

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**Одеський національний технологічний університет**  
**Університет Інформатики і прикладних знань, м.Лодзь, Польща**  
**Національний технічний університет України «Київський**  
**політехнічний інститут»**  
**Навчально-науковий інститут комп'ютерних систем і технологій**  
**«Індустрія 4.0» ім. П.М. Платонова**

**XXIII Всеукраїнська науково-технічна конференція**  
**молодих вчених, аспірантів та студентів**

**«СТАН, ДОСЯГНЕННЯ ТА ПЕРСПЕКТИВИ**  
**ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ»**

*Матеріали конференції*



Одеса

**20-21 квітня 2023 р.**

Стан, досягнення та перспективи інформаційних систем і технологій / Матеріали XXIII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів. Одеса, 20-21 квітня 2023 р. - Одеса, Видавництво ОНТУ, 2023 р. – 449 с.

Збірник включає матеріали доповідей учасників конференції, які об'єднані за тематичними напрямками конференції.

Збірник буде корисним як для фахівців і працівників фірм, зайнятих в області ІТ, так і для викладачів, магістрів і студентів вищих навчальних закладів, які навчаються за напрямками і спеціальностями програмного забезпечення, обчислювальної техніки і автоматизованих систем, прикладної математики та обробки інформації, буде корисним професіоналам з комп'ютерного моделювання та розробки комп'ютерних ігор.

Результати досліджень у збірнику представляють собою своєрідний зріз сучасного стану справ в перерахованих галузях знань, який може допомогти як фахівцям, так і студентам університетів скласти загальну картину розвитку інформаційних технологій та пов'язаних з ними питань.

Наукові праці згруповані за напрямками роботи конференції та наведені в алфавітному порядку прізвищ авторів.

Матеріали (тези доповідей) друкуються в авторській редакції. Відповідальність за якість та зміст публікацій несе автор.

Матеріали подано українською та англійською мовами.

Редактор збірника Котлик С.В.

