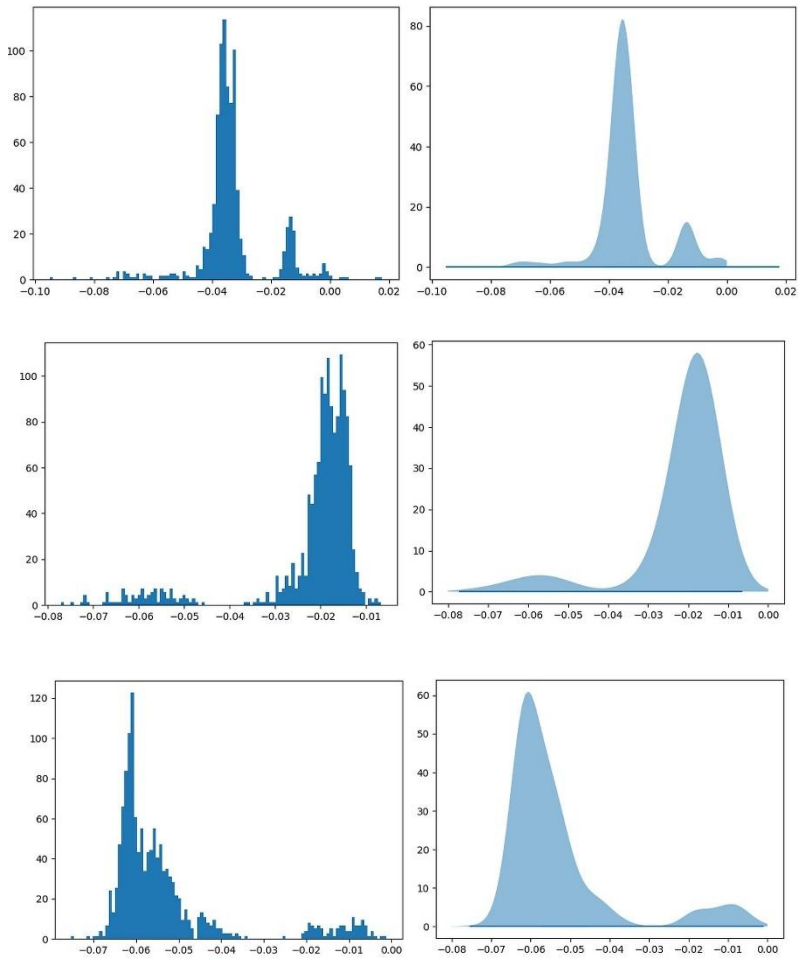


Рис. 4.4. Розподіли щільності даних зображень по осях простору представлення (ліворуч: гістограма, праворуч: наближення методом ядерної щільності)

Наведені характеристики загального розподілу у просторі подання двох суттєво різних типів даних: Інтернет та зображень, підтверджують висновок, зроблений у попередніх розділах про структуризацію в процесі неконтрольованого генеративного навчання, зокрема, виділення неоднорідностей структури щільності з областями підвищеної щільності та розрядженими регіонами.

4.2.2. Латентні розподіли вибірок категорій

Набір даних Інтернет. На графіках (рис. 4.5) подано латентні розподіли класів даних Інтернет. Пряма візуалізація підтверджує припущення про певний розподіл (тобто освіченим кінцевим числом пов'язаних областей у просторі представлення) для більшості розглянутих класів даних.

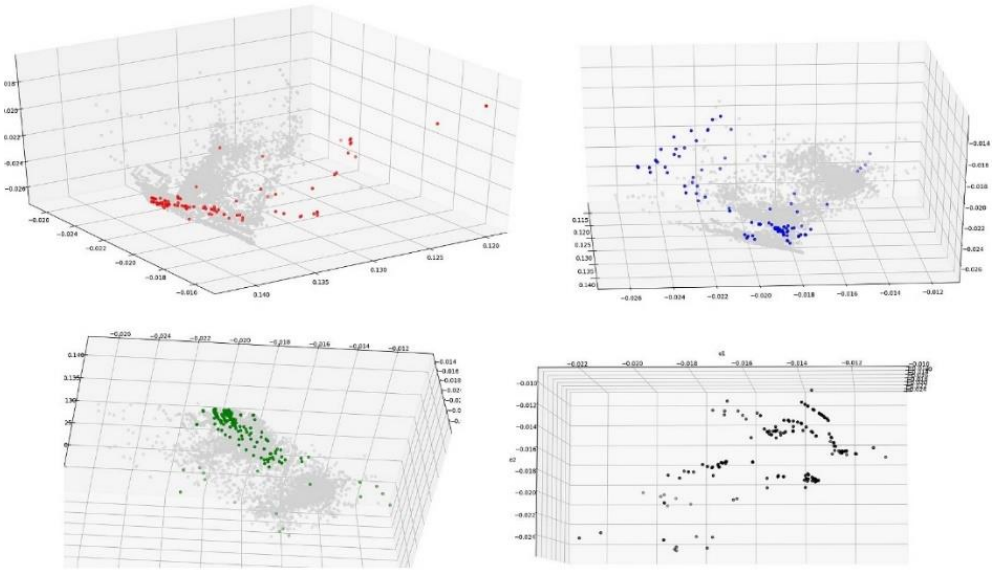


Рис. 4.5. Латентні розподілу категорій, дані Інтернет (за годинниковою стрілкою, згори зліва, Інтернет додатки: Месенджер, Торрент, протокол HTTP, Онлайн гра)

Характеристики латентних розподілів даних категорій виміряні для найбільш представлених у наборі класів методами, описаними вище, представлені в табл. 4.5.

Таблиця 4.5

Характеристики латентних розподілів категорій, набір Інтернет

Клас (додаток)	Тип	Кількість зразків	Відносний розмір (1)	Структура	Концен-трація	Щільність
DNS	мережевий протокол	1000	0.005	3	0.11	7.5
NTP	мережевий протокол	200	0.01	4	0.15	7.0
SMTP	емейл	300	0.1	3	0.11	6.0
Telnet	віддалена сесія	200	1×10^{-4}	1	0.036	9.5
MSN	месенджер	580	0.015	2	0.07	7.0
Escale	онлайн гра	200	0.015	4	0.15	7.0
XBox	онлайн гра	460	0.2	6	0.21	6.0
BitTorrent	сервіс файлів	670	0.5	10	0.36	4.0
Streaming	стрімінг	370	0.1	5	0.18	6.5
HTTPS	Гипертекст	4800	0.4	> 10	0.4	4.0

(1) Відносно розміру загального набору

Як видно з результатів, для більшості виміряних додатків латентна область розподілу зразків класів складалася з невеликої кількості чітко визначених кластерів, хоча характеристики розподілів могли відрізнятися між класами, від дуже компактних розподілів з невеликим числом маленьких і щільних кластерів (DNS, NTP, Telnet) до більш розріджених з великою кількістю кластерів (Торент, HTTP). Це спостереження може бути пояснено характеорм звя'зку між відомими категоріями та натуральними кластерми даних: якщо в одних випадках, таких як відеогра тощо, марка (label) програми суттєво визначає характер та мережеву поведінку сесії Інтернет і, відповідно, її характеристичний " портрет " у просторі представлення, то, наприклад, протокол HTTP, що належить до другого типу розподілів може нести різні додатки з суттєво різними характеристиками, які можуть бути віднесені до різних латентних кластерів.

Це спостереження дозволяє зробити важливий висновок: хоча відомі класи вхідних даних та приховані, або натуральні концепти можуть бути взаємопов'язані та корельовані, характер та ступінь залежності між ними не можуть прийматися за належне у всіх випадках.

Набір даних зображень. Візуалізації розподілів класів зображень у просторі представлення генеративних моделей представлені на діаграмі (рис. 4.6). Як видно, розподіли в більшості випадків представлений компактними чітко визначеними областями розподілу вибірок категорій.

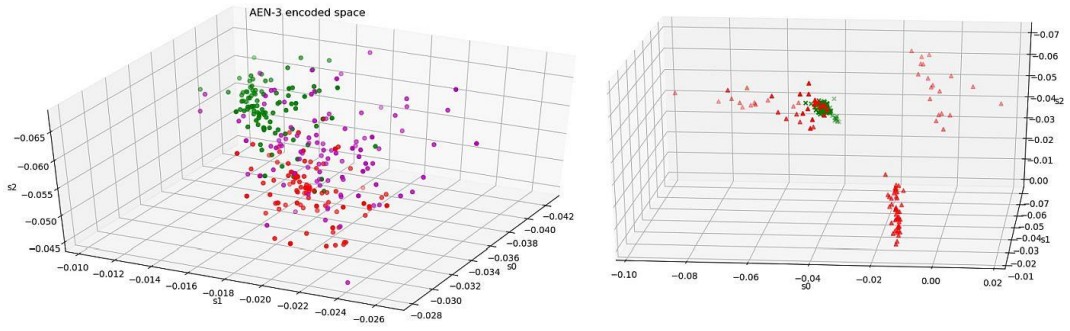


Рис. 4.6. Латентні розподіли класів зображень: класи "ліс"; "поле"; "вода" (ліворуч), "сліди транспорту"; "транспортні засоби" (праворуч)

Латентні розподіли щільності вибірок класів у координатах простору представлення для обраних класів наведені на діаграмі 4.7, де як і у разі загальних даних, спостерігається характеристична структура щільності розподілу.