|  |     |
|--|-----|
| <b>Яковенко М.І., Корнієнко Ю.К. (Одеський національний технологічний університет)</b>   |     |
| <b>Розділ 5: Комп'ютерні телекомунікаційні мережі та технології</b>  | 300 |
| 1. Алгоритм попередньої обробки зображень для алгоритму QOI. <b>Доценко Д., Крайник Я. М. (Чорноморський національний університет імені Петра Могили)</b>  | 300 |
| 2. Аналіз сучасних архітектур GPU. <b>Завальнюк Є.К., Романюк О.Н., Снігур А.В., Шевчук Р. П. (Вінницький національний технічний університет, Західноукраїнський національний університет)</b>   | 302 |
| 3. Дослідження інструментальних засобів розробки програмного забезпечення для електронної комерції. <b>Клівчук Д.К. (Волинський національний університет імені Лесі Українки)</b>  | 304 |
| 4. Основні принципи роботи сучасних навігаційних систем. <b>Наголюк Д. О. (Донецький національний університет імені Василя Стуса)</b>  | 305 |
| 5. Сучасний стан і перспективи розвитку глобальних мереж інфокомунікацій. <b>Нєнов О. Л. (Одеський національний технологічний університет)</b>   | 307 |
| 6. Розробка захищеної корпоративної локальної мережі. <b>Рижков М.С., Сахарова С.В., Нєнов О.Л. (Одеський національний технологічний університет)</b>  | 309 |
| 7. Вимірювання параметрів оптичних компоненті мережі. <b>Сахарова С.В., Рибалов Б.О. (Одеський національний технологічний університет)</b>   | 311 |
| 8. Аналіз сучасних HTML-редакторів. <b>Терешко Д. С., Романюк О. Н., Романюк О. В. (Вінницький національний технічний університет)</b>   | 313 |
| 9. Оптимізація роботи алгоритму розподілу навантаження між серверами в мережі шляхом поєднання Rest і Soap. <b>Тоха В.В. (Вінницький національний технічний університет)</b>   | 314 |
| 10. Автоматизація процесу перебудови характеристик частотно-залежних компонент при обробці сигналів датчиків у робототехнічних системах. <b>Чумаченко Н.К., Бадерко І.В., Ситніков В.С. (Національний університет "Одеська політехніка")</b>                 | 317 |
| 11. Розробка мережевого фільтра на базі міні комп'ютера Raspberry Pi. <b>Шевчук М.С., Іванова Л.В., Сахарова С.В. (Одеський національний технологічний університет, Одеський технічний фаховий коледж ОНТУ)</b>  | 319 |
| <b>Розділ 6: Штучний інтелект і автоматизація робототехнічних систем</b>   | 322 |
| 1. Terms clustering hybrid service with word2vec, k-means, and majorclust algorithms for knowledge processing systems with cloud-based architecture. <b>Malakhov K.S. (Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine)</b> | 322 |
| 2. Safety and ethics in the use of automated systems. <b>Rysbek Akerke. (University "Turan", Kazakhstan)</b>   | 324 |
| 3. Exploring extramae: a scalable self-supervised approach to synthetic time series generation. <b>Аблець А. В. (Криворізький національний університет)</b>  | 325 |
| 4. Синтетичні набори даних в штучному інтелекті. <b>Антонова А.Р., Юрченко І.С. (Одеський національний технологічний університет)</b>  | 326 |
| 5. Використання штучного інтелекту у 3D-модельованні. <b>Бойцова М.П., Бойцова О.С. (Одеський національний технологічний університет)</b>  | 328 |
| 6. Розробка сайту психологічної допомоги на базі штучного інтелекту . <b>Босенко Л.С., Болтач С.В. (Одеський національний технологічний університет)</b>   | 330 |
| 7. Програма для відстеження пози та рухів людини на основі аналізу відео потоку з використанням MediaPipe. <b>Вишневський В., Рябенський В., Вишневський В. (Національний Університет Кораблебудування ім. адмірала Макарова)</b>                            | 332 |
| 8. Використання штучного інтелекту в освіті: переваги, виклики та можливості. <b>Горбачов О.С. (Донбаська державна машинобудівна академія)</b>   | 334 |
| 9. Огляд метода знаходження оптимальної розкладки клавіатури за допомогою генеративного алгоритму штучного інтелекту (гаші). <b>Горільський Е.О., Шаповалова Н. Н. (Криворізький національний університет)</b>   | 335 |

responsible for learning the latent representations of both masked and unmasked patches. The decoder then maps the complete latent representation back into the feature space, ultimately generating synthetic time series data that closely resembles the original data. The key innovation of ExtraMAE lies in its self-supervised learning approach. Unlike unsupervised methods such as GANs, ExtraMAE directly learns the temporal dynamics of the original time series data through the extrapolator. This method allows ExtraMAE to better capture the underlying structure and dependencies of the original data, which ultimately leads to higher fidelity synthetic time series.

**Benefits of ExtraMAE.** One of the most significant advantages of ExtraMAE is its ability to generate synthetic data that maintains the temporal dynamics of the original data. This high-fidelity data can be used as a substitute for real-world data in various applications, such as finance, healthcare, and environmental monitoring. By utilizing synthetic data generated by ExtraMAE, researchers, and practitioners can overcome data scarcity issues and develop more effective models in their respective domains. ExtraMAE's self-supervised learning approach allows it to outperform state-of-the-art benchmarks in time series generation. Its lightweight architecture ensures that the model is both fast and scalable, making it suitable for a wide range of applications. Additionally, ExtraMAE can be easily adapted for various downstream tasks, such as time series classification, prediction, and imputation[3]. Its versatility and performance make ExtraMAE a powerful tool for synthetic time series generation.

**Conclusion.** The ExtraMAE model provides a promising solution for synthetic time series generation, effectively addressing the lack of data in various domains. By capturing temporal dynamics efficiently and excelling in downstream tasks, it offers a valuable tool for researchers and practitioners working with time series data. ExtraMAE's success in this field could potentially usher in a new era of self-supervised time series generation.

#### **References.**

- [1] Samuel A. Assefa. Generating synthetic data in finance: Opportunities, challenges and pitfalls. InfoSciRN: Data Protection (Topic), 2020.
- [2] Chenguang Fang and Chen Wang. Time series data imputation: A survey on deep learning approaches. arXiv preprint arXiv:2011.11347, 2020.
- [3] Zhe Li, Pengyun Wang, Zhongwen Rao, Lujia Pan, Zenglin Xu. Masked Autoencoder with Extrapolator (ExtraMAE): A Scalable Self-Supervised Model for Time Series Generation. 2022.
- [4] Ian Goodfellow, et al. Generative adversarial nets. Advances in neural information processing systems, 27, 2014.
- [5] Jinsung Yoon, Daniel Jarrett, and Mihaela van der Schaar. 2019. Time-series Generative Adversarial Networks. In NeurIPS.

УДК 004.89

### **СИНТЕТИЧНІ НАБОРИ ДАНИХ В ШТУЧНОМУ ІНТЕЛЕКТІ**

**АНТОНОВА А.Р., ЮРЧЕНКО І.С.** (allaantonova62@gmail.com)

Одеський національний технологічний університет

*Авторами здійснено аналіз тлумачення поняття «синтетичні набори даних», існуючих шляхів моделювання і генерації даних, та проблем, які такі комп'ютерні програми вирішують.*

**Постановка проблеми.** Синтетичні набори даних в штучному інтелекті - це набори даних, створені шляхом моделювання та генерації даних, що є штучно створеними екземплярами реальних даних. Ці дані можуть бути використані для навчання та тестування моделей машинного навчання та інших алгоритмів штучного інтелекту.

**Виклад основного матеріалу.** Існує декілька методів побудови синтетичних наборів даних у ШІ. Існує декілька підходів, таких як:

- Генеративні моделі - це моделі, які можуть генерувати нові дані на основі вхідних даних. Наприклад, генеративні відбивачі (GAN) - це нейронні мережі, які навчаються генерувати нові дані, що нагадують реальні дані. Наприклад GAN можуть бути використані для навчання моделей машинного навчання на задачах класифікації, детекції об'єктів тощо.
- Бутстреп - це метод випадкового дублювання даних для створення нових наборів даних. Цей метод може бути використаний у фінансовій аналітиці для створення синтетичних даних, які можуть відображати різні сценарії розвитку фінансового ринку. Це може допомогти аналітикам приймати рішення щодо інвестування грошей на основі вірогідних прогнозів
- Симуляції - це метод, що використовується для створення синтетичних даних, які відображають реальний світ за допомогою комп'ютерних моделей. Цей метод може бути використаний у сфері автоперевезень для створення синтетичних даних, які можуть відображати різні дорожні сценарії, що допоможе розробникам автомобільних систем автоматичного керування навчати моделі керувати автомобілями в різних умовах.
- Аугментація даних - це метод, який дозволяє створювати нові набори даних, додавши до існуючих даних деякі зміни, наприклад, зміна контрастності, розміру або нахилу зображення. Цей метод може бути використаний у галузі медичної діагностики, де недостатня кількість даних для навчання моделей може ускладнити їхнє навчання. Застосування аугментації даних може допомогти краще навчити моделі розпізнавати патології на медичних зображеннях.

При побудові моделей для створення синтетичних даних для навчання штучного інтелекту можна використовувати наступні принципи:

- Розуміння цілей: необхідно розуміти, які дані потрібні для навчання моделі, щоб дані, що згенерували, відповідали цілям навчання.
- Використання реалістичних даних: дані, що генеруються, повинні бути достатньо реалістичними і схожими на реальні дані, щоб модель могла ефективно навчатися на них.
- Врахування статистичних властивостей: синтетичні дані повинні мати схожі статистичні властивості з реальними даними, щоб модель могла узагальнювати нові дані.
- Використання різних алгоритмів: можна використовувати різні алгоритми для генерації даних, як-от генеративні моделі, нейромережеві моделі, статистичні моделі тощо.
- Оцінка якості: необхідно оцінювати якість згенерованих даних за допомогою різних метрик, щоб переконатися в їхній придатності для навчання моделі.
- Використання контрольної вибірки: можна використовувати контрольну вибірку, щоб порівнювати результати навчання моделі на згенерованих даних із результатами навчання на реальних даних.
- Облік конфіденційності: при генерації синтетичних даних необхідно враховувати конфіденційність та захищеність даних, щоб уникнути витоків конфіденційної інформації.
- Використання різноманітності: можна використовувати різні параметри для генерації синтетичних даних, щоб збільшити їх різноманітність та збагатити навчальний набір даних.