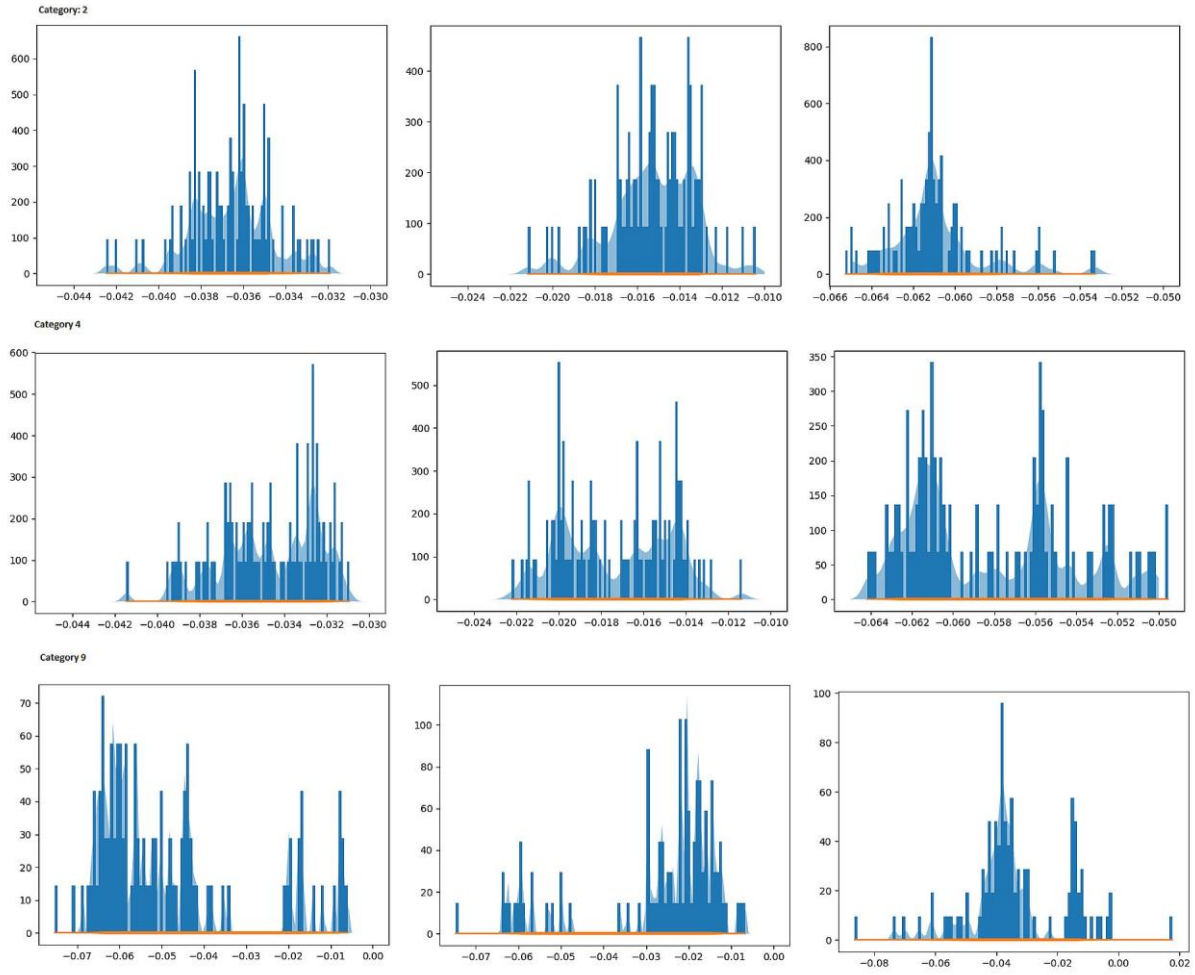


Рис. 4.7. Розподіли даних категорій по осях простору представлення. Зверху вниз: класи "поле", "дороги", "транспортні засоби"

У таблиці 4.6 представлені основні характеристики латентних розподілів класів даних зображень.

Таблиця 4.6

Характеристики латентних розподілів класів зображень.

Клас зображень	**Число зразків**	**Відносний розмір** (1)	**Структура**	**Концен-трація**	**Щільність** (2)
Будинки	~ 100	0.005	3	0.17	211
Поле	-	0.18	2	0.11	246
Ліс	-	0.16	2	0.11	357
Водойма	~ 100	0.19	2	0.11	375
Дороги	-	0.23	4	0.22	228

Закінчення табл. 4.6

Сліди транспорту	-	0.27	6	0.33	174
Велика структура	-	0.28	6	0.33	292
Об'єкт	~ 100	0.78	7	0.39	135

(1) Відносно розміру загального набору
(2) Відносно середньої (уніформ) щільності розподілу, 1.0

Як і у випадку даних Інтернет для деяких категорій спостерігається виражена категоризація областей розподілу у просторі представлення. Для низки категорій розподіли більш “розмиті”, що може свідчити про недостатню глибину архітектури або представництво набору даних. У цьому випадку потрібне подальше дослідження для створення ефективних представлень.

4.2.3. Поділ областей розподілу категорій при генеративного навчання

Цікаві висновки про генеративну категоризацію дозволяє зробити гістограмування латентних розподілів класів на тлі загального розподілу. Приклад такого розподілу для програми DNS (один з мережевих протоколів Інтернет) представлений на наступній гістограмі (рис. 4.8.а)) латентного розподілу густини, де зразки класу (синій колір) виділені на загальній гістограмі розподілу щільності (вища щільність відповідає відтінкам червоного кольору, низька – зеленого).

Візуалізація латентних розподілів щільності в координатах представлення дозволяє помітити що зразки різних класів знаходяться в різних геометричних областях простору генеративного представленння створеного моделлю.

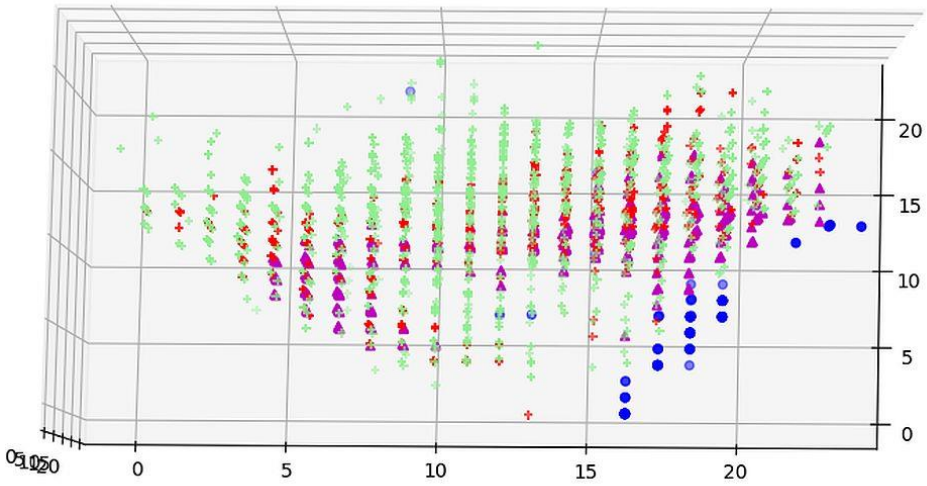


Рис. 4.8. а) Латентний розподіл класу (DNS) на тлі загальної гістограми щільності, набір Інтернет

На діаграмі проекцій розподілу добре видно, що розподіли класів розташовані в особливих просторових областях представлення, тут може бути використана аналогія “поличок у шафі” (рис. 4.8. б)).

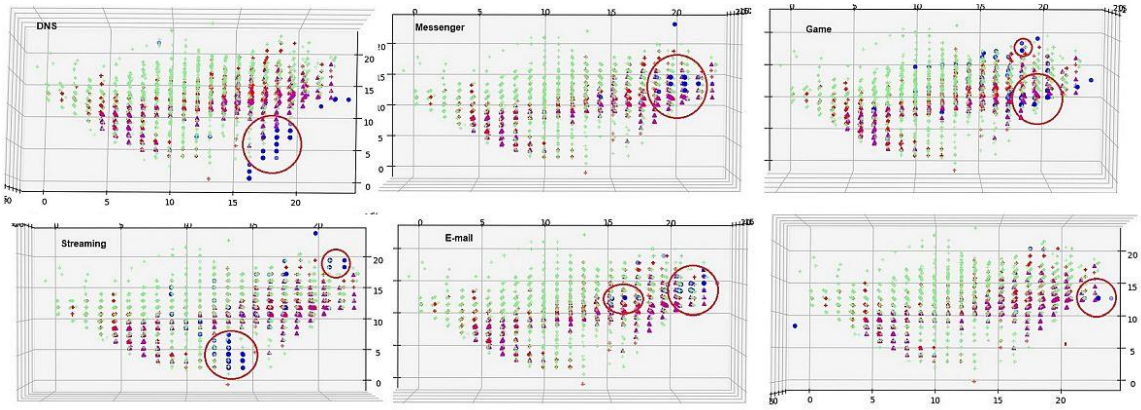


Рис. 4.8.б) Генеративна категоризація при неконтрольованому навчанні (проекції), набір Інтернету (синій: додаток DNS)

Аналогічні результати генеративної категоризації було отримано з набором даних зображень (рис. 4.9).

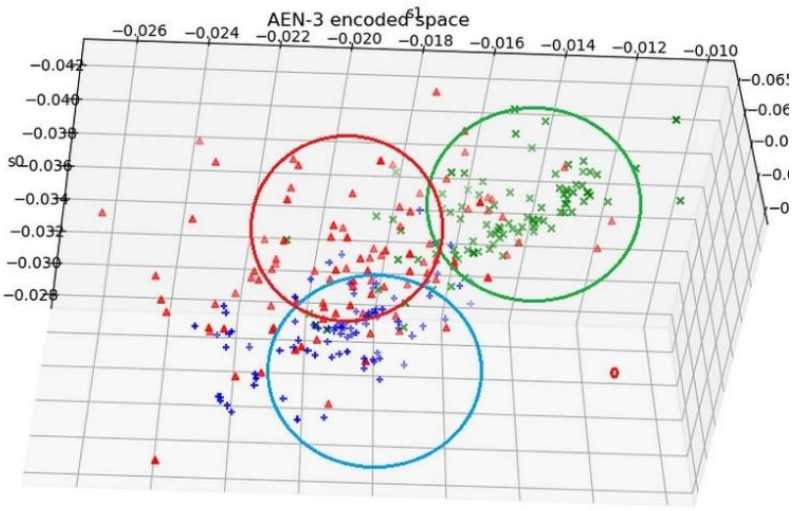


Рис. 4.9. Генеративна категоризація при неконтрольованому навчанні, дані зображень

Спостереження ефекту генеративної категоризації в представленнях моделей неконтрольованого генеративного навчання на двох масивах даних різних за типом та характером дозволяє зробити висновок про загальний характер ефекту, та є прямим експериментальним підтвердженням теореми генеративної категоризації та висновків теоретичної частини роботи.

4.3. Класифікація в просторі генеративних представлень

Вимірювання точності класифікації відомих класів у просторі представлення генеративних моделей дозволяє простежити кореляційний зв'язок між ефективністю генеративної категоризації та відомими типами вхідних даних.

У наведених нижче експериментах класифікатор відомих класів даних Інтернет навчався обмеженим набором маркованих зразків у просторі представлення (розмір навчального набору 30 - 100 зразків). У наступних стадіях результати класифікації за класами були зведені з результатами вимірювання категаризаційних параметрів латентних розподілів класів.

У табл. 4.7 та 4.8 наведено результати класифікації після навчання на латентних наборах даних Інтернет та зображень із класифікаторами двох типів: глибокої нейронної мережі контрольованого навчання (DNN) та класифікатора найближчих сусідів (kNN). Навчальний набір класів складав 300 маркованих

зразків класу та фону. В експерименті вимірювалася помилка першого та другого типу (вказані як чутливість, вибірковість), усереднені по 100 тестам зі 100 випадково обраних зразків класу та поза класом (тобто 100х крос валідація). Результати усереднені за 10–20 сесіями навчання. Рівень впевненості 95% при інтервалі 0.3%.

Таблиця 4.7

Класифікація при контрольованому навчанні в латентному просторі, набір Інтернет (метрика точності: чутливість / вибірковість)

Клас (додаток) Інтернет	**Точність, класифікатор DNN**	**Точність, класифікатор kNN**
DNS	92.0 / 8.3	92.5 / 7.3
NTP	100 / 11.5	99.5 / 13.0
Telnet	96.8 / 9.2	97.9 / 11.6
XBox (онлайн-гра)	83.9 / 7.6	87.8 / 10.2
Месенджер	88.9 / 8.2	89.9 / 5.0
Мейл	90.4 / 19.4	92.0 / 20.6
Відео / аудіо стрім	98.7 / 1.8	91.0 / 6.9

Таблиця 4.8

Класифікація при контрольованому навчанні в латентному просторі, набір зображень (метрика точності: чутливість / вибірковість)

Клас зображень	**Точність, класифікатор DNN**	**Точність, класифікатор kNN**
Будинки	92.0 / 9.1	96.3 / 7.3
Поле	94.0 / 8.3	93.5 / 7.7
Ліс	100 / 11.5	99.5 / 13.0
Водойма	97.3 / 6.7	98.2 / 10.1
Дороги	96.8 / 9.2	97.9 / 11.6
Сліди транспорту	83.9 / 7.6	87.8 / 10.2
Велика структура	88.9 / 8.2	89.9 / 5.0
Об'єкт	90.4 / 19.4	92.0 / 20.6

Результати щодо ефективності класифікації в просторах представлень генеративних моделей у цьому розділі підтверджують попередні результати про переваги використання неконтрольованих представлень для підвищення ефективності класифікації та показують, що існування генеративної структури корельованої з характерними типами вхідних даних дозволяє або підвищити точність класифікації, або зберегти її при значному зниженні розміру навчальних наборів.

Кореляція геометричної категоризації та точності класифікації при контрольованому навчанні на латентних наборах. Як зазначалося раніше, ефект генеративної категоризації в неконтрольованих представленнях може мати прямий вплив на передбачувальну силу класифікаторів навчених на маркованих латентних вибірках класів. Цей зв'язок був досліджений та виміряний безпосередньо в розділі 3.

Результати по кореляції категоризаційних властивостей неконтрольованих представлень, створених моделями генеративного навчання і точності класифікації при контрольованому навчанні на латентних наборах підтверджують висновок про кореляційну залежність між інформаційною структурою (ландшафтом), що виникає при неконтрольованому генеративному навчанні і характерними типами вхідних даних.

4.4. Методи навчання на генеративному ландшафті представлень даних Інтернет

Однією з головних цілей роботи було емпіричне випробування методів навчання з використанням генеративної структури (ландшафту) представлень заснованих на висновках теоретичної частини роботи з категоризації генеративних представлень.

Перевірка ефективності методів з наборами даних Інтернет та інших типів та характеристик важлива для підтвердження їх придатності до використання із загальними даними різних типів та формування інформаційної технології навчання на основі генеративної структури (ландшафту) представлень.

4.4.1. Метод розпізнавання відомих класів на основі генеративного ландшафту з мінімальними наборами навчання

Метод навчання “сигнальної вибірки” визначений та обґрунтований у розділі 2.5.2, був застосований до наборів даних Інтернет та зображень. Початкове навчання проводилося з мінімальними наборами від 1 до 10 представників класу. При створенні наборів навчання використовувалися реальні зразки концепту та штучні зразки створені на основі структури щільності генеративного ландшафту.

У табл. 4.9 та табл. 4.10 представлені результати застосування методу з наборами даних Інтернет та зображень. "Сигнальна точність" позначає точність класифікатора навченого з використанням початкового набору (тобто нульової ітерації навчання) від одного до декількох позитивних зразків класу, в той час як "Навчена точність" відноситься до точності класифікатора після декількох ітерацій навчання (як правило, 10 ітерацій, 10 відомих зразків класу та фону). У програмній реалізації методу використовувалися 2-5 штучних позитивних зразків на кластер класу та 1-2 штучних негативних зразків на кластер поза класом, як обговорювалося вище. Наведено результати найкращого набору параметрів.

"Фактор категоризації" є мірою категоризації латентного розподілу класу, в діапазоні 1 – 10 на підставі основних характеристик компактності розподілу: відносний розмір, структура, щільність як визначено в розділі 2.3.

Таблиця 4.9

Ефективність навчання на генеративному ландшафті методом сигнальної вибірки, набір Інтернет (метрика точності: F1)

Додаток	Сигнальна точність	Навчена точність	Фактор категоризації
DNS	0.861	0.915	8.5
Месенджер	0.748	0.921	7.5
Мейл	0.702	0.93	7.2
Відео / аудіо стрім	0.798	0.892	6.9
XBox (онлайн-гра)	0.782	0.808	6.5
HTTP	0.312	0.762	4.5

*Результати усереднені за 20 сесіями навчання. Рівень впевненості 95% при інтервалі 0.3%

Таблиця 4.10

Ефективність навчання на генеративному ландшафті методом сигнальної вибірки, набір зображень (метрика точності: чутливість/вибірковість)

Категорія	Сигнальна точність	Навчена точність	Фактор категоризації
Будинки	90.5 / 93.1	92.5 / 92.7	7.5
Поле	99.4 / 86.8	100 / 88.9	8.0
Ліс	97.8 / 87.9	97.9 / 90.8	8.0
Водойма	97.3 / 93.3	98.2 / 89.9	8.5
Дороги	84.8 / 86.2	87.9 / 92.4	7.0
Сліди транспорту	83.9 / 93.4	89.9 / 95.0	6.5
Велика структура	87.7 / 81.8	90.4 / 90.6	6.0
Об’єкт	77.2 / 76.9	91.0 / 93.1	4.5

Як можна бачити з результатів наведених вище, для більшості вимірюваних класів в обох наборах ефективність навчання з сигнальним набором всього з одного зразка була значно краща за випадковий вибір (при кількості класів близько 10 випадковий результат був би навколо 0.1). Також важливо відзначити, що метод зберігає високий рівень вибірковості по відношенню до помилки першого типу, чого не завжди вдається досягти стандартними методами навчання, в яких точність результуючої класифікації залежить від представницькості даних у навчальному наборі і дефіцит негативних маркованих даних може суттєво спричиняти помилку першого типу.

З отриманих результатів можна зробити висновок, що метод ітеративного навчання з початковим “сигнальним” набором зразків працює найефективніше з добре категорованими концептами, тобто мають компактний регіон розподілу у просторі представлення з невеликою кількістю характерних кластерів.

Зворотним прикладом є випадки програми HTTP (гіпертекст) у наборі Інтернет або класу “Об'єкти” у наборі зображень. В обох випадках латентні розподіли були розмиті як за геометричними параметрами поширення та щільності, так і за кількістю характерних структур (кластерів щільності). Водночас ефективність навчання в експериментах була значно нижчою для таких типів

даних. У таких випадках може знадобитися подальше уточнення даних та/або генеративної архітектури, наприклад, глибини, розрідженості та інших архітектурних особливостей для створення ефективних представлень вхідних даних. Наприклад, у випадку протоколу Гіпертекст (HTTP, дані Інтернет), як зазначалося раніше, дані відмічені одним класом можуть включати значну кількість незалежних додатків і уточнення маркування даних по додатку швидше ніж зовнішньому протоколу сесії може призвести до розбиття області розподілу категорії протоколу на кілька добре визначених регіонів розподілу додатків.

4.4.2. Метод розпізнавання натуральних концептів даних трафіку Інтернет

У цьому розділі проведено експериментальну перевірку методу повністю неконтрольованого навчання розпізнавання натуральних концептів даних на основі генеративного ландшафту, визначеного в розділі 2. Результати перевірки методу набору даних зображень наведено в табл. 4.11.

Таблиця 4.11

Ефективність методу неконтрольованого навчання розпізнавання натуральних концептів, набір зображень (метрика точності: чутливість / вибірковість; класифікатор: kNN)

Натуральний концепт (кластер)	Точність класифікації	Чисельність концепту (1)	Щільність області концепту (2)
0 “Ліс-вода”	99.0 / 97.0	0.24	1.5×10^3
1 “Поле”	99.0 / 96.5	0.17	1.0×10^3
2 “Вода”	98.8 / 98.0	0.06	1.1×10^3
6 “Об’єкт”	98.0 / 92.0	0.035	380
16 “Дороги”	98.2 / 93.1	0.028	570
20 Не ідентифіковано	98.0 / 97.0	0.026	550

Результати усереднені за 10 сесіями навчання
(1) Відносно розміру загальної вибірки (2) Відносно розміру середньої щільності (1.0)

Можна зазначити, що згідно з опублікованими результатами у деяких випадках спостерігається сильна кореляція між натуральними концептами неконтрольованого ландшафту та відомими класами вхідних даних [35, 104]. У

таких випадках *класифікатор концепту без навіть мінімальних даних відомого класу*, може показувати хороші результати з його класифікації. Ці результати є додатковим аргументом на користь зв'язку між відомими класами та натуральними категоріями даних.

Суттєва перевага запропонованого методу полягає в тому, що він не вимагає навіть мінімальних наборів даних категорій для успішного навчання, як випливає з наведених вище результатів, і може застосовуватися в аналізі структури даних при повній відсутності попередньої інформації про їх зміст, наприклад, в нових ситуаціях, оточенні і аналогічних.

4.5. Порівняльний аналіз запропонованих методів навчання генеративних систем з результатами загальноприйнятих методів класифікації даних Інтернет

В роботі проведено порівняльний аналіз застосування методу навчання на основі генеративної структури представлень з результатами навчання класифікації даних Інтернет опублікованими в літературі [28, 29] (табл. 4.12). Представлені для загальновідомих методів контрольованого навчання результати класифікації трафіку Інтернет відносяться до випадку загальномережевого застосування, *коли мережі набору навчання та вимірювання точності класифікації відрізняються*.

Таблиця 4.12

Порівняльні результати з загальноприйнятими методами класифікації трафіку Інтернет

Клас Інтернет	**Ландшафтний метод, початкова точність [(1)]**	**Ландшафтний метод кінцева точність [(1)]**	**Методи контрольованого навчання [(2)]**
Месенджер, мейл	0.75	0.92	0.6 – 0.84
Інтернет-протоколи	0.86	0.92	0.78 – 0.91
Streaming	0.71	0.84	–
Загально	0.78	0.91	0.75 – 0.92

[(1)] Розмір навчального набору, початкова точність: 3–10 позитивних зразків класу; кінцева точність: 100 зразків класу (10 ітерацій навчання); метрика точності F1.

[(2)] Набір навчання відрізнявся від набору верифікації; набір навчання від сотень до кількох тисяч зразків класів.

З результатів порівняння точності класифікації можна зробити висновок, що методи навчання з використанням генеративного ландшафту дозволяють значно зменшити залежність результатів навчання від вибору навчального набору та досягати результатів навчання порівнянних із результатами традиційних методів контрольованого навчання зі значно меншими наборами навчальних даних.

4.6. Інформаційна технологія навчання з мінімальними наборами даних із використанням генеративного ландшафту

Результати експериментальної частини роботи, включаючи експериментальну перевірку методів навчання на генеративному ландшафті дозволяють визначити *інформаційну технологію навчання з мінімальними даними з використанням генеративного ландшафту*, зведенням та формалізацією процесів, описаних у теоретичній та експериментальній частинах роботи.

Інформаційна технологія включає та детально визначає стадії створення інформативних генеративних представлень, визначення їх структури, побудови навчальних наборів на основі структури генеративного ландшафту та ітеративного навчання класифікаторів концептів, розроблених та перевірених у роботі, та повністю визначає та описує процес навчання з використанням генеративного ландшафту, розроблений і підтверджений експериментальними результатами роботи.

Загальна схема процесу обробки інформації технології навчання з мінімальними навчальними наборами наведена на рис. 4.10.

Рис. 4.10. Загальна блок-схема обробки інформації при навчанні з використанням генеративного ландшафту представлень

Запропонована технологія поєднує в єдиний загальний процес такі етапи:

– підготовки та попередньої обробки даних;

– генеративного навчання та створення генеративних представлень;

– виявлення структури генеративних представлень методами неконтрольованої кластеризації;

– створення навчальних наборів класів та концептів на основі виявленої генеративної структури представлення;

– навчання класифікаторів класів і концептів на основі генеративного ландшафта в ітераціях навчання.

Таким чином, технологія узагальнює процеси навчання на основі генеративної структури представлень для даних різних типів та сфер застосування. Технологія вимагає мінімальних наборів відомих навчання, до окремих зразків відомих класів. Детальна діаграма процесу обробки даних технології представлено на рис. 4.11.

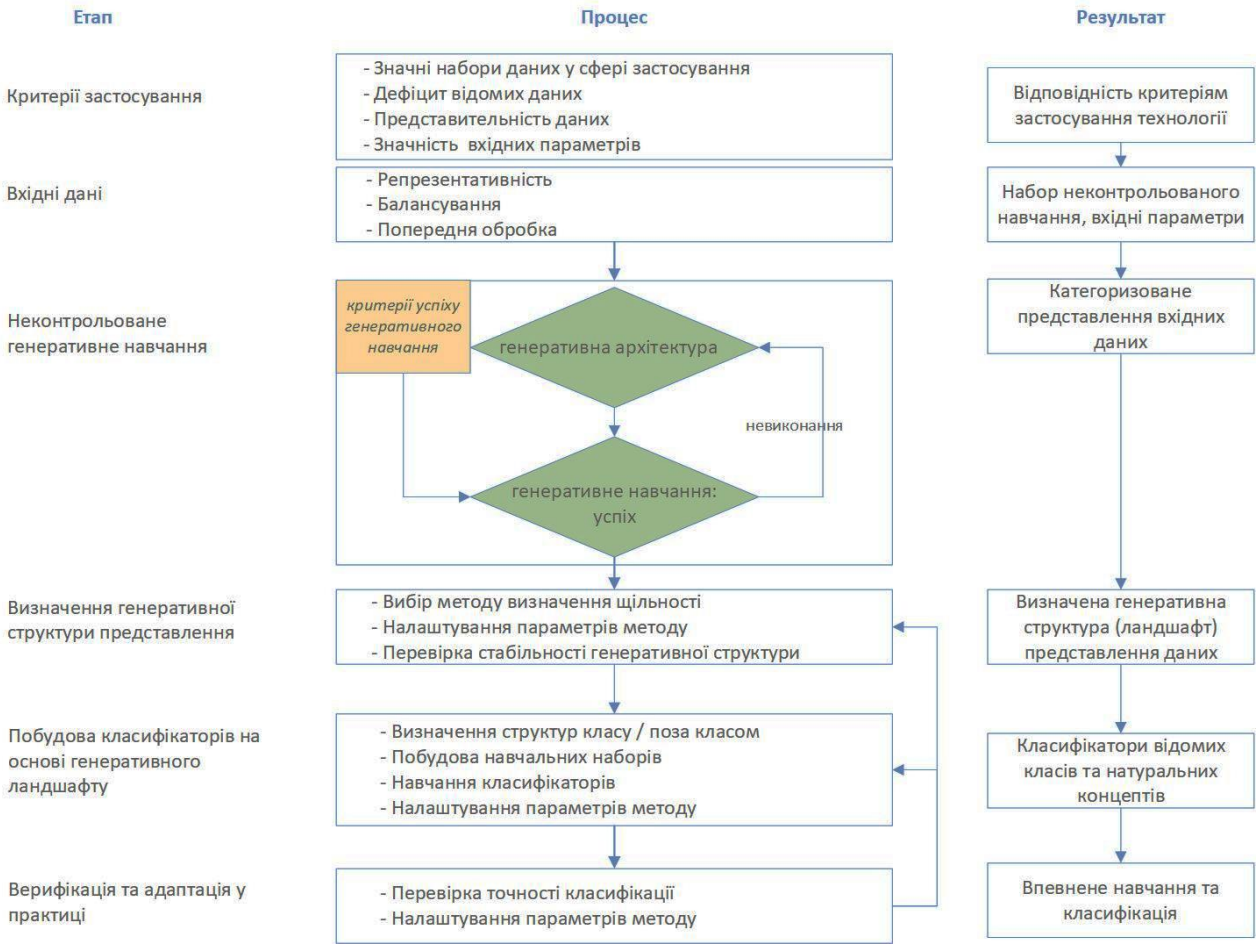


Рис. 4.11. Інформаційна технологія навчання з використанням генеративного ландшафту представлень

При перевірці з даними Інтернет та зображень застосування технології дозволило досягти точності, порівняної з опублікованими результатами з мінімальними наборами даних (30 – 100 відомих зразків, проти тисяч при використанні традиційних методів контрольованого навчання). При цьому результат навчання не залежить від якості навчальних наборів отриманих з інших мереж. Навіть із одиничними відомими зразками дозволяє досягти значної точності розпізнавання деяких класів.

Технологія дозволяє досягти високого рівня точності розпізнавання натуральних категорій без відомих даних і може застосовуватися та мати важливе значення при аналізі даних невідомої структури та/або походження.

Методи, використані в запропонованій технології були перевірені з наборами даних різних типів та галузях застосування: дані Інтернет та зображень місцевості, що підтверджує потенціал використання технології у різних галузях застосування з різними типами даних.

ВИСНОВКИ

В роботі досліджувалися представлення даних пакетів трафіку Інтернет тощо, створені моделями неконтрольованого генеративного навчання та методи навчання на основі структури генеративних представлень з мінімальними наборами навчальних даних, розроблені на їх основі.

Розв'язано задачу обґрунтування, розробки та програмної реалізації моделей неконтрольованого навчання, які дозволяють створювати структуровані інформативні представлення складних даних в процесі неконтрольованого навчання та їх використання в інноваційних методах навчання з використанням визначеної структури даних представлень. У роботі проведено теоретичні та експериментальні дослідження та досягнуто наступних результатів.

1. Удосконалено математичну модель розподілів даних пакетів трафіку Інтернет у генеративних представленнях завдяки теоретичному аналізу зв’язку між геометричними властивостями розподілів і ефективністю навчання генеративних моделей, що дозволило забезпечити теоретичну основу методів навчання на основі генеративного ландшафту представлень та підвищити ефективність виділення інформативних факторів даних.

2. Розроблено та реалізовано оригінальні нейромережеві моделі глибокого неконтрольованого генеративного навчання типу автоенкодера з різким стиском латентного шару в середовищі програмування Python що дозволило створювати інформативні представлення даних пакетів трафіку Інтернет низької розмірності зі збереженням ключових характеристик розподілу даних.

3. Доведена теорема про категоризацію генеративних представлень (при певних припущеннях та умовах), що на підставі методів варіаційного аналізу забезпечує теоретичне обґрунтування методів навчання на основі генеративної структури представлень даних Інтернет та інших типів.

4. Отримано, оброблено та підготовлено до використання навчальні набори даних Інтернет та візуальних даних у середовищі програмування Python що дозволило стабільно вивчати моделі та аналізувати генеративні представлення даних пакетів трафіку Інтернет.

5. Знайдено розв'язання задачі визначення структури щільності генеративних представлень неконтрольованими методами без вимоги відомих даних навчання засноване на теоретичних основах процесу створення структурованих генеративних представлень та застосування методів неконтрольованої кластеризації, багатовимірних гістограм що дозволило підвищити стабільність розпізнавання структур щільності генеративних представлень від рівня теоретичної можливості до рівня стабільного використання в інформаційній технології з успішністю генеративного навчання та визначення структури щільності вище 80%.

6. Розв'язано задачу стабільного навчання розпізнавання відомих класів даних пакетів трафіку Інтернет на основі структури щільності генеративних представлень, що за рахунок розроблених концептуальної та математичної моделі, методів теорії генеративних представлень та теореми про категоризацію генеративних представлень забезпечують: достатню точність розпізнавання, на рівні відомих методів контрольованого навчання; зменшення залежності від джерела отримання навчальних даних; зменшення обсягу навчальних даних, у 10–100 разів в порівнянні з відомими методами контрольованого навчання. Отримано програмну реалізацію методів в середовищі програмування Python на основі сучасних пакетів та бібліотек глибокого навчання, машинного навчання, аналізу та обробки даних.

7. Розроблено повністю неконтрольовані методи розпізнавання натуральних концептів даних пакетів трафіку Інтернет, без вимог даних навчання відомих класів на основі структури щільності генеративних представлень, з точністю на рівні відомих методів контрольованого навчання. Отримано програмну реалізацію методів в середовищі програмування Python на основі сучасних пакетів та бібліотек глибокого навчання, машинного навчання, аналізу та обробки даних.

8. Визначено, формалізовано та виконано концептну реалізацію інформаційної технології розпізнавання класів трафіку Інтернет класу архітектур глибокого навчання на основі запропонованих методів визначення структури щільності генеративних представлень, що дозволяє автоматизувати процес навчання та

використання запропонованих моделей і методів розпізнавання класів трафіку Інтернет з мінімальними навчальними наборами при збереженні або підвішенні точності розпізнавання під час використання з данними загальних мереж Інтернет в порівнянні з методами контрольованого навчання. Отримано програмне концептне виконання інформаційної технології навчання розпізнаванню даних пакетів трафіку Інтернет у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання.

Запропоновані в роботі методи мають низку переваг порівняно з відомими підходами в областях: точності розпізнавання відомих класів, досягаючи і навіть перевершуючи результати сучасних методів контрольованого навчання при класифікації даних загального джерела; вимоги до навчальних наборів, необхідні розміри наборів зменшено в 10–100 разів; стабільності навчання, успіх навчання не залежить від джерела навчального набору; універсальності застосування, ефективність методів підтверджена з наборами даних суттєво різних за характером та джерелом.

За результатами роботи розв'язано поставлену задачу впевненого навчання розпізнаванню даних пакетів трафіку Інтернет та інших типів із зменшеними наборами навчальних даних відомих класів та суттєвого зменшення залежності успіху навчання від джерела даних. Розроблені у роботі методи та інформаційна технологія навчання на основі генеративної структури інформативних представлень показали значну ефективність при навчанні з мінімальними наборами навчальних даних. При цьому процес навчання є гнучким та ітеративним з можливістю навчання у міру емпіричного досвіду та/або наявності навчальних даних. Успішне застосування методів неконтрольованого навчання для розв'язання задачі роботи, а також експериментальні результати роботи дають підставу розраховувати, що запропоновані методи та перевірені в роботі можуть бути використані в ширшому колі завдань та додатків із різними типами даних і задач розпізнавання.

Подальший розвиток запропонованих у роботі методів може включати напрямки:

– розробки методів автоматичної генерації потенційно необмежених наборів маркованих даних на основі генеративної кластерної структури представлень даних Інтернет та інших типів, яка може бути визначена методами, розробленими в роботі;

– розробка нових методів виявлення новизни на основі натуральних концептуальних структур в представленнях, створених моделями генеративного навчання, які можуть бути визначені методами, розробленими в роботі;

– можливі розвитки методів роботи у розробці систем навчання здатних до мимовільного, ітеративного та гнучкого навчання безпосередньо зі спостережень за довкіллям з мінімальними наборами початкових даних навчання.

Одним із можливих напрямів майбутніх досліджень є створення ефективних генеративних представлень складніших даних. Це може вимагати створення складніших генеративних архітектур. Наприклад, у моделей, використовуваних у роботі, простір представлення створено активаціями нейронів єдиного шару нейронної мережі. Це одна з найпростіших можливих архітектур і цілком ймовірно, що у разі більш складних даних вона має бути вдосконалена. Можливі варіанти складніших архітектур включають шари з обмеженням розрядження [91], представлення створені багатьма шарами тощо. Розповсюдження та розвиток методів визначення характеристичних структур у таких більш складних представленнях та навчання на генеративному ландшафті таких представлень це також суттєвий напрямок подальшої роботи.

У практичному напрямку, методи та алгоритми навчання на генеративному ландшафті представлень можуть бути уточнені та оптимізовані. Наприклад, більш ефективні стратегії побудови латентних навчальних наборів можуть призвести до подальшого підвищення ефективності методів, розроблених у роботі.

Отже, перспективними напрямками досліджувань можуть бути паралелі та загальні принципи у навчанні штучних та біологічних систем обробки інформації. Результати теоретичної та експериментальної частин роботи вказують, що аналогії та паралелі можуть йти далі поверхневої подібності та відображати загальні принципи та аналогічні процеси та архітектури обробки інформації.

Ці питання відкривають можливість цікавих нових напрямів спільних досліджень з боку машинного навчання та нейробіології, включаючи такі напрями, як формування абстрактних категорій та заснованих на них поведінки [105, 106]; аналогії та паралелі між нейромережевими системами біологічних та штучних систем [107, 108]; можлива роль методів та процесів неконтрольованого навчання у формуванні концептуальних структур [109, 110] та інші.

Взагалі, як показали результати роботи, дослідження та використання інформативних представлень моделей генеративного навчання можуть мати значні перспективи як у теоретичних дослідженнях так і в практичних застосуваннях.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. He K., Zhang X., Ren S., Sun J.: Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. arXiv:1502.01852 2015.

2. Kavukcuoglu K., Sermanet P., Boureau Y.L., Gregor K., Matheu M., LeCun Y.: Learning convolutional feature hierarchies for visual recognition. In: 23rd International Conference in Neural Information Processing Systems, Vancouver, Canada, 1 1090–1098 2010.

3. Prystavka P., Rogatyuk A.: The mathematical foundations of foreign object recognition in the video from unmanned aircraft. Proceedings of National Aviation University, 3 (64) 133–139 2015.

4. Prystavka P., Cholyshkina O., Dolgikh S., Karpenko D.: Automated object recognition system based on convolutional autoencoder. In: 10th International Conference on Advanced Computer Information Technologies (ACIT-2020), Deggendorf, Germany, 830–833 2020.

5. Lunga D., Prasad S., Crawford M., Ersoy O.: Manifold-learning-based feature extraction for classification of hyperspectral data: a review of advances in manifold learning. IEEE Signal Processing Magazine, 31(1) 55–66 2014.

6. Gondara, L.: Medical image denoising using convolutional denoising autoencoders. In: 16th IEEE International Conference on Data Mining Workshops (ICDMW), Barcelona, Spain, 241–246 2016.

7. Dolgikh, S.: Unsupervised clustering in epidemiological factor analysis. The Open Bioinformatics Journal, 14 (1) 63–72 2021.

8. A P, S. C., Lauly, S., Larochelle, H., Khapra, M.M.., Ravindran B. et al.: An autoencoder approach to learning bilingual word representations. In: 27th International Conference on Neural Information Processing Systems (NIPS'14), Montreal, Canada, 2 1853–1861 2014.

9. F. Gingoli et al.: Picking up the truth from the ground for Internet traffic. ACM Computer Communication Review, 39 (5) 12–18 2009.

10. Seddigh N., Nandy B., Bennet D., Ren Y., Dolgikh S. et al.: A framework & system for classification of encrypted network traffic using Machine Learning. In: 15th International Conference on Network and Service Management (CNSM), Halifax Canada 2019, 1–5 2019.

11. Zhou C., Paffenroth R.C.: Anomaly detection with robust deep autoencoders. In: 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, Canada, 665–674 2017.

12. Gnatyuk S., Hu Z., Sydorenko V., Marek A. et al.: Critical aviation information systems: identification and protection. In: Cases on Modern Computer Systems in Aviation, IGI Global, 341–366 2019.

13. Dychka, I., Chernyshev, D., Tereikovskyi, I., Tereikovska, L., Pogorelov, V.: Malware detection using Artificial Neural Networks. In: Hu, Z., Petoukhov, S., Dychka, I., He, M. (eds) Advances in Computer Science for Engineering and Education II. ICCSEEA 2019. Advances in Intelligent Systems and Computing, vol 938. Springer, Cham 2019.

14. Hubskyi, O., Babenko, T., Myrutenko, L., Oksiiuk, O.: Detection of SQL injection attack using Neural Networks. In: Shkarlet, S., Morozov, A., Palagin, A. (eds) Mathematical Modeling and Simulation of Systems (MODS'2020). Advances in Intelligent Systems and Computing, 1265. Springer, Cham. 2020.

15. Rasch R., A. Kott A., Forbus K.D.: Incorporating AI into military decision making: an experiment. IEEE Intelligent Systems, 18 (4) 18–26 2003.

16. Araujo, T., Helberger, N., Kruikemeier, S. et al.: In AI we trust? Perceptions about automated decision-making by artificial intelligence. AI & Society 35 611–623 2020.

17. Palko D., L. Myrutenko L., T. Babenko T., Bigdan, A.: Model of Information Security critical incident risk assessment. In: 2020 IEEE International Conference on Problems of Infocommunications, Science and Technology Kharkiv, Ukraine, 157-161 2020.

18. Pimentel M., Clifton D., Clifton L., Tarassenko L.: A review of novelty detection. Signal Processing, 99 215–249 2014.

19. Albertini M. K., de Mello R.F.: A self-organizing neural network approach to novelty detection, ACM Symposium on Applied Computation (SAC), Seoul, South Korea, 462–466 2007.

20. Campbell, M.: Knowledge discovery in deep blue. Communications of the ACM, 42 (11) 65–67 1999.

21. Silver D., Shrittwieser J., Simonyan K., Antonoglou I., Huang A. et al.: Mastering the game of Go without human knowledge. Nature, 550 354–359 2017.

22. Team uses AI to complete Beethoven's unfinished masterpiece. NPR News, 2.10.2021.

23. Shi J., J. Xu J., Y. Yao Y., Xu B.: Concept learning through deep reinforcement learning with memory augmented neural networks. Neural Networks, 110 47–54 2019.

24. Hu, Z., Tereikovskyi, I., Korystin, O., Mihaylenko, V., Tereikovska, L.: Two-layer perceptron for voice recognition of speaker's identity. In: Hu, Z., Petoukhov, S., Dychka, I., He, M. (eds) Advances in Computer Science for Engineering and Education III. ICCSEEA 2020. Advances in Intelligent Systems and Computing, 1247. Springer, Cham 2020.

25. Alshammari R., Zincir-Heywood N.: Can encrypted traffic be identified without port numbers, IP addresses and payload inspection? Journal of Computer Networks, 55 (6) 1326–1350 2011.

26. T. Wiradinata T. and A. Paramita A.: Clustering and feature selection technique for improving Internet traffic classification using K-NN. Journal of Advances in Computer Networks, 4 (1) 24–27 2016.

27. R. Bar-Yanai R., M. Langberg M., D. Peleg D., L. Roditty L.: Realtime classification for encrypted traffic. In: 9th International Conference on Experimental Algorithms, Naples, 373–385 2010.

28. T. Bujlow T., V. Carela-Español V. Barlet-Ros P.: Comparison of Deep Packet Inspection (DPI) tools for traffic classification. Technical report, UPC-DAC-RR-CBA-2013-3 ed.3 Universitat Politècnica de Catalunya, Barcelona, Spain, June 2013.

29. Ubik, S., Zeidl, P.: Evaluating application–layer classification using a Machine Learning technique over different high-speed networks. In: 5th International Conference on Systems and Networks Communications, 387–391 2010.

30. Kim H., Claffy KC., Fomenkov M. et al. Internet traffic classification demystified: myths, caveats, and the best practices. In: 2008 ACM CoNEXT Conference, 1–12 2008.

31. Hinton, G., Sejnowski, T.: Unsupervised Learning: Foundations of Neural Computation. MIT Press, 1999.

32. Hornik K., Stinchcombe M., White H.: Multilayer feedforward neural networks are universal approximators. Neural Networks, 2 (5) 359–366 1989.

33. Горелик А.Л., Скрипкин В.А.: Методы распознования. Москва "Высшая школа", 1984.

34. Higgins I., Matthey L., Glorot X., Pal A., Uria B. et al.: Early visual concept learning with unsupervised deep learning. arXiv 1606.05579 [cs.LG] 2016.

35. Le Q.V., Ranzato M.A., Monga R., Devin M., Chen K. et al.: Building high level features using large scale unsupervised learning. arXiv 1112.6209 [cs.LG] 2012.

36. Dolgikh, S.: Spontaneous concept learning with deep autoencoder. International Journal of Computational Intelligence Systems, 12 (1) 1–12 2018.

37. Fischer A., Igel C.: Training restricted Boltzmann machines: an introduction. Pattern Recognition, 47 25–39 2014.

38. Hinton G.E., Osindero S., Teh Y.W.: A fast learning algorithm for deep belief nets. Neural Computation, 18 (7) 1527–1554 2006.

39. Kriegel H.-P., Kroger P., Sander J., Zimek A.: Density-based clustering. WIREs Data Mining and Knowledge Discovery, 1 (3) 231–240 2011.

40. Бодянский Е. В., Дейнеко А. А., Куценко Я. В.: Ядерная кластеризация на основе обобщенной регрессионной нейронной сети и самоорганизующейся карты Т. Кохонена. Інформаційно-керуючі системи на залізничному транспорті, 3. 15–22 2015.

41. Kohonen T., Honkela T.: Kohonen network. Scholarpedia, 2 (1) 1568 2007.

42. Бодянский Е. В., Дейнеко А. А., Куценко Я. В.: Ядерная самоорганизующаяся карта на основе радиально-базисной нейронной сети. Електротехнічні та комп'ютерні системи, 20 97–105 2015.

43. Bengio, Y.: Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2 (1) 1–127 2009.

44. Welling M., and D.P. Kingma D.P.: An introduction to variational autoencoders. Foundations and Trends in Machine Learning, 12 (4) 307–392 2019.

45. Le, Q.V.: A tutorial on deep learning: autoencoders, convolutional neural networks and recurrent neural networks. Stanford University, 2015.

46. Creswell A., White T., Dumoulin V., Arulkumaran K., Sengupta B. et al.: Generative adversarial networks: an overview. IEEE Signal Processing Magazine, 35 (1) 53–65 2018.

47. Partaourides, H., Chatzis, S.P.: Asymmetric deep generative models: Neurocomputing, 241 90–96 2017.

48. Coates A., Lee H., Ng A.Y.: An analysis of single-layer networks in unsupervised feature learning. In: 14th International Conference on Artificial Intelligence and Statistics (AISTATS), Lauderdale, USA, 15 215–223 2011.

49. Vapnik, V.N.: The Nature of Statistical Learning Theory. 2nd ed., Springer Verlag 2000.

50. Beyer L., Zhai X., Oliver A., Kolesnikov A.: S4L: Self-Supervised Semi-Supervised Learning. IEEE/CVF International Conference on Computer Vision (ICCV-2019), 1476–1485 2019.

51. Zhou X., Belkin M.: Semi-supervised learning. In: Academic Press Library in Signal Processing, Elsevier, 1 1239–1269 2014.

52. Kaelbling, L.P., Littman, M.L., Moore, A.W.: Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research, 4 237–285 1996.

53. Bengio Y., Courville A., Vincent, P.: Representation Learning: a review and new perspectives. arXiv:1206.5538 [cs.LG] 2014.

54. Hotelling, H.: Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24 417–441 2013.

55. Werman M., Peleg S., Rosenfeld A.: A distance metric for multidimensional histograms, Computer Vision, Graphics and Image Processing, 32(3) 328–336 2005.

56. Parzen, E.: On Estimation of a probability density function and mode. The Annals of Mathematical Statistics, 33 (3) 1065–1076 1962.

57. Rosenblatt, M.: Remarks on some nonparametric estimates of a density function. The Annals of Mathematical Statistics, 27 (3) 832–837 1956.

58. Epanechnikov, V.A.: Non-parametric estimation of a multivariate probability density. Theory of Probability and Its Applications, 14 153–158 1969.

59. Silverman, B.W.: Density Estimation for Statistics and Data Analysis. London: Chapman & Hall/CRC, 45 1996.

60. Everitt, B.: Cluster analysis. Chichester, West Sussex, U.K (Wiley) 2011.

61. Dunn, J.: Well separated clusters and optimal fuzzy partitions. Journal of Cybernetics, 4 95–104 1994.

62. Estivill-Castro, V.: Why so many clustering algorithms – A Position Paper. ACM SIGKDD Explorations Newsletter, 4 (1) 65–75 2002.

63. Sander, J., Ester, M., Kriegel, H-P., Xu, X.: Density-based clustering in spatial databases: the algorithm GDBSCAN and its Applications. Data Mining and Knowledge Discovery, Berlin Springer-Verlag, 2 (2) 169–194 1998.

64. Fukunaga K., Hostetler L.D.: The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on Information Theory, 21 (1) 32–40 1975.

65. Menardi G., Azzalini A.: An advancement in clustering via non-parametric density estimation. Statistics and Computing, 24 753–767 2014.

66. Zeng N., H. Zhang H., Song B., Liu W., Li Y. et al.: Facial expression recognition via learning deep sparse autoencoders. Neurocomputing, 273 643–649, 2018.

67. Kingma D., Ba J.: Adam: a method for stochastic optimization. arXiv:1412.6980v8 2015.

68. Huang G., Sun Y., Liu Z., Sedra D., Weinberger K.: Deep networks with stochastic depth, arXiv:1603.09382 2016.

69. Cybenko, G.: Approximation by superposition of sigmoidal function. Mathematics of Control, Signals, and Systems, 2 303–314 1989.

70. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60 (6) 84–90 2017.

71. Spall, J.C.: Introduction to stochastic search and optimization: estimation, simulation, and control. Hoboken, New Jersey: Wiley 2013.

72. Hinton G.E.: Training products of experts by minimizing contrastive divergence. Neural Computation, 14 (8) 1771–1800 2002.

73. Hopfield, J. J.: Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79 (8) 2554–2558 1982.

74. Bodyanskiy Ye., Shafronenko A., Volkova V.: Adaptive fuzzy probabilistic clustering of incomplete data. International Journal "Information Models and Analyses" 2 (2), 37–49 2013.

75. Hassabis D., Kumaran D., Summerfield C., Botvinick M.: Neuroscience inspired Artificial Intelligence, Neuron 95 (2) 245–258 2017.

76. Yoshida T., K. Ohki K.: Natural images are reliably represented by sparse and variable populations of neurons in visual cortex. Nature Communications, 11 872 2020.

77. Bao X., Gjorgiea E., Shanahan L.K., Howard J. D., T. Kahnt T., Gottfried, J.A.: Grid-like neural representations support olfactory navigation of a two-dimensional odor space. Neuron, 102 (5) 1066–1075, 2019.

78. Rice, J.: Mathematical statistics and data analysis. Belmont, CA Brooks/Cole Cengage Learning, 138 2007.

79. Cover, T.M., Thomas, J.A.: Elements of information theory. John Wiley & Sons, Ltd., 13–55. 2005.

80. Robins, J., Wasserman, L.: Conditioning, likelihood, and coherence: a review of some foundational concepts. Journal of American Statistical Association, 95 (452) 1340–1346 2000.

81. Variational inference. Algorithms for Inference, Massachusetts Institute of Technology 2014.

82. S. Mandt, Hoffman M.D., Blei D.M.: Stochastic gradient descent as approximate Bayesian inference. Journal of Machine Learning Research, 18 1–35 2017.

83. Ranzato M.A., Boureau Y-L., S. Chopra S., LeCun Y.: A unified energy-based framework for unsupervised learning. In: 11th International Conference on Artificial Intelligence and Statistics, 2 371–379, 2007.

84. Krause O., Fischer A., Igel C.: Algorithms for estimating the partition function of restricted Boltzmann machines. Artificial Intelligence, 278 103195 2020.

85. Stinson, P: Generative modeling and inference in directed and undirected neural networks. Columbia University, 2020.

86. Raschka, S.: Python machine learning: machine learning and deep learning with python, scikit-learn, and tensorflow. Packt Publishing, Birmingham, 37–38 2019.

87. Keras: the Python deep learning API. Online: https://keras.io

88. Prystavka P., Dolgikh S., Cholyshkina O., Kozachuk O.: Latent representations of terrain in aerial image classification. In: 17th International Conference on ICT in Education, Research and Industrial Applications, Kherson, Ukraine 2021 CEUR-WS.org, 3013 86–95, 2021.

89. Dolgikh, S: Topology of conceptual representations in unsupervised generative models. In: 26th International Conference on Information Society and University Studies (IVUS 2021) Kaunas, CEUR-WS.org, 2915 150–157 2021.

90. Rumelhart, D. E., Hinton, G. E., Williams, R. J.: Learning representations by back-propagating errors. Nature, 323 533–536 1996.

91. Bach F., Jenatton R., Mairal J., Obozinski G.: Optimization with sparsity-inducing penalties. Foundations and Trends in Machine Learning, 4 (1) 1 – 106 2011.

92. Nielsen, M. Neural Networks and Deep Learning. Online: http://neuralnetworksanddeeplearning.com/ 2019.

93. Dolgikh, S.: Categorization in unsupervised generative self-learning systems. International Journal of Modern Education and Computer Science, 3 68–78 2021.

94. Eastman, P.: Introduction to Statistical Mechanics. Stanford University 2014.

95. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 46 (3) 175–185 1992.

96. Cortes, C., Vapnik V.N.: Support vector networks. Machine Learning, 20 (3) 273–297 1995.

97. Dolgikh, S: Categorized representations and general learning. In: 10th International Conference «Theory and Application of Soft Computing, Computing with Words and Perceptions ICSCCW-2019», Prague, August 27-28 2019, 93–100 2019.

98. Hearley, K.B., Oakes, L.M.: Experience and distribution of attention: Pet exposure and infants' scanning of animal images. Journal of Cognitive Development, 16 (1) 11–30 2015.

99. Matplotlib – Visualization with Python. Online: https://matplotlib.org.

100. WITS passive datasets, Waikato University, Waikato, New Zealand Online: https://wand.net.nz/wits, last accessed 2020/02/10.

101. Alshammari R., Zincir-Heywood A.: Investigating two different approaches for encrypted traffic classification. In: 6th Annual Conference on Privacy, Security and Trust, Fredericton, Canada, 156–166 2007.

102. Приставка П. и др.: Навчальний набір даних “Аерозйомка”. Кафедра Прикладной математики, Факультет информационных технологий Национальный авиационный университет (НАУ) онлайн https://drive.google.com/file/d/1BAmSRbYUyCnrPYn-jpHI7l_6qsNmc9o6/view 2018-2021 2021.

103. Snell, J., Swersky K., Zemel. R.S.: Prototypical networks for few-shot learning. arXiv:1703.05175 2017.

104. Dolgikh, S.: Unsupervised generative learning and native explanatory frameworks. 11th International Conference on Advanced Computer Information Technologies (ACIT-2021) Deggendorf, Germany, 748–752 2021.

105. Roth G., Dicke U.: Evolution of the brain and intelligence. Trends in Cognitive Science, 9 (5) 250–257 2005.

106. Feinberg T.E., Mallatt, J.: The nature of primary consciousness. A new synthesis. Consciousness and Cognition, 43 113–127, 2016.

107. Garm, A., Poussart, Y., Parkefelt, L., Ekström, P., Nilsson, D-E.: The ring nerve of the box jellyfish Tripedalia cystophora. Cell and Tissue Research, 329 (1) 147–157 2007.

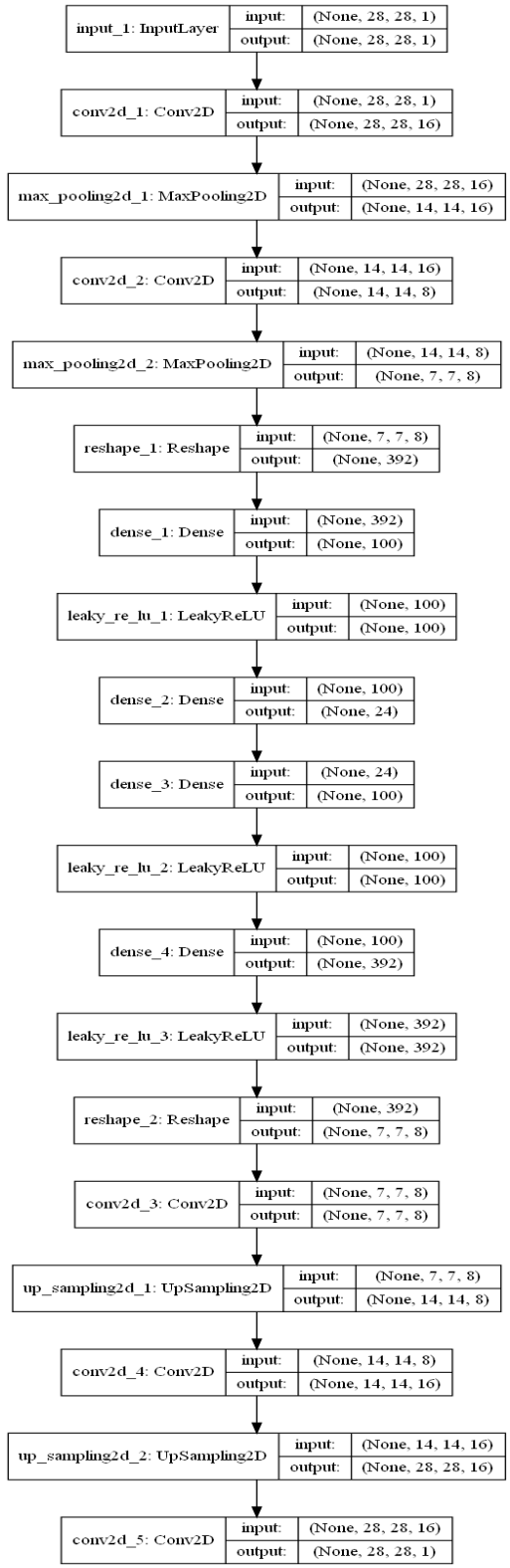
108. Nevens J., P. Van Eecke P. Beuls K.: From continuous observations to symbolic concepts: a discrimination-based strategy for grounded concept learning. Frontiers in Robotics and AI, 7 84 2020.

109. Dolgikh, S.: Synchronized conceptual representations in unsupervised generative learning. 13th International Conference on Soft Computing and Pattern Recognition (SoCPaR-2021), MirLabs USA, December 15-17 2021, Lecture Notes in Networks and Systems, Springer, 417 23–32 2022.

110. Wang Q., Young S., Harwood A., Ong C.S.: Discriminative concept learning network: reveal high-level differential concepts from shallow architecture. In: International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 1–9 2015.

ДОДАТОК А

Діаграма архітектури моделей глибокого автоенкодера зі стисненням



ДОДАТОК Б

Акт впровадження

АКТ

впровадження у науково-дослідну діяльність
результатів дисертаційної роботи Долгих Сергія Миколайовича
«Інформаційна технологія розпізнавання мережевих даних Інтернет на основі генеративних нейромережевих моделей» на здобуття кандидата технічних наук

Комісія у складі: голова – науковий керівник Науково-дослідної лабораторії протидії кіберзагрозам в авіаційній галузі д.т.н., проф. Гнатюк С.О., молодший науковий співробітник Дорожинський С.А., молодший науковий співробітник Поліщук Ю.Я. склали даний акт про те, що результати дисертаційної роботи Долгих Сергія Миколайовича впровадженні у науково-дослідну діяльність та використовуються в Науково-дослідній лабораторії протидії кіберзагрозам в авіаційній галузі, у рамках виконання Науково-дослідної роботи «Інтелектуалізована система захищеного передавання пакетних даних на базі розвідувально-пошукового безпілотного літального апарату» (№ держреєстрації: 0122U002361).

№ з/п	Назва роботи, що впроваджується	Форма впровадження	Ефективність від впровадження
	1	2	3
1.	Програмна реалізація методу навчання розпізнавання відомих класів даних пакетів трафіку Інтернет з використанням структури щільності генеративних представлень у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання.	Наукова стаття в рамках НДР 421-ДБ22	Дозволяє стабільне навчання розпізнавання відомих класів даних Інтернет з мінімальними даними класів на рівні сучасних методів контрольованого навчання при використанні загальних мережевих даних.
2.	Програмна реалізація методу навчання розпізнавання натуральних концептів даних пакетів трафіку Інтернет з використанням структури щільності генеративних представлень, у програмному середовищі Python з використанням пакетів та бібліотек машинного навчання та обробки даних.	Наукова стаття в рамках НДР 421-ДБ22	Дозволяє стабільне навчання розпізнавання натуральних типів (концептів) даних Інтернету без вимоги відомих даних навчання.

Голова комісії,

Науковий керівник
НДЛ протидії кіберзагрозам
в авіаційній галузі — Сергій ГНАТЮК

Члени комісії:

м.н.с. НДЛ протидії кіберзагрозам
в авіаційній галузі — Сергій ДОРОЖИНСЬКИЙ

м.н.с. НДЛ протидії кіберзагрозам
в авіаційній галузі — Юлія ПОЛІЩУК