Основна причина використання синтетичних наборів даних полягає у вирішенні проблеми недостатньої кількості реальних даних для навчання моделей машинного навчання. Наприклад, у галузі медицини, де доступні обмежені медичні дані, можуть бути створені синтетичні дані для покращення розпізнавання та діагностики різних хвороб.

Ще одним прикладом використання синтетичних наборів даних може бути в області автомобільної промисловості, де можуть бути створені синтетичні набори даних для навчання автономних автомобілів на різних типах доріг та в різних погодних умовах.

Оцінка якості синтетичних наборів даних також є важливим аспектом іскусственного інтелекту. Один із методів оцінки полягає в порівнянні статистичних характеристик синтетичних даних з реальними даними. Також може використовуватись оцінка

ефективності моделі машинного навчання, навченої на синтетичних даних, на тестових даних, отриманих з реальних даних.

Методи оцінювання синтетичних наборів даних можуть включати такі методи, як оцінювання розподілів даних, порівняння статистичних метрик (таких як середнє значення, медіана, дисперсія тощо) та візуальний аналіз синтетичних даних порівняно з реальними даними.

Оцінювання розподілів даних - це процес порівняння розподілу синтетичних даних з розподілом реальних даних для того, щоб визначити, наскільки точно синтетичні дані відображають реальні дані. Це важливий етап в розробці моделей машинного навчання, оскільки точність моделі залежить від якості набору даних, на якому вона навчається.

Порівняння статистичних метрик також є важливим етапом в оцінюванні синтетичних даних.

- Середнє значення - це середня величина всіх значень у наборі даних.
- Медіана - це значення, що знаходиться посередині впорядкованого набору даних, коли вони впорядковані за зростанням або спаданням.
- Дисперсія - це міра того, наскільки великі відхилення значень від середнього значення.

Візуальний аналіз синтетичних даних порівняно з реальними даними передбачає використання різних графічних методів, таких як діаграми розсіювання, гістограми, boxplot тощо, для того, щоб порівняти розподіл синтетичних даних з реальними даними. Графіки можуть допомогти виявити різниці між даними та розподілами та дати можливість досліджувати взаємозв'язки між різними змінними.

Загалом створення синтетичних наборів даних у штучному інтелекті - це важливий інструмент, який може допомогти розробникам штучного інтелекту покращити ефективність своїх моделей та алгоритмів за відсутності достатньої кількості реальних даних.

### **СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:**

1. <https://www.syntho.ai/uk/what-is-synthetic-data/>
2. <https://www.techtarget.com/searchcio/definition/synthetic-data>
3. <https://gretel.ai/blog/how-to-generate-synthetic-data-tools-and-techniques-to-create-interchangeable-datasets>

**УДК 004.89:004.056.5(477)(043.2)**

### **ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У 3D-МОДЕЛЮВАННІ**

**БОЙЦОВА М.П., БОЙЦОВА О.С.**

Одеський національний технологічний університет

Використання штучного інтелекту під час творчого процесу є потужним інструментом, однак слід пам'ятати, що моделювання - складний багатоетапний процес, тому формулювання правильних запитів може стати важким завданням. Зараз одною з найпопулярніших нейромереж в Україні є Chat GPT, однак він може надавати відповіді лише у форматі текстових команд, однак у 3D-моделюванні необхідно оперувати з зображеннями із текстурами та, власне, й самими моделями.

Щоб вирішити цю проблему, потрібна інша неймережа, яка виступатиме посередником між вами та Chat GPT і буде виконувати ваші запити в обраній програмі для моделювання чи у просторі роботи зі штучним інтелектом, надаючи вам файли із 3D-моделями. OpenAI, на базі якої працює Chat GPT, пропонує зв'язану з ним неймережу Point-E, яка створює моделі за текстовим описом